

COMITÉ NATIONAL FRANÇAIS DE RADIOÉLECTRICITÉ SCIENTIFIQUE
UNION RADIO SCIENTIFIQUE INTERNATIONALE
SIÈGE SOCIAL : ACADEMIE DES SCIENCES, 23 QUAI DE CONTI, PARIS 6^{ME}

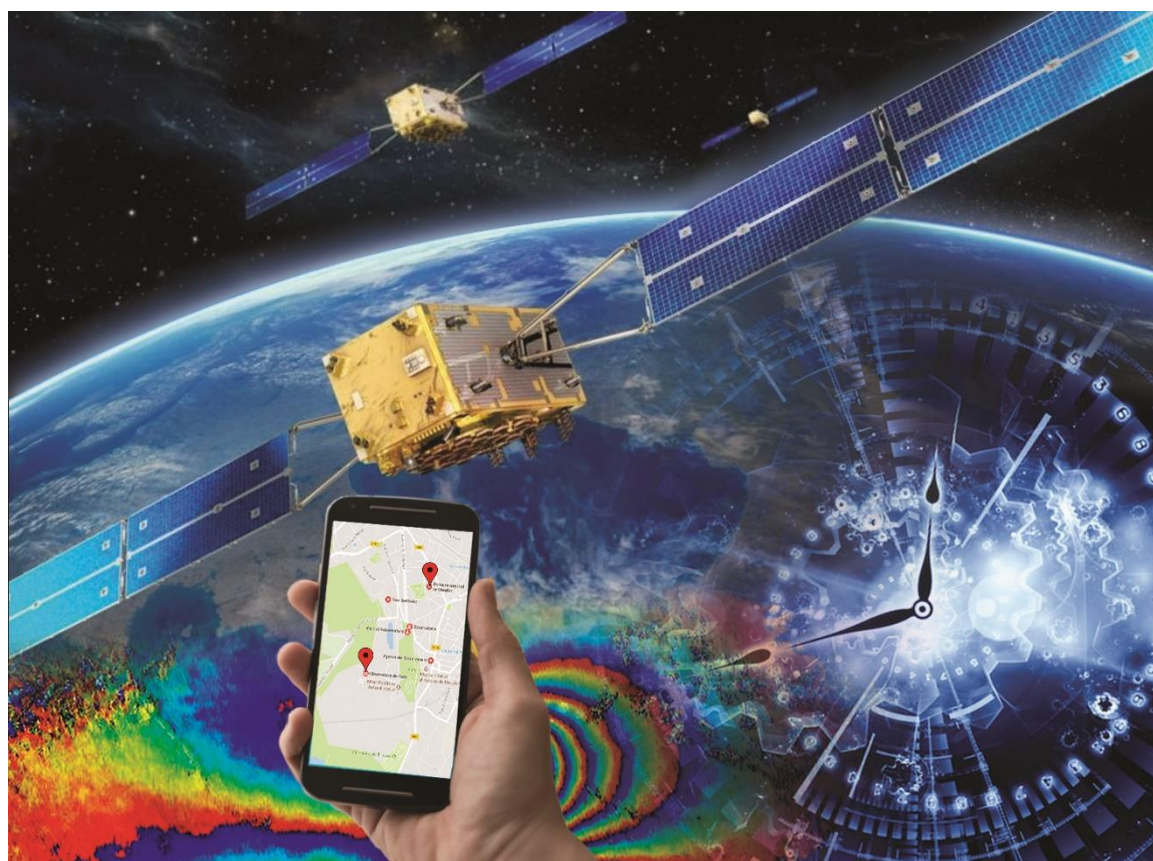


JOURNÉES SCIENTIFIQUES, *WORKSHOP*

GÉOLOCALISATION ET NAVIGATION
DANS L'ESPACE ET LE TEMPS
***GEOLOCATION AND NAVIGATION
IN SPACE AND TIME***

28 / 29 MARS, 2018

OBSERVATOIRE DE PARIS, SITE DE MEUDON



Actes Proceedings

Crédit illustration de couverture : Atelier Isatis – Dijon
Cover image Credit: Atelier Isatis - Dijon

Table of contents

Influence de l'environnement sur la géolocalisation et la navigation / Environmental effects on geolocation and navigation	1
Conf. invit. : L'inconnue ionosphère dans l'équation de la géolocalisation : problématique de la réduction, Blelly Pierre-Louis	2
Electromagnetic Wave Propagation in Precipitation Media and its Applications in: Radar Remote Sensing, Earth-Satellite Links, and Terrestrial Telecommunication, Chandra Madhu	6
Support Vector Machine Hydrometeor Classification for dual-polarization radar: application to avionic systems and meteorology, Roberto Nicoletta [et al.]	8
Solar radio emission as a disturbance of aeronautical radionavigation, Klein Karl-Ludwig [et al.]	15
Radars et antennes tout numérique / Radars and full-digital antennas	18
Conf. invit. : Benefits of space-time diversity for radar, Le Chevalier Francois [et al.]	19
Potentialités des antennes à formateur de faisceaux réduit pour applications spatiales en bande Ku, Milbrandt Jorick [et al.]	28
Système radar FMCW pour l'identification des transpondeurs, Bouhlel Nizar [et al.]	36
Amélioration de la résolution angulaire d'un radar FMCW panoramique, Fouettilloux Romain [et al.]	44

GNSS	48
Conf. invit. : GNSS : une révolution pour le géopositionnement précis, Perosanz Felix	49
Conf. invit. : Positionnement GNSS statique et cinématique appliqué à l'étude des déformations tectoniques de la Terre, Briole Pierre	54
Compact Antennas in Cavities for GNSS Applications, García Gámez Laura [et al.]	59
les destins liés des théories et des systèmes radioélectriques, Fuerxer Pierre	62
Investigations on TLE time series accuracy for GNSS satellites, over short and long time scales, Deleflie Florent [et al.]	67
Filtrage de Kalman et estimateurs / Kalman filter and estimators	68
Conf. invit. : Filtrage de Kalman invariant pour la navigation inertielle hybridée GPS, Bonnabel Silvére	69
Invariant Extended Kalman Filter for target tracking, Pilté Marion [et al.]	73
Time-Lapse Estimation for Optical Telescope Sequences, Campbell Mark [et al.]	81
Géolocalisation et navigation en milieu difficile / Geolocation and navigation in difficult environment	89
Conf. invit. : Super-resolution Sparse Channel Estimation for Localization – Prospects and Issues, Fleury Bernard	90
Uncertainty Estimation in AoA based Localization using PCE, Van Der Vorst Thomas [et al.]	95
A semi parametric model for RSSI-based localization, Elgui Kevin [et al.]	102
Le positionnement en milieux contraints : signaux d'opportunité ou infrastructure spécifique ?, Samama Nel [et al.]	107
Conf. invit. : Avancées et défis de la localisation véhiculaire coopérative, Denis Benoît [et al.]	115
Géolocalisation et navigation à l'intérieur des bâtiments, Barbot Jean-Pierre . . .	120

GNSS Integrity Enhancement for urban Transport Applications by Error Characterization and Fault Detection and Exclusion (FDE), Zhu Ni [et al.]	128
Solution d'intégration de COTS offrant un service de navigation piéton collaboratif fiable, résilient et sécurisé, Heurquier Dominique [et al.]	139
Système autonome orienté données : Du signal à l'information, Tanzi Tullio [et al.]	145
Synchronisation et calibration entre un lidar 3D et une centrale inertielle pour la localisation précise d'un véhicule autonome, Merriaux Pierre [et al.]	147
Géolocalisation HF et ionosphère / HF Geolocation and ionosphere	157
Passive HF Geolocation Using TDoA Based Receiver Network, Jain Ankit [et al.]	158
High frequencies wave propagation in the Earth ionosphere, Foucault Etienne [et al.]	164
Métrologie des horloges / Metrology aspects of clocks	173
Conf. invit. : Horloges atomiques - Etat de l'art et enjeux, Dimarcq Noel	174
Galileo gravitational redshift test with eccentric satellites, Delva Pacôme [et al.] .	178
Hardware Implementation of Quad-channel Clock Synchronization System for Nançay Radioheliograph A/D Converter, Ait Mansour El Houssain [et al.]	179
Transfert de temps par lien laser spatial, Belli Alexandre [et al.]	188
Author Index	189

**Influence de l'environnement sur la
géolocalisation et la navigation /
Environmental effects on geolocation
and navigation**

**L'inconnue Ionosphère dans l'équation de la géolocalisation :
problématique de la réduction**
The unknown Ionosphere in positioning equation: how to solve it

Pierre-Louis BLELLY¹

¹ *Institut de Recherche en Astrophysique et Planétologie (IRAP),
pierre-louis.blelly@irap.omp.eu*

Mots clés (en français et en anglais) : Ionosphère, modélisation numérique, assimilation de données.


Présentation en session plénière / Plenary session communication

Résumé / Abstract

Le positionnement GPS a été très tôt confronté au problème constitué par les couches ionosphériques et l'a contourné en utilisant soit un système bi-fréquence, soit un système monofréquence et un modèle basique de l'ionosphère, afin de déterminer le TEC associé à l'ionosphère. Cette approche a permis de développer des systèmes performants qui, de fait, souffrent du défaut originel, à savoir une faible connaissance de l'ionosphère, qui peut être circonvenue pour les conditions « normales » d'évolution de l'ionosphère. Or, l'ionosphère présente très souvent un comportement « anormal » lié aux couplages entre le Soleil et l'environnement terrestre. En conséquence, quelle que soient leurs performances, ces systèmes ont un plafond de verre concernant leur niveau de disponibilité, qu'il ne sera pas possible de franchir tant que l'inconnue ionosphère ne sera pas correctement prise en compte dans l'équation GPS, que ce soit en terme de caractérisation du milieu ou en terme de prévision d'évolution du milieu. L'objet de cette présentation est donc en premier lieu de sensibiliser la communauté du positionnement à la problématique ionosphérique, liée fondamentalement au caractère plasma du milieu, placé dans un environnement fortement magnétisé et réactif (chimiquement), et qui est marquée par une forte dynamique temporelle et spatiale. Ensuite, nous montrerons que le comportement, parfois chaotique, de ce milieu ionisé peut néanmoins être assez bien appréhendé et que les efforts de modélisation de cette dynamique menés par la communauté permettent de mieux caractériser la structure ionosphérique observée, en fournissant des paramètres clés pour l'expliquer. Enfin, le développement d'outils de modélisation contrôlés par ces paramètres, eux-mêmes contraints par les mesures temps réel actuellement disponibles, constitue une perspective pertinente pour une prise en compte adaptée de l'ionosphère, qui permettra de lever cette inconnue ionosphérique dans l'équation GPS.


The unknown “Ionosphere” in positioning equation: how to solve it

Pierre-Louis BLELLY
IRAP




OUTLINE

- Structure of the ionosphere
- Dynamics
- Couplings in the MIT system
- Connection with scintillations
- Conclusion



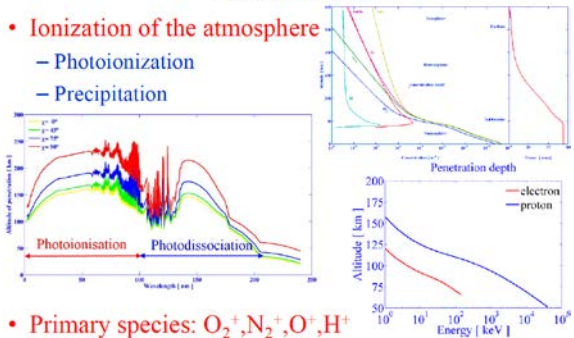

Structure

- Production
- Chemistry
- Vertical structure



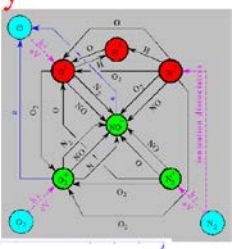
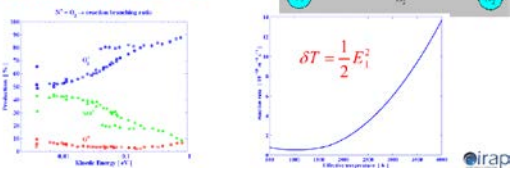

Production

- Ionization of the atmosphere
 - Photoionization
 - Precipitation
- Primary species: O_2^+, N_2^+, O^+, H^+

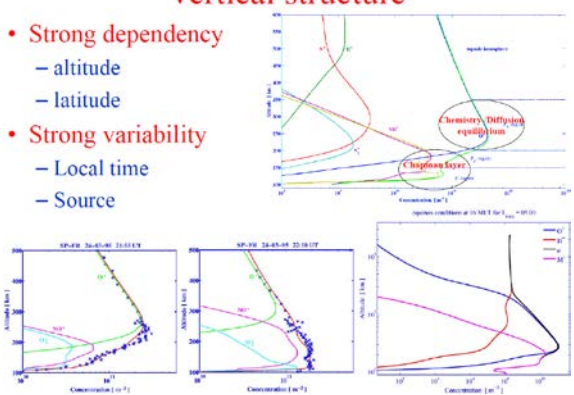

Chemistry

- Complex chemistry
- Highly variable with energy
 - Reaction rates
 - Products
- Alteration of initial structure
- Secondary species : NO^+


Vertical structure

- Strong dependency
 - altitude
 - latitude
- Strong variability
 - Local time
 - Source

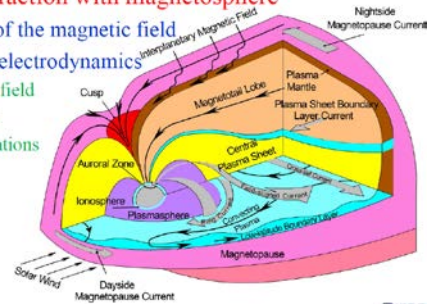

Dynamics

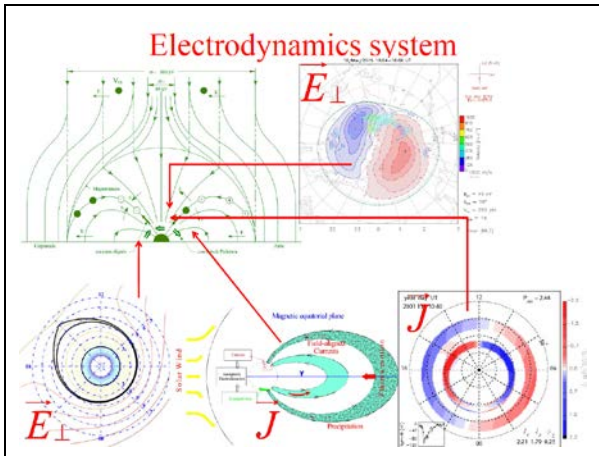
- Magnetospheric system
- Electrodynamics system
- Impact on the ionosphere



Dynamics

- Control by the solar wind
- Strong interaction with magnetosphere
 - Key role of the magnetic field
 - Complex electrodynamic
 - Electric field
 - Current
 - precipitations

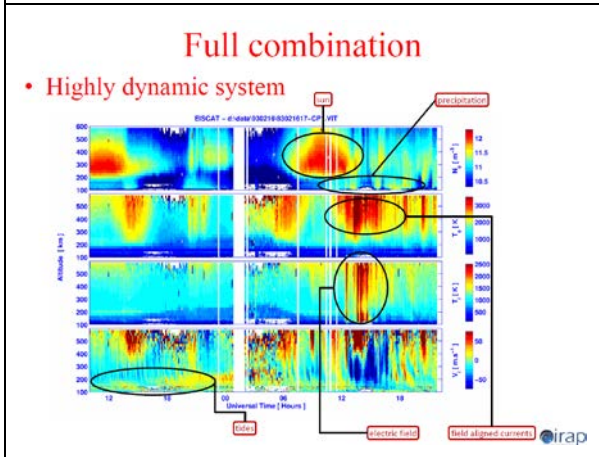





Impact at ionospheric level

- Ionospheric plasma convection
 - Drag on atmosphere
 - Frictional heating
 - Enhancement of the chemistry
- Splitted motion
 - 1D along magnetic field B
 - 2D perpendicularly to B
- Field aligned transport

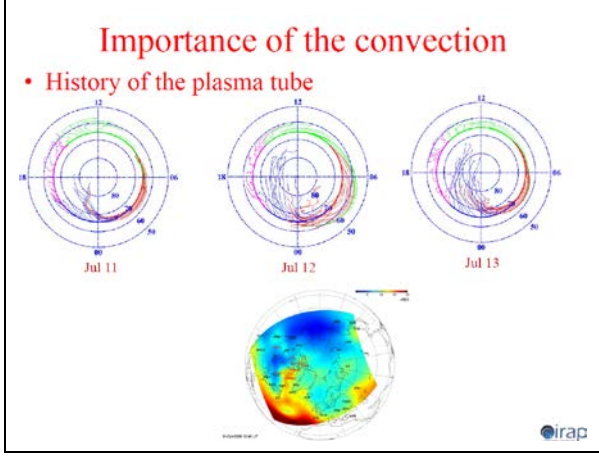
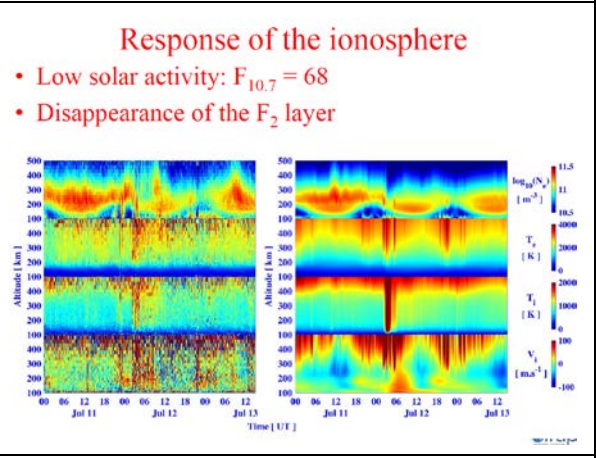
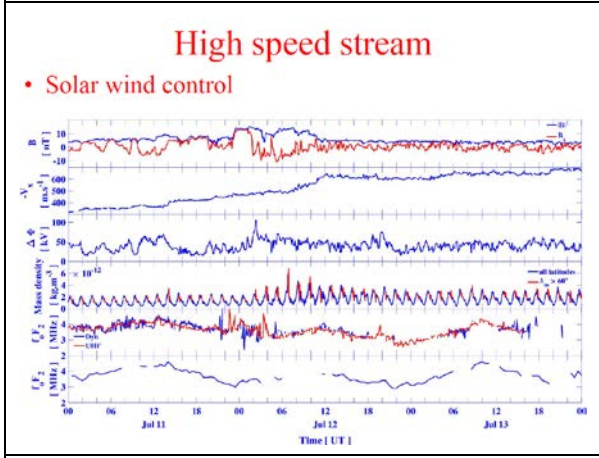
The diagram shows the impact of magnetospheric convection on the ionosphere. It includes a polar plot of the ionosphere and a vector diagram showing the perpendicular velocity (V_{\perp}) and field-aligned velocity (V_{\parallel}). Labels include 'ionospheric plasma convection', 'V_perp', 'V_parallel', and 'field aligned transport'.



Couplings in MIT

- High speed stream
- Chemistry
- Vertical structure

The diagram illustrates couplings in the magnetosphere-ionosphere system. It shows the interaction between the magnetosphere and the ionosphere, highlighting the high-speed stream, chemistry, and vertical structure. Labels include 'couplings in MIT' and 'irap'.



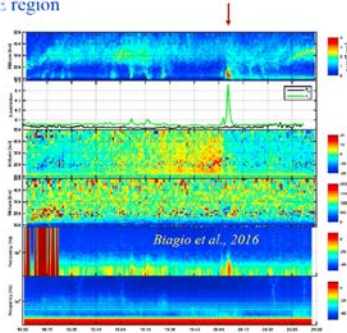
Connection with scintillations

- Ionospheric structure
- Large scale irregularities
- Magnetospheric activity

The diagram illustrates the connection between magnetospheric activity and ionospheric scintillations. It shows how magnetospheric activity leads to large-scale irregularities in the ionosphere, which cause scintillations. Labels include 'Connection with scintillations', 'Ionospheric structure', 'Large scale irregularities', 'Magnetospheric activity', and 'irap'.

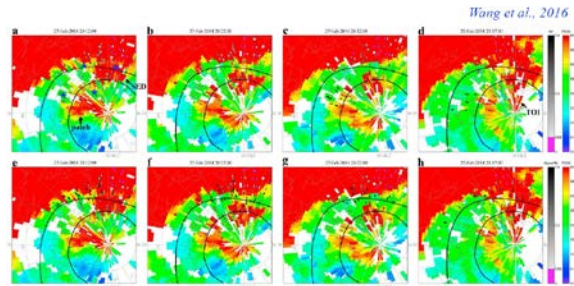
Ionospheric structure

- Large scale ionization structures
 - Enhancement of the E region
 - Substorm onset
 - Vertical extension
 - No cascade to smaller
 - Phase without amplitude
 - Recombination rate
 - Plasma instability



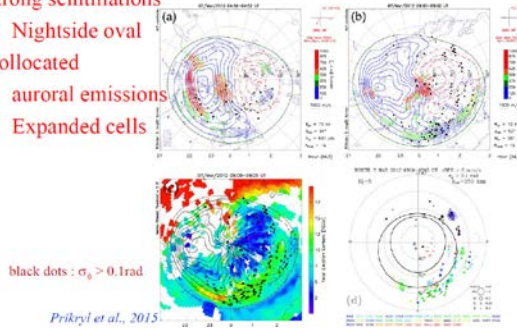
Large scale irregularities

- Effect of the convection
 - Plasma patches over the polar cap



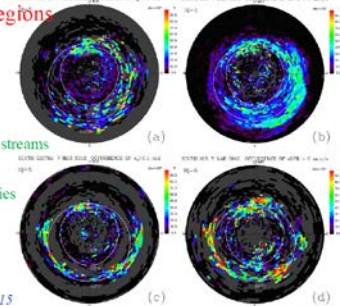
Magnetospheric activity

- Geomagnetic storm
- Strong scintillations
 - Nightside oval
- Collocated
 - auroral emissions
 - Expanded cells



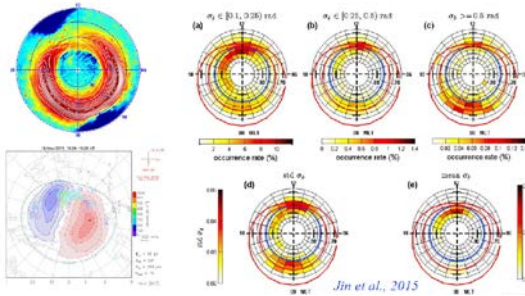
Magnetospheric activity

- Impact in both hemisphere
- Collocated with auroral oval
- Enhanced in specific regions
 - Cusp
 - TOI
 - Tongue of ionization
 - SAPs
 - Subauroral polarization streams
 - SEDs
 - Storms enhanced densities
 - AURO
 - Auroral oval



Auroral oval

- Scintillation occurrence level
 - Mainly noon and midnight sectors



Conclusion

- Highly dynamical system
 - Strong couplings
 - Electrodynamics
 - Chemistry
 - Precipitation
- Scintillations
 - Plasma instabilities
 - Magnetic field
 - electro-dynamics
 - Ionospheric irregularities
 - Convection
 - plasma transport over large distances
 - Precipitation
 - Cusp
 - Density enhancement
- Strong modelling effort
 - Better understanding of couplings intrication
 - Provides global conditions for scintillations
 - Ability to simulate some irregularities
- Efforts to combine scintillation measurement and magnetospheric activity
- Extension towards equatorial region
 - Same medium
 - Different magnetic and electro-dynamics configuration
 - Other instabilities



Electromagnetic Wave Propagation in Precipitation Media and its Applications in: Radar Remote Sensing, Earth-Satellite Links, and Terrestrial Telecommunication

Propagation des ondes EM et précipitations : Applications en télédétection Radar, Satellite et télécommunication

Madhu Chandra¹

¹ Chemnitz University of Technology, Chemnitz, Germany,
madhu.chandra@etit.tu-chemnitz.de

*Keywords (in English and French): rain attenuation, phase-shift, signal degradation, polarisation dependence.
Atténuation de la pluie, décalage de phase, dégradation du signal, dépendance de la polarisation*

Résumé / Abstract

If the propagation of electromagnetic waves in ‘free-space’ is taken as a standard reference, the corresponding wave propagation through precipitation filled media can experience several physical effects: wave attenuation, phase shifts exceeding the ones due to free-space propagation delay, change of polarisation state (so-called wave depolarisation), dispersion over the signal bandwidth, scintillations, and the addition of radiometric noise. These effects, in turn, are capable of significantly impairing the quality, strength, and even the character of information carrying radio signals. As one might expect, the named propagation effects are strongly dependent on the frequency of the EM-waves and the nature (hydrometeor type, particle size distribution) of the precipitation medium.

In view of the fact that link-budgeting of wireless signals is deemed to include the influence of these propagation effects and their resulting impairment, the assessment of such effects, and the constraints they set, has become a challenging issue for propagation scientists and engineers. The need to estimate or account for propagation effects in precipitation media arises in several scenarios: terrestrial microwave links, earth-satellite links, active and passive remote sensing, and, contrary to popular opinion, even in mobile communications. The inclusion of higher frequencies, well beyond 10 GHz, and higher bandwidths well into the GHz region has only exacerbated the problem. At this juncture, it is perhaps instructive to ask which types of precipitations may generate propagation effects more significantly than others. Generally, wet precipitation media, i.e. the media containing wet hydrometeors such as raindrops or wet sleet or wet hail, contribute to propagation effects far more strongly than their corresponding dry counterparts such as snowflakes, dry hail and powdery snow. As it turns out, the modelling and simulation of the various propagation effects arising from wave propagation in diverse media is a daunting task that demands an interdisciplinary mix of atmospheric physics, EM-scattering, systems engineering, and propagation science. In view of the broadness of this field and its inherent complexity, the present state-of-the-art offers maximum insight into the propagation effects due to one of the most frequent contributors: the rain medium! Fortunately (less so when you are holidaying!), rain is also one of the most prevalent wet precipitations across the globe and, therefore, for practical purposes, the most relevant medium when it comes to the estimation and assessment of propagation effects in wireless communications, remote sensing, and GNSS applications. A review of the propagation properties of rain and the signal distortions it generates will therefore constitute the main theme of this contribution. Quantitative and qualitative features of such rain induced impairments will be illustrated using real and simulated data. Also, the most established physical models of these propagation effects will be critically reviewed against the future needs.

Briefly, the main features to be addressed in this spirit will be: signal attenuation, and precipitation-induced-phase-shifts. Both features, which underlie the various named propagation effects, will be considered in dependence of polarisation and frequency. At this stage, the reader might ask: what are the physical mechanisms that lead to such effects? The origin of signal attenuation may be found in the losses incurred by the ‘travelling wave’ due to absorption and scattering by raindrops. Both features can be physically accounted for by the imaginary part of the complex forward scattering amplitude of raindrops (assuming $\exp(j\omega t)$ harmonic time-dependence phasors). Similarly, the origin of phase shifts may be found in the real part of the complex forward scattering amplitude that ultimately leads to an increase in the refractive index of the rain medium. The fact that the complex forward scattering amplitude depends on the electrical size (measured with respect to the effective wavelength) and the dielectric constant of the raindrop accounts for the strong frequency dependence of the propagation effects in rain. Now it remains to be seen why the rain media is capable of generating polarisation dependent attenuation, phase shifts and depolarisation. The root cause of these unique features lies in the scattering anisotropy of raindrops that have non-spherical shapes. Contrary to the popular belief that raindrops have a ‘tear drop’ shape, the true shape of raindrops is more like oblate spheroids with their minor axis parallel to the gravity vector. This shape anisotropy and overall alignment of rain hydrometeors leads to different values of refractive indices along the horizontal and vertical polarisations. It is worth noting that this anisotropy of rain media only came to light with the advent of microwave links and weather radars! A noteworthy application of this polarisation dependence is the assessment of co-to-cross-polar-discrimination in polarisation diversity systems.

At this stage, the reader might ask: what are the typical magnitudes of such propagation effects in rain? Not unexpectedly, the magnitudes of propagation effects depend not only on the frequency and the complex dielectric constant of the raindrops but also on the so-called raindrop-size-distribution function. This elusive quantity poses a major challenge in the field of radar and radio meteorology. Keeping these factors in mind, qualitatively it may be stated that rain attenuation is capable of leading to a total fall-out of earth-satellite and terrestrial links. In remote sensing, however, an entire sensor-data-take, for example, a SAR image may be rendered unusable. A detailed quantitative treatment of these effects will be a subject of the presentation and the full paper. Nevertheless, in this regard, we can afford the reader a brief glimpse on the matter with the help of the following table.

Rainrate [mm/hr]	Specific Attenuation [dB/km]		
	L-Band	C-Band	X-Band
2.5	< 0.001	< 0.01	0.05
0.5	< 0.001	0.05	0.3
20	< 0.01	0.2	2.0
150	0.1	1.0	6.0

Table 1: Rain Attenuation (dB/km) in dependence of frequency bands and rainrate

A detailed discussion will acquaint the reader with the corresponding properties at higher frequencies. Also, examples of phase shifts, attenuation and polarisation dependent behaviour of wave propagation in rain will be illustrated with measurements made by weather radars. This feature will also bring into fore the nearly forgotten application of weather radars, namely, their function as a versatile tool for obtaining and estimating propagation effects in precipitation media of all kinds, not just rain. Finally, the issue of multiple scattering will be raised and its importance in generating radiometric noise and cross-polar power will be elucidated.

References

1. J.A Allnutt, ‘Satellite-to-Ground radio-wave propagation’, IEE Publications, ISBN 0-86341-157-6,1989.
2. ITU-618, International Telecommunication Union, ITU-R Recommendation P.618-5, Propagation data and prediction methods required for the design of earth-space telecommunication systems, Geneva, 1997b.
3. T. Oguchi, Electromagnetic wave propagation and scattering in rain and other hydrometeors, IEEE Proc., 71: 1029-1078, 1983.
4. R. L. Olsen, A review of theories of coherent radio-wave propagation through precipitation media of randomly oriented scatterers, and the role of multiple scattering, Radio Science, 17: 913-928, 1982.

Support Vector Machine Hydrometeor Classification for dual-polarization radar: application to avionic systems and meteorology

Support Vector Machine Hydrometeor Classification pour radar à double polarisation : application aux systèmes avioniques et à la météorologie

N. Roberto¹, L. Baldini^{1,2}, E. Adirosi^{1,2}, L. Facheris³, F. Cuccoli², A. Lupidi^{2,4}, and A. Garzelli^{2,5}

¹ CNR-Institute of Atmospheric Sciences and Climate, Rome, Italy, {n.roberto, l.baldini, e.adirosi}@isac.cnr.it

²National Laboratory of Radar Surveillance System (RaSS), Interuniversity Consortium for Telecommunication (CNIT), Pisa.

³Department of Information Engineering, University of Florence, Florence, Italy, {luca.facheris, fabrizio.cuccoli}@unifi.it

⁴Department of Information Engineering, University of Pisa, Pisa 56122, Italy, a.lupidi@iet.unipi.it

⁵Department of Information Engineering, University of Siena, Siena 53100, Italy, andrea.garzelli@unisi.it

Keywords: weather radar, hydrometeor classification, dual polarization, avionic radar; *radar météorologique, classification des hydrométéores, double polarisation, radar avionique*

Abstract

During the last few decades, most of the worldwide weather radar infrastructure has been upgraded to dual-polarization. Plans to demonstrate the utility of such technology for civil aircraft radars were undertaken by the EU Clean Sky Joint Technology Initiative (www.cleansky.eu). Dual polarization offers appealing advantages for avionic applications, such as the ability to correct X-band attenuation due precipitation and the possibility of implementing automated Hydrometeor Classification Algorithms (HCA). HCA can be useful to detect the presence of dangerous weather conditions related to the presence of hydrometeors like hail or graupel, typically associated with convection. Most of implemented HCAs rely on Fuzzy Logic (FL) methods. Supervised learning models based on a Support Vector Machine (SVM) are widely used for classifying remote sensing imageries and offer advantages in terms of computing time. The major problem of the SVM HCA is performing the learning phase. This process has been implemented off-line using results obtained by a FL classifier. Effectiveness of the SVM HCA has been tested by using simulated scenarios relative to an intense convective event that occurred on 15 October 2012 in the Southern Mediterranean and with real data collected by C-band dual-polarization ground-based radars.

Résumé

Ces dernières décennies, à l'échelle mondiale, la plupart des radars météorologiques ont été équipés de double polarisation. Plusieurs études ont été entreprises dans le cadre de l'initiative EU clean Sky Joint Technology (www.cleansky.eu) afin d'illustrer l'utilité de cette technologie pour les radars des avions de ligne. La double polarisation apporte des avantages intéressants pour les applications avioniques, tels que la capacité de corriger de l'atténuation liée aux précipitations pour la bande X ainsi que la possibilité d'implémenter des algorithmes automatisés de classification des hydrométéores (HCA). Les HCAs peuvent être utiles pour détecter la présence de conditions météorologiques dangereuses liées à la présence d'hydrométéores tels que la grêle ou les graupels, ceux-ci étant habituellement associés à la convection. La plupart des HCAs utilisent des méthodes de logique floue (FL). Les techniques d'apprentissages supervisées basées sur l'utilisation d'un Support Vector Machine (SVM) sont largement utilisées pour la classification des images de télédétection et sont très performantes en terme de temps de calcul. Le principal problème des HCA utilisant un SVM est la mise en forme de la base de données d'apprentissage. Cette étape a été implémentée à part en utilisant la classification issue d'un algorithme de FL. Les performances du SVM-HCA ont été évaluées au moyen de simulations d'un événement de convection intense qui s'est déroulé le 15 octobre 2012 dans le sud de la Méditerranée et avec des observations issues de radars sols en bande C équipés de double polarisation.

1 Introduction

Within the 7th framework programme of the European Commission, several projects of the Clean Sky Joint Technology Initiative (JTI) aimed at improving the weather radars that are used on board the civil aircrafts to detect (and avoid) adverse weather along the flight. State-of-the-art systems commercially available are pulsed single polarization X-bands radars that estimate the effective radar reflectivity factor, which is displayed in a few

colours coded according to specific standards and suggest proper actions: areas in red must be avoided. In addition, some radars have Doppler capabilities and detect the turbulence associated to the motion of cloud particles within a certain distance. Dual polarization that nowadays can be considered as the technology standard for ground-based operational and research weather radars (during the last few decades, most of worldwide weather radar infrastructure has been upgraded to dual-polarization.), has not used yet for aircraft systems. However, dual-polarization has interesting characteristics that can make it appealing for civil aircraft applications ([1][2]). First, dual-polarization can enable quite robust methods for compensating the X-band attenuation caused by propagation through precipitation. Attenuation can lead to underestimate the severity of some convective cell and the associated risk when the measured reflectivity is lowered by attenuation. Second, the possibility of implementing automated hydrometeor classification procedures to detect the presence of dangerous weather conditions such as those characterized by specific hydrometeors associated to the insurgence or to the presence of convection.

This paper focuses on this second aspect. Among the meteorological products and applications based on dual-polarization radar measurements, the identification of hydrometeor type has become one of the most successful. Different Hydrometeor Classification Algorithms (HCA) have been proposed along the years, but almost all of them are based on empirical approaches allowing to deal with approximate knowledge of the actual relationships between radar measurements and microphysics of precipitation within a resolution volume, such as the fuzzy logic (FL) methods. References [3] and [4] are early papers on HCA, while [5]-[7] are more recent studies aiming at operational implementations in operational radars at different bands. Such classifiers achieve hydrometeor classification by analyzing the set of dual-polarization measurements pertinent to a single radar measurement sampling. Attempts to improve the robustness of the fuzzy logic approach exploit measurements in the neighbouring bins to enhance the spatial coherency of classification maps [8]. A similar approach, but based on convolutional neural network deep learning approach is explained in [9]. Unsupervised clustering to define the number of classes and the membership functions has been explored [10], while a semi-supervised approach, exploiting fuzzy logic for defining the constraints of an unsupervised approach has been proposed [11]. Actually, a few HCA implementations based on supervised learning methods can be found in the literature. Neural networks have been used in [3] and [9] to adjust membership functions based on observed data. Supervised learning models based on a Support Vector Machine (SVM) [12] are widely used for classifying different remote sensing imageries (see [13] for a review), including meteorological satellite observations ([14][15]). Therefore, the SVM approach is worth to be investigated for hydrometeor classification using dual-polarization radar measurements, also for avionic applications.

An SVM classification algorithm consists of two phases: the training phase, in which the relation between input variables (i.e. radar measurements) and class labels output (i.e. hydrometeor classes) is established, and the predict phase, in which the most likely class is assigned to an input sample measurements (Fig. 1). The supervised algorithm is typically trained once and out of the classification prediction scheme. The SVM classifier usually proceeds for the two-classes problem in three stages. The first one is the basic margin classifier where separation between two classes is tried with a linear decision boundary (a hyperplane). For problems that are not separable by a linear decision, in a second step the classifier is modified to tolerate misclassification (soft margin) by introducing a penalty. At the last stage, the classifier can be nonlinearly generalized by using a kernel method. The support vector machine can be regarded as a linear classifier that finds the best linear separation between samples belonging to two classes. This binary scheme can be extended for multiclass classification by combining several binary classifiers, whereas one-step multiclass solutions are generally more expensive from a computational view. An accurate strategy used for this purpose is the one-against-all (OAA) that is used in this work. It consists in applying binary SVM to separate members of one class to members of other classes. The algorithm is based on a parallel architecture composed by N SVMs, one for each class.

In the framework of research activities developed in Clean Sky, a reference architecture has been defined that designates a decision support system (DSS) to host end-user applications using dual polarization measurements as input. Avionic devices acting as DSS on board of civil aircrafts must be certificated and must respect strict specifications (e.g., [16]) and have relatively poor performances in terms of both computational and storage capabilities. The EFB by Astronautics NEXIS Flight-Intelligence System co., has been adopted for development. Classification output can be both displayed to the pilot (see Fig. 1, right), associated to the label of hydrometeors (or to a set of new labels) or associated to a level of risk related to the presence of a given type of hydrometeor. Ultimately, the level of risk can be handle by a process that helps a pilot to choose an optimal route (Fig. 1, center and right where different display mode available to the pilot are shown) [17]. However, the HCS should be also accurate, at least not less than a standard fuzzy logic HCA.

This paper summarizes some basic concepts of SVM and the solution that has been adopted apply this approach to dual-polarization weather radars [23]. Testing with a simulated scenario allows to reveal important properties of SVM that must be taken into account in implementations.

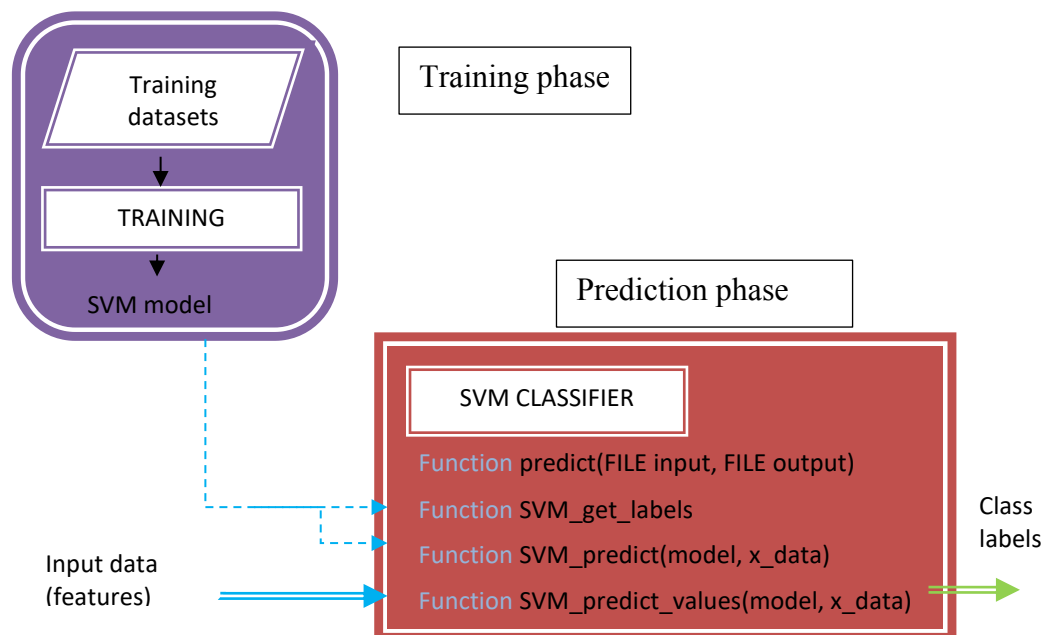


Figure 1: Scheme of implementing Support Vector Machine

Finally, testing with actual dual polarization data indicates that the SVM HCA developed and shown in this paper is generic and can be effectively used for ground-based weather radars. The useful approach used for hydrometeor classification for ground-based radar can be time consuming and could not be enough efficient for avionic weather radar. If the aircraft speed is approximately 800 km h^{-1} and that the radar antenna sweeps 120° in azimuth with 3° angular resolution in 10 s, with 150 km of maximum range and 150 m range resolution, each sweep implies 1000 range bins for about 40 azimuth lines implying labeling 1000 samples at least every 250 ms, or equivalently, at least $4000 \text{ labels s}^{-1}$. The SVM HCA proposed in [1] was found to be able to generate 7000 labels/s (running on an x86_64 Linux machine Eight core, Intel-i7@2,80GHz processor). These results highlight the SVM algorithm capability in terms of computational time and indicate it as a candidate for hydrometeors classification on an EFB machine.

2 Support Vector Machine Implementation of HCA

The SVM training phase provides the grounds of the HCA by defining the support vectors, defined through unknown linear relations between observations for which labels are known. This phase, in the case of hydrometeor classification, could involve processing of large sets of data. Since fuzzy logic is a well established method for HCA, the training phase is conducted using an FL-HCA. Fuzzy logic HCAs have as input the radar measurements, as output the hydrometeor classes and membership functions that define the degree of truth of the class of hydrometeor given some inputs. Membership functions are determined both by investigating physical scattering and propagation behavior of the different hydrometeors, by analyzing empirical measurements, or by synthesizing the knowledge available from existing studies. The FL-HCA used in this paper is in use at Institute of Atmospheric Sciences and Climate of the Italian National Research Council (ISAC-CNR) and is tailored to identify meteorological targets, with a particular emphasis on the hydrometeors involved in convective events [19]. It is not far from different schemes described in the literature and therefore it is reasonable to consider SVM results as not significantly influenced by the choice of FL classifier. Input measurements are the radar reflectivity factor (Z_h in dBZ), differential reflectivity (Z_{dr} , in dB), specific differential phase shift (K_{dp} in deg/km), copolar correlation coefficient (ρ_{hv}), and the height of the 0° isothermal. The output class set is the following: *i) Rain*: includes light, moderate, and heavy rain; in rain, Z_h varies from 10 to 60 dBZ and Z_{dr} is positive, due to the oblateness of rain drops; *ii) Dry Snow*: includes all the non-wet ice particles detectable above the 0°C isothermal, such as aggregates, plates, and columns, associated to low values of Z_h ; *iii) Wet Snow*: this class typically refers to melting ice particles that form the “melting layer” where ice particles melts below the 0° isothermal to form raindrops; *iv) Graupel*: they are iced particles detectable both above (in convection) and within the melting layer, but can reach the reach ground for downdraft; they produce $Z_{dr} \sim 0$, or even negative values for conical graupel; *v) Hail*: the presence of hailstones characterizes deep convective events and can be found from the ground up to heights well above the melting layer; hailstones have high variability in size and

shape, but typically determines high Z_h , but Z_{dr} and K_{dp} values close to 0; vi) *Hail Mix* is a class that accounts for mixture of raindrops lifted by updraft, supercooled raindrops, hail, small hail, and graupel. The typical radar signature of this class is high values of Z_h and low values of ρ_{hv} , a measurement that is sensitive to heterogeneity of scatterers within the radar sample volumes.



Figure 1: Position of EFB on the cockpit (left) (from [2]) and display of the classification results with identification of risk zones and route suggested to pilot. Two different display and zoom options can be used (center and right panels)

To perform a meaningful training phase, the training set should be representative of the physical phenomena to be classified. A large variety of cases should form the training set. Some processing steps are required to select a training set of radar measurements that are not contaminated by clutter and are compensated for biases due to attenuation or partial beam blocking. For measurements collected at high elevation angle, a compensation for polarimetric measurements is also necessary [20]. One of the input of the HCA is the level of 0° isothermal. It can be estimated also using radar, but in this study is considered as estimated by radio soundings for ground-based systems or (using some assumptions) from measurements collected by devices on-board an airplane, or from weather model. The SVM algorithm is trained on classification maps generated by a FL classification approach. For each class the SVM descriptors are extracted. Radar data are scaled into a specific range because, in general, they (and in general, input features) have values that ranges in different scales and therefore are not commensurable. For nonlinear SVM using radial basis function (RBF) kernel, the penalty and the kernel parameters, usually denoted with symbols C and γ , determining the capability of the algorithm to separate the hyperplanes, need to be estimated. The better these parameters are identified, the more accurate is the prediction of the classifier. Their identification is performed during the training phase using a cross-validation (CV) procedure.

3 Validation with a simulated datasets

The Weather Research and Forecasting model (WRF) numerical weather prediction code outputs the atmospheric quantities from which the radar measurements are simulated. The Milbrandt-Yau 2-moment microphysics option of WRF [21] is used to obtain microphysical parameters such as mixing ratio (q in Kg/Kg) and concentration of particle (N in m^{-3}) for 4 different classes (rain, snow, graupel and hail). A corresponding scenario of radar measurements is obtained following the procedure described in [22][23]. Since the hydrometeor classes provided by the weather model used are four, and the classes of the SVM HCA are six, some HCA classes are obtained mixing two WRF classes. Difference are in snow class (HCA split snow in two classes: wet snow and dry snow), and in Mix Hail class (this class is present only in HCA and it is a mixed of rain and hail).

The WRF model is run for an intense precipitation events occurred in the afternoon 15 October 2012 in South Mediterranean and off the coasts of Tunisia. The obtained scenario is composed of a cube of $800 \times 682 \times 52$ elements of $0.300 \text{ km} \times 0.300 \text{ km} \times 0.150 \text{ km}$ for which the microphysical characteristics, i.e. q and N , and meteorological parameters (pressure, temperature, wind speed and direction) are provided by WRF and by the simulator in [22], radar measurements are calculated at frequency of 9.353 GHz. An SVM algorithm has been implemented for the WRF scenario using the radar measurements and the height of 0° level given by the WRF model. The training set is obtained random sampling the elements of the WRF cube. A training set composed by 104 samples (l) was randomly sampled, while a test set of 3×106 samples (m , randomly sampled excluding the samples belong to the training set) were used to test the SVM-HCA. In order to get an estimate of the generalization performance of a model typical steps are (i) find optimal tuning parameters using cross-validation

(CV), (ii) train a model using these optimal parameters on the full training set and (iii) test this model on the test set. For the model used in this work, the RBF kernel is identified by the parameters C and γ . The better these parameters are identified, the more accurate is the classifier prediction. A CV is therefore used also to get an estimation of the generalization error when certain tuning parameters are used. A “grid-search” on C and γ using CV is computed. Various pairs of (C, γ) values are tried and the one with the best CV accuracy is picked.

It is important to estimate how large a training set should be to have an accurate model. To this purpose, a large subset of training set is obtained. It is expected that larger training subsets increase the CV accuracy. It was found that over a certain size (it is 10^4 in this case as reported in [16], the CV accuracy is stabilized. This number is considered to be the number of observations necessary to train the SVM model with an high accuracy (94.8%) and in a reasonable time processing (less than 1 minute). It is important to assure that all the classes in the training are well represented by the SV. In fact, when the training is made using random sampling the classes that have low occurrences should not be sampled enough to establish a robust relation. In order to allow a good representation of all the classes in the model the training set should be quasi-equally divided among the classes.

The validation of the classes predicted by the SVM with respect to the WRF microphysical expresses in terms of the q parameter is assessed. Several metrics can be used, namely the confusion matrix (CM) calculated between the truth and the predicted hydrometeor classes applied at the same test set. The elements $CM(i, j)$ contain the number of observations classified in i th class, which in reality belong to the j th class. The diagonal contains the same classification for both algorithms (correct classification). Given CM, the global performance of SVM is quantified by the overall accuracy (OA) and Cohen’s Kappa (K). In spite of the differences between the WRF and SVM classes definition, good performance are obtained for SVM classifications using the metric above. The two classifiers are in good agreements with an $OA= 74\%$ and $K= 0.62$. This result indicates that the SVM well predict the WRF classes via the FL reference.

4 Examples of application of SVM HCS to ground-based weather radar

The SVM model trained by FL classification presented in previous sections has been applied to real polarimetric ground based measurements collected at C-band by the Polar 55C dual-polarization radar of CNR-ISAC installed at the Rome branch of ISAC. Considered are convective events occurred in October 2012 [24] and October 2015 [25]. The training phase was implemented using the Fuzzy Logic classification scheme [19]. Fig. 3 compares hydrometeor classification results obtained with SVM (left) and FL (right) classifier for a sweep collected on 14 October 2015 at 01:05 UTC (top) and a vertical cut at 45° of the radar volume (bottom) composed by a sequence of sweeps containing that sweep. Output of the classifier are quite similar as expected, since the SVM classifier was trained with the FL. In the plot showing the vertical section, the figure obtained with SVM presents a lower number of “nodata” and more a spatially homogeneous graupel region is found in a precise region above the melting layer.

5 Conclusions

This work has presented a supervised hydrometeor classification approach based on a learning machine method, namely, the Support Vector Machine. The development of this SVM HCA was stimulated by the perspective application for a dual-polarization weather radar for civil aircrafts. However, the architecture is a generic one and the resulting classifier can be applied to the dual-polarization weather radars widely adopted by weather services. The implementation presented requires 4 dual-polarization radar measurements and the height of the 0°C level. Having established robust relations between input variables and known samples during the training phase by means of a well consolidated fuzzy logic HCA, in the prediction phase, the SVM HCA predicts output classes for a given sample. Validation has been conducted both with realistic scenarios generated by the numerical weather prediction model WRF (in this case a “truth” reference is available) and with real radar measurements through comparison with an FL output. Besides the good performance of SVM in terms of hydrometeor identification, the SVM approach has also exhibited a superior computational efficiency with respect to the FL approach. These aspects recommend to use the SVM HCA in quasi-real-time tasks. The approach used to develop the HCA can be easily extended to include more input variables and an extended set of output variables.

6 Acknowledgments

The authors acknowledge the support of the European Union Clean Sky Joint Technology Initiative, activity “System for Green Operation” through the projects KLEAN (Project ID: 306927), X-Wald (Project ID: 619236), and Win-FC (Project ID: 632536). Authors thank the Department of Civil Protection of Italy for providing radar

data. Content of this paper reflects only the authors' views. Clean Sky JU and European Union are not liable for any use that may be made of the information contained therein.

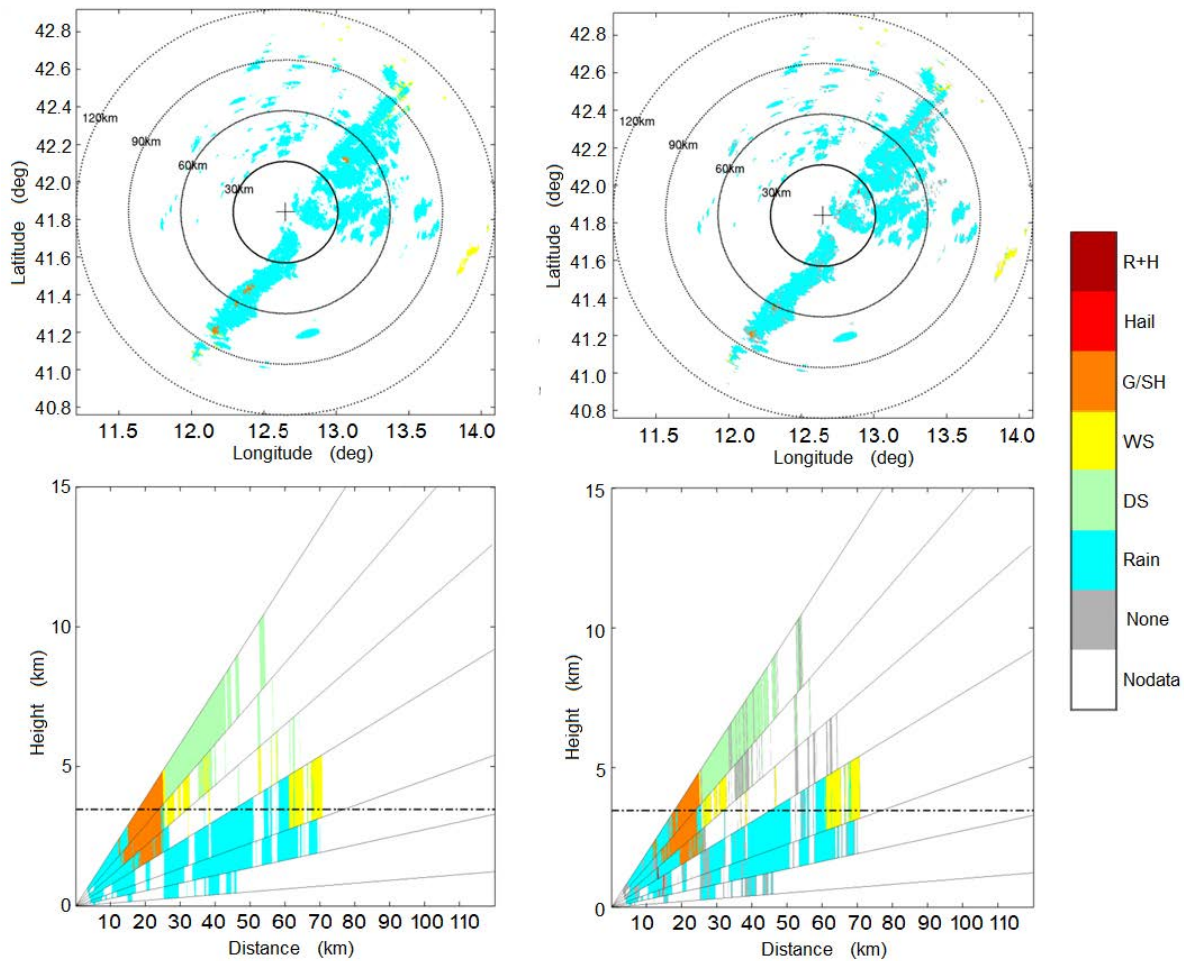


Figure 3: Example of SVM (left) and FL (right) classification for a radar volume collected by the Polar 55C radar on 14 October 2015 at 13:05 UTC. On top a sweep collected at 1.6° elevation angle is shown. On bottom, a vertical cut at 45° azimuth is reconstructed using six sweeps at different elevations is shown. “R+H” stands for hail mix; “G/SH” stands stands for wet snow. “DS” stands stands for wet snow; “DS” stands for dry snow; “none” stands for not classified; and “nodat” stands for no data.

References

- [1] Cuccoli A, Lupidi L, Facheris L, Baldini "Polarimetric radar for civil aircrafts to support pilots' decision in bad weather conditions, 2015 IEEE Int. Geosci. Remote Sens. Symp pp. 2327-2330, 2015.
- [2] Lupidi A., F. Cuccoli and S. Lischi, Polarimetric radar for flight support system with a reconfigurable electronic flight bag, 2016 IEEE Metrology for Aerospace (MetroAeroSpace), Florence, pp. 108-112 2016.
- [3] Liu, H., Chandrasekar, V, Classification of hydrometeor type based on multiparameter radar measurements: development of a fuzzy logic and neuro fuzzy systems and in-situ verification. J. Atmos. Ocean. Technol. 17, pp. 140–164, 2000.
- [4] Zrnić, D.S., Ryzhkov, A., Straka, J., Liu, Y., Vivekanandan, J. Testing a procedure for automatic classification of hydrometeor types. J. Atmos. Ocean. Technol. 18, pp. 892–913, 2001.
- [5] Dolan, B., Rutledge, S.A. A theory-based hydrometeor identification algorithm for X-band polarimetric radars. J. Atmos. Ocean. Technol. 26, pp. 2071–2088, 2009.

- [6] Dolan, B., Rutledge, S.A., Lim, S., Chandrasekar, V., Thurai, M. A robust C-band hydrometeor identification algorithm and application to a long-term polarimetric radar dataset. *J. Appl. Meteorol. Climatol.* 52, pp. 2162–2186, 2013.
- [7] Al-Sakka, H., Boumahmoud, A.A., Fradon, B., Frasier, S.J., Tabary, P., A new fuzzy logic Hydrometeor Classification Scheme applied to the French X-, C-, and S-band polarimetric radars. *J. Appl. Meteorol. Climatol.* 52, pp. 2328–2344, 2013
- [8] Bechini, R., Chandrasekar, V., A semisupervised robust hydrometeor classification method for dual-polarization radar applications. *J. Atmos. Ocean. Technol.* 32, pp. 22–47, 2015.
- [9] Wang, H., Ran, Y., Deng, Y. et al. *J Wireless Com Network.* 2017: 173. <https://doi.org/10.1186/s13638-017-0965-5>
- [10] Grazioli, J., Tuia, D., Berne, A., Hydrometeor classification from polarimetric radar measurements: A clustering approach. *Atmos. Meas. Tech.*, 8, pp. 149–170, 2015.
- [11] Besic, N., Figueras i Ventura, J., Grazioli, J., Gabella, M., Germann, U., Berne, A., Hydrometeor classification through statistical clustering of polarimetric radar measurements: A semi-supervised approach. *Atmos. Meas. Tech.* 9, pp. 4425–4445, 2016.
- [12] Scholkopf, B., Smola, A.J., *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond.* MIT Press: Cambridge, MA, USA, 2001.
- [13] Mountrakis, G., Im, J., Ogole, C., Support Vector Machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* 66, pp. 247–259, 2011.
- [14] Lee, Y., Wahba, G., Ackerman, S.A., Cloud classification of satellite radiance data by multicategory Support Vector Machines. *J. Atmos. Ocean. Technol.* 21, pp. 159–169, 2004.
- [15] Rossi R., Basili F., Del Frate F., Luciani M., Mesiano F., Techniques based on Support Vector Machines for cloud detection on quickbird satellite imagery, *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, pp. 515–518, 2011.
- [16] Federal Aviation Administration: Guidelines for the Certification, Airworthiness, and Operational Use of Electronic Flight Bags. https://www.faa.gov/documentLibrary/media/Advisory_Circular/AC_120-76C.pdf.
- [17] Serafino, G., Multi-objective aircraft trajectory optimization for weather avoidance and emissions reduction. In *Int. Workshop on modelling and simulation for autonomous systems*; Cham, Switzerland, pp. 226–239, 2015.
- [18] Roberto N., Baldini L., Adirosi, E., Facheris L., Cuccoli F., Lupidi, A., Garzelli, A., A Support Vector Machine hydrometeor classification algorithm for dual-polarization Radar, *Atmosphere*, 8, pp. 134, 2017.
- [19] Baldini, L.; Gorgucci E., Chandrasekar V., Petersen, W., Implementations of CSU hydrometeor classification scheme for C-band polarimetric radars. *32nd Conf. on Radar Meteorology*, Albuquerque, N. M., Amer. Meteor. Soc., 2005.
- [20] Baldini, L., Roberto, N., Gorgucci, E., Fritz, J., Chandrasekar, V., Analysis of dual polarization images of precipitating clouds collected by the COSMO SkyMed constellation. *Atmos. Res.* 144, pp. 21–37, 2013.
- [21] Milbrandt, J.A., Yau, M.K., A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, 62, pp. 3065–3081, 2005.
- [22] Lupidi, A., Lischi, S., Berizzi, F., Cuccoli, F., Roberto, N., Baldini, L., Validation of the advanced polarimetric Doppler weather radar simulator with Polar55C Real Observations. In *Proc. of the 15th International Radar Symposium (IRS)*, Gdansk, Poland, 16–18 June 2014.
- [23] Roberto, N., Baldini, L., Adirosi, E., Lischi, S., Lupidi, A., Cuccoli, F., Barcaroli, E., Facheris, L. Test and validation of particle classification based on meteorological model and weather simulator. In *Proc. of the 13th European Radar Conference (EuRAD)*, London, UK, 11–13 October 2016.
- [24] Roberto, N., Adirosi, E., Baldini, L., Casella, D., Dietrich, S., Gatlin, P., Panegrossi, G., Petracca, M., Sanò, P., Tokay, A., Multi-sensor analysis of convective activity in central Italy during the HyMeX SOP 1.1. *Atmos. Meas. Tech.*, 9, pp. 535–552, 2016.
- [25] Montopoli, M., Roberto, N., Adirosi, E., Gorgucci, E., Baldini, L., Investigation of weather radar quantitative precipitation estimation methodologies in complex orography. *Atmosphere*, 8, pp. 34, 2017.

Solar radio emission as a disturbance of aeronautical radionavigation *Perturbation du contrôle radar aérien par un sursaut radio solaire*

Karl-Ludwig Klein¹, Christophe Marqué², and Christian Monstein³

¹LESIA-UMR 8109 and Station de radioastronomie de Nançay - Observatoire de Paris, CNRS (further affiliations: PSL Res. Univ., Univ. P & M Curie and Paris-Diderot, Univ. Orléans, OSUC), F-92190 Meudon, France; ludwig.klein@obspm.fr

²Solar-Terrestrial Center of Excellence - SIDC, Royal Observatory of Belgium, Avenue Circulaire 3, B-1180 Brussels, Belgium; christophe.marque@oma.be

³Institute for Particle Physics and Astrophysics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH- 8093 Zürich, Switzerland; monstein@astro.phys.ethz.ch

Keywords: air traffic control radar, space weather, solar radio emission

Mots-clefs: radar de contrôle aérien, météorologie de l'espace, émission radio solaire

Abstract:

We present an overview of a solar radio burst that perturbed air traffic control radar in Sweden and other European countries on 4th November 2015. The spectral features of the radio burst are described. It is shown that it is particularly bright because of coherent emission processes. Such events are not predictable today. A systematic study of such events and the understanding of their impact on aviation will need a close cooperation between research and air traffic authorities.

Résumé:

Nous présentons un sursaut radio solaire qui a perturbé le contrôle radar aérien en Suède et d'autres pays européens, le 4 Novembre 2015. Nous décrivons les caractéristiques du spectre radio, montrant que l'émission est particulièrement forte parce qu'elle est cohérente. Il n'est pas possible aujourd'hui de prévoir de telles événements. Une collaboration étroite entre les chercheurs et les autorités de contrôle aérien sera nécessaire pour une étude systématique et la compréhension de l'importance pour l'aviation qu'a ce type d'émission radio.

1 An air traffic control incident in Sweden, 4th November 2015

On 4th November 2015 secondary air traffic control radar was strongly disturbed in Sweden and some other European countries. At some Swedish airports airplanes were not allowed to start during about one hour, and incoming aircraft were diverted. Belgian air traffic control noted disturbances when the radar antennas were pointing sunward, in particular in two time intervals lasting some minutes each, near 14:20 and 14:46 UT. The reasons for the disturbances were searched in a space weather event. The Sun was rapidly suspected as the source of the problem. It occurred at a time when the Sun was close to the horizon in Stockholm.

2 The solar radio burst

Secondary air traffic control radar is operated at frequencies of 1030 and 1090 MHz. A major solar flare occurred on the Sun at the time of the disturbances. The time evolution of the X-ray emission is shown in the bottom panel of Figure 1. The observations are from the *Geostationary Operational Environmental Satellite* (GOES) spacecraft operated by NOAA. The X-rays reflect the heating of a region in the corona overlying a sunspot group, from the usual ambient temperature of one to two million K to a few tens of millions of K. The time history shows an initial impulsive heating during the rise, lasting about 10 minutes, and a subsequent decrease of the heat input and a long-lasting cooling phase during the decay of the emission. The peak is reached near 13:50 UT. The top panel of Figure 1 shows the whole Sun radio flux observed by spectrographs at Humain (Royal Observatory Belgium; 1000-1500 MHz) and Nançay (Paris Observatory; ORFEES, 140-1000 MHz). Observations were also taken at Bleien (ETH Zürich), but are not shown here. The observations are presented as a dynamic radio spectrum. The radio flux is given by the shading in the time-frequency plane, dark shading meaning bright emission.

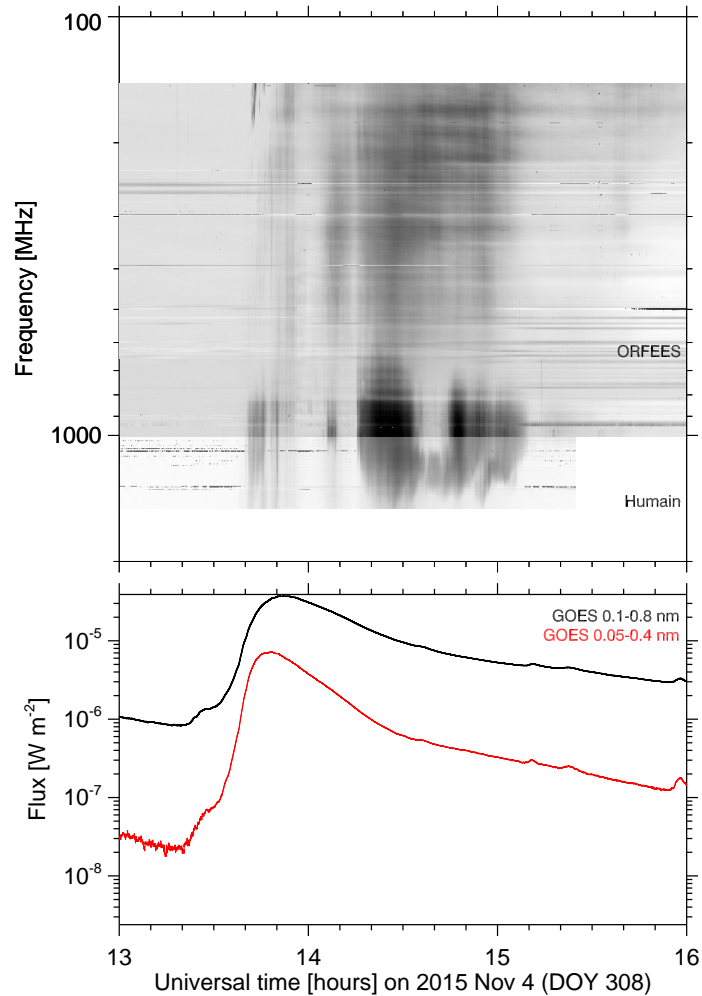


Figure 1 – Time history of the solar X-ray (bottom panel; GOES) and radio emission (top panel; spectrographs at Humain and Nançay radio observatories) on 4th November 2015. Strong radio fluxes are indicated by dark shading. Horizontal stripes are residuals of terrestrial emitters, which could not be fully remove from the data.

Like the X-rays, the radio emission has a duration of more than an hour. The timing is, however, quite different from the X-rays: the strongest radio emission occurs during the decay of the soft X-ray burst. The emission is brightest at frequencies between about 800 and 1400 MHz. A more detailed analysis of the Humain observations shows flux densities as high as 150,000 sfu (1 sfu = solar flux unit is $10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$) in the range 1030-1090 MHz, which is indeed exceptionally high. The time evolution shows two particularly bright periods, between 14:20 and 14:30 UT and 14:40 and 14:50 UT. Upon closer inspection the time history is actually a series of very short bursts across all or part of the (800-1400) MHz band, with individual durations below 1 s. The exceptional brightness of this burst in a frequency range that includes the operational frequencies of the air traffic control radar, and the time coincidence between the brightest radio emission in this range and the radar disturbances identified in Belgium are strong arguments that the solar radio burst was the reason of these disturbances.

The question is then how the Sun can produce such a high flux density at radio frequencies. Like other solar radio bursts at frequencies near and below 1000 MHz, this event shows a strong modulation both in frequency and in time, with narrow-band features and short individual bursts. This is not consistent with incoherent synchrotron emission from mildly relativistic electrons (a few hundreds of keV to some MeV), which is the mechanism considered at work at higher frequencies. Bright emission rather suggests some coherent, maser-like feature, while series of short bursts suggest that the conditions for this behaviour are not maintained for long times, but are reproduced in a repetitive way. Spectral fine structure finally suggests that the emission occur at some characteristic frequency of the coronal plasma. The basic idea is that energetic electrons confined in coronal magnetic fields develop anisotropic velocity distributions, with less electrons moving at small angles with respect to the magnetic field direction than at large angles (i.e., 90°). This is natural in a magnetic trap built by magnetic fields emanating from the solar interior, since electrons moving along the magnetic field do not experience the Lorentz force and can freely travel into the low and dense layers of the solar atmosphere, where they lose their energy by collisions. In this situation maser-like processes develop, where excess energy in the vertical motion of the energetic electrons is transferred to the plasma, creating upper hybrid waves, or even directly to electromagnetic waves that escape from the corona, creating electron cyclotron maser emission. Both the directly escaping electromagnetic waves and electromagnetic waves resulting from coupling processes of the upper hybrid waves can achieve high energy densities, i.e. strong radio bursts. The details of these processes are far from being fully understood. But favorable conditions, with energetic electrons being trapped in coronal magnetic loops, occur in the aftermath of the development of a coronal mass ejection (CME), where a large-scale magnetic field structure is driven out of the corona. The magnetic field then relaxes through magnetic reconnection, which creates new magnetic loops and also provides the context of the electron acceleration. The rail burst on 4th November 2015 was indeed accompanied by such a CME.

The radio emission on 4th November 2015 was associated in time with rapidly evolving activity occurring above an area with strong solar magnetic fields. This may be an argument for cyclotron maser emission, which is favored when the cyclotron frequency is higher than the plasma frequency. However, it is not clear if this can actually be the case in the solar corona. The type of the radio emission is not unique to this specific flare. Similar, and even stronger, radio bursts were reported in the literature to have interfered with GPS signals. The record holder was nearly two orders of magnitudes stronger than the 4th November event. For the time being we are not aware of any indication observable before the burst, for instance a particular configuration of sunspots, that could allow us to predict the strong radio emission.

3 Conclusion

The disturbance of Secondary Control Radar in Sweden and elsewhere on 4th November 2015 is generally attributed to space weather. From the timing and the exceptional nature of the emission, the solar radio burst is a most plausible candidate. The emission is particularly strong due to its coherent character, produced by anisotropic distributions of energetic electrons that are trapped in coronal magnetic fields. We have so far no means to forecast this kind of radio emission.

Explanations to the public are still lacking of why the air traffic was much more strongly disturbed in Sweden than in other countries, where the Sun was also in the beam of the radar antennas. The radio burst was very strong, but still stronger events were detected in the past. They were visible as disturbances of GNSS communications, but do not seem to have produced major problems with air traffic control radar. There is a need for further research and a thorough assessment of past incidents, which must have been recorded by air traffic control authorities, but which imply data that are not made public. This requires a close cooperation between the authorities and research. In the meantime, well-calibrated monitoring instruments of whole Sun radio fluxes covering the UHF band could at least provide a real-time identification of the origin of a disturbance.

This contribution is based on an article with the same title, authored by Christophe Marqué, Karl-Ludwig Klein, Christian Monstein, Hermann Opengoorth, Antti Pulkkinen, Stephan Buchert, Säm Krucker, Rudiger Van Hoof, and Peter Thulesen, which has been submitted to the *Journal of Space Weather and Space Climate*. Work at Paris Observatory was supported by the Agence Nationale pour la Recherche (ANR/ASTRID, DGA) project Outils radioastronomiques pour la météorologie de l'espace (ORME, contract No. ANR-14-ASTR-0027). KLK acknowledges helpful discussions with B. Collard, B. de Courville, L. Lapène, J.-Y. Prado, R. Rosso, and B. Roturier.

Radars et antennes tout numérique / Radars and full-digital antennas

Benefits of space-time diversity for radar *Les bienfaits de la diversité spatio-temporelle en radar*

François LE CHEVALIER

Emeritus Professor, Delft University of Technology, The Netherlands ; E-mail :F.LeChevalier@TUDelft.nl

Nikita PETROV

Researcher, Delft University of Technology, The Netherlands ; E-mail :N.Petrov@TUDelft.nl

Keywords (in English and French): Radar, Diversity, Wideband, MIMO; Radar, Diversité, Large bande, MIMO..

Abstract

When designing a new radar system, standard resolution trade-offs play a major role, providing the basic parameters of the radar, such as size, update rate, and range. Besides, diversity has long been used for mitigating fading effects due to the fluctuation of targets and clutter.

However, with the arrival of more flexible systems, using multiple parallel channels on transmit and receive, and wider instantaneous bandwidths, these standard trade-offs are becoming less simple – and more flexible. In this communication, we will analyze the benefits of diversity and its relations with range, Doppler, and angle, for detection and location of moving targets with wideband / wide beam radar systems. The idea is to contribute to a better understanding of the real benefits of agile transmissions for detection/localization of moving targets, focusing on range, velocity and angular measurement improvements, and on the benefits for detection of moving targets.

Special attention will be given to the quality of the different wideband wide-beam sensor modes for long range surveillance, and new results on detection of moving targets in clutter will be provided to demonstrate the effectiveness of these new architectures for small targets detection at long range, in difficult environments.

Résumé

Lors de la conception d'un nouveau système radar, les compromis classiques de résolution jouent un rôle majeur, pilotant les paramètres de base du radar, tels que la taille, la cadence de renouvellement et la portée. Par ailleurs, la diversité a longtemps été utilisée pour atténuer les effets de fading dus à la fluctuation des cibles et du fouillis.

Cependant, avec l'arrivée de systèmes plus souples, l'utilisation de plusieurs canaux parallèles en transmission et en réception, et des largeurs de bande instantanées plus larges, ces compromis standard deviennent moins simples - et plus flexibles. Dans cette communication, nous analyserons les apports de la diversité et ses relations avec la distance, le Doppler et l'angle, pour la détection et la localisation de cibles en mouvement avec des systèmes radar à large bande / large faisceau. L'idée est de contribuer à une meilleure compréhension des avantages réels des transmissions agiles pour la détection / localisation de cibles mobiles, en se concentrant sur les améliorations de portée, de vitesse et de mesure angulaire, et sur les avantages pour la détection des cibles en mouvement.

Une attention particulière sera accordée à la qualité des différents modes de détection pour la surveillance à longue portée, et de nouveaux résultats sur la détection des cibles mobiles en présence de fouillis seront présentés pour démontrer l'efficacité de ces nouvelles architectures pour la détection de petites cibles à longue portée, dans des environnements difficiles.

1 Introduction

1.1 Objectives

In order to design a surveillance radar¹, a critical point is the “illumination time”, also called the “time on target”: this time duration should be long enough to allow Doppler analysis, and to gain a sufficient signal to noise ratio (SNR), but also sufficiently small to allow a fast update rate, required by the user. This well-known trade-off between update rate and velocity resolution also involves the antenna beamwidth (the wider the beam, the better the velocity resolution, for a given update rate), and the clutter rejection capability (the wider the beam, the higher the clutter level), and has also direct consequences on the power budget (the wider the beam, the lower the antenna gain, but also the higher the coherent integration gain, for a given update rate).

These intricate relations between beamwidths, velocity resolution, and power budget (hence, range) are getting even more complex when taking into account the fluctuation characteristics of the targets and clutter, since

¹ This paper focuses on the basic issue of detection of moving targets with a ground surveillance radar, taking into account ground clutter.

performances can be improved through an increased averaging of clutter and target echoes – averaging which may itself be eased through widening of the beam, or longer illumination time. Such improvements are often more difficult to analyze, because they arise through modifications of the clutter and targets distribution functions, more complex than mere mean or standard deviation modifications (clutter and targets being generally not gaussian, averaging several samples generally changes the resulting distribution functions).

Moreover, modern radar systems generally operate over significant relative bandwidths – typically 10% – which can be exploited either coherently, with wider instantaneous bandwidths (providing improved range resolutions), or non-coherently, with a collection of measurements in different sub-bands. Again, these different operating modes have consequences on the power budget, but also on the distribution functions of targets and clutter.

The purpose of this paper is to try and clarify, with intuitive reasoning rather than precise equations, the order of magnitudes of these competing effects, so as to provide the designer with some basic insight necessary for building new radar architectures, involving multiple transmitting / receiving channels and arbitrary waveform agility.

1.2 Canonical problem

Detection being a 2-hypothesis problem (H_0 : no target, H_1 : a target), it basically comes down to comparing a certain quantity X , function of the received signals and of the expected situations (e.g. energy of the output of a matched filter), to a threshold depending on the required probability of detection P_d and probability of false alarm P_{fa} . This situation is shown on Figure 1, where the position of the threshold T defines the Probability of detection P_d (area with oblique lines) and the probability of false alarm P_{fa} (area with horizontal lines).

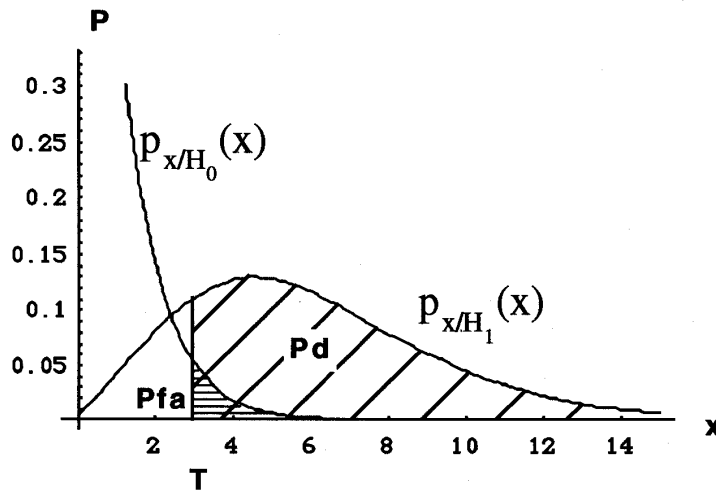


Figure 1: Detection and False alarm, after thresholding of the quantity X with threshold T (from [1]).

Obviously, the shape of the probability density functions p_{x/H_i} (probability of the received signal, under hypothesis H_i) is critical here. Using *diversity* is a means to improve the separation: generally speaking, averaging quantities is a way to reduce the spread of each probability density functions, and to bring it closer to a Gaussian (central limit theorem); Using *coherent integration*, or more generally matched filtering, is a way to increase the mean value of X under hypothesis H_1 . Both techniques thus improve the separation, in different ways: our objective here is to clarify these effects, and their consequences, for typical situations.

2 Standard Detection

Statistical Detection of radar fluctuating targets in the presence of noise is limited by the presence of noise and by the fact that the target may provide only very small signals for certain presentation angles or frequencies of illumination (a phenomenon also known as target fading in the literature). In order to mitigate target fading most radars use frequency agility:

- (1) They transmit successive bursts at different carrier frequencies;
- (2) When received, each burst is coherently processed as usual (Doppler filtering in each range cell);

(3) The outputs of these coherent summations are non-coherently summed (sum of the modulus, or the squared modulus), before final detection thresholding is applied.

This way improves the signal to noise ratio through each coherent burst processing. The resulting non-coherent summation allows consideration of observations at different frequencies, involving different target radar cross sections (RCS). In practice, these different bursts also generally use different repetition frequencies, allowing removal of the ambiguities in range and Doppler [1].

For a high required probability of detection, “some” non-coherent integration is preferable, in order to avoid getting trapped in a low RCS zone, especially for highly fluctuating targets (e.g. Swerling 1 targets, in the standard classification of targets fluctuations [1], [2]). “Some” means that coherent integration must first be used to get a sufficient signal-to-noise ratio (SNR) which should typically be larger than 0dB after coherent integration, so that it is not too much degraded by the modulus operation which, as every nonlinear operation, severely reduces the detection capability if it is done at low SNR. This explains the often used “Golden Rule”: first improve SNR through coherent integration, and then mitigate the low RCS zones by sending a few bursts with frequency agility from burst to burst. The price to pay for that noncoherent integration (and the associated diversity gain) when the available time-on-target is limited, is a lower Doppler resolution because of shorter coherent bursts.

The next question is: how many bursts are required during the time on target? Figure 2 gives a generic answer by showing the required SNR per burst as a function of the number of bursts on the target: $N=6$, or 10 , for Swerling targets [1], [2], [5]. The diversity gain can be defined as the difference in the required SNR per sample, between the coherent summation case and the noncoherent summation case. It clearly appears that the diversity gain is maximum for $N=6$: between 2.6 and 7 dB, depending on the required probability of detection, for this case of Swerling case 1 and 2. There is still a gain for higher resolutions, but it is much smaller.

Similar analyses with Swerling case 3 and 4 targets show that the gain is lower—as expected, since the fluctuations of Swerling 3 targets are smaller than those of Swerling 1—but still exists, at least for detection probabilities larger than 0.8 (i.e., gain between 2.7 and 1 dB). The diversity gain would then become a loss for very high resolution of Swerling 3 targets (1 dB loss for a target analyzed in 30 cells and a required $P_D = 0.8$, not shown in the figure).

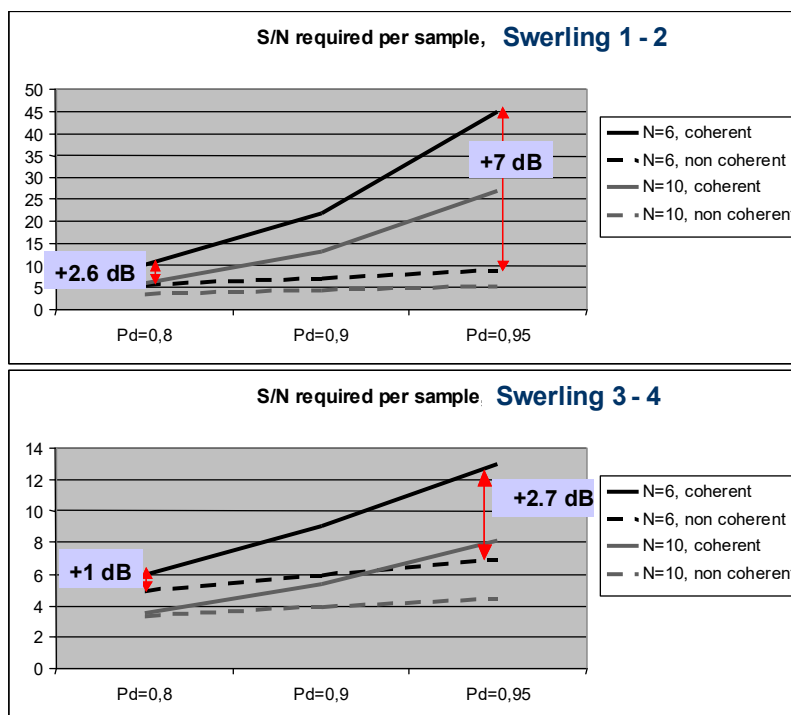


Figure 2: The effect of spatial resolution cells for fluctuating target detection. The traces show noncoherent versus coherent integration for a $P_{fa} = 10^{-6}$.

This basic analysis of diversity leads to the following conclusions:

- The more the target fluctuates, the higher the diversity gain.
- Higher requirements in the probability of detection P_d lead to a higher diversity gain.

- The number of bursts should be between 5 and 10, not more.

A very similar reasoning could apply in the angular / spatial domain, as just shown here in the range / frequency domain. For a multistatic system with a few radar sites, some non-coherent integration will nicely complement coherent integration of the signals received by each site. Depending on the exact signature of the target, spatial diversity and frequency diversity could be preferable: frequency diversity when the scatterers are distributed in range, spatial diversity when the scatterers are distributed in angle. A good solution, if possible, consists in combining both, for example with two or three frequencies per site, and two or three transmitting and/or receiving sites. However, it should be emphasized that frequency diversity is very generally an existing feature on most medium/long range monostatic radars (because most of them use multi-bursts operation, for ambiguity/eclipses removal), whereas spatial diversity, requiring multisite implementations, is only applicable for specific situations, such as passive radars as discussed by Cherniakov in [3] and Chernyak in [4]).

3 Wideband radar detection of moving targets

An essential limitation for standard narrowband radars using bursts of periodic pulses comes from the well-known pulsed radar range-Doppler ambiguity relation, which states that the ambiguous velocity V_a and the ambiguous range D_a are related by $D_a \times V_a = \lambda \times c / 4$. That relation means many ambiguities, either in range or velocity (or both), need consideration. This in turn implies the transmission of successive pulse trains with different repetition frequencies, requiring more time to be spent on target for ambiguity and blind speeds removal (without a corresponding gain in Doppler resolution, since the successive coherent pulse trains are then processed incoherently).

An alternative solution [1] consists in improving the range resolution (through increasing the instantaneous bandwidth) so that the moving target range variation (range walk or range migration) during the pulse train becomes non-negligible compared with the range resolution. Such radars may use bursts with a low pulse repetition frequency (no range ambiguities) and wideband pulses such that the range walk phenomena during the whole burst is significant enough to remove the velocity ambiguity (the range walk being a non-ambiguous measurement of the radial velocity). It then becomes possible to detect the target and measure range and velocity with only one long coherent pulse burst.

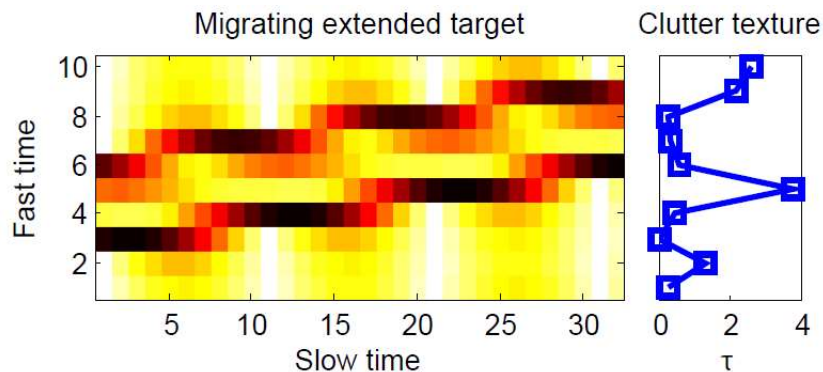


Figure 3. Range migrating extended target in spiky clutter

With such wideband radars, using Parseval's theorem, we first observe that the energy in the squared modulus of the impulse response (range profile) is the same as the energy in the squared modulus of the frequency response. So, summing the energy of the impulse response along the length of the target is equivalent, from a detection point of view, to summing the energy of the corresponding frequency response.

Integration along the range profile of the target for a pre-assumed length of the targets of interest (e.g., 15 m for air targets) is a way to combine coherent integration, used to obtain the range profile with its associated Doppler spectra in each range cell, with noncoherent integration. In other words, for wideband radars, coherent integration time — and the clutter separation that it provides — needs not be reduced to take benefit of diversity gain. Thus summing the bursts in each range cell of a narrowband agile radar is equivalent to summing the samples of the range profile of a high range resolution radar.

An appropriate detector for such wideband radars has been designed [6], ensuring CFAR performance with respect to the clutter texture, speckle correlation matrix, and target velocity. Comparison in Figure 4 includes the CFAR detector DIM-LRT (for Dependent Interference Model, Likelihood Ratio Test), and the clairvoyant detector

(assuming known correlation matrix of clutter). The loss of the proposed detector in comparison to the clairvoyant one is about 1 dB in each scenario. The analysis shows that target detection performance does not depend on clutter spatial correlation γ , but the detection performance depends on target velocity. Thus, the detection gain for the target with velocity $v_0 = 15$ m/s, which obeys a range-walk of about 3 range cells during the CPI, with respect to the stationary one is about 7 dB in K-distributed clutter with shape parameter $\nu = 0.5$.

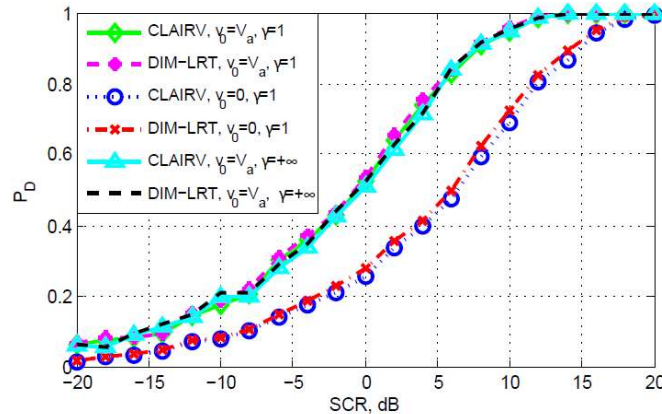


Figure 4. Detection probability of range migrating target in CG clutter with: $v = 0$ m/s and $v = 15$ m/s, SCR is the power of clutter after whitening, exponential correlation in range with $\gamma = 1$ and $\gamma = +\infty$; $PFA = 10^{-5}$.

The parameters of the radar are: $f_c = 10$ GHz, $B = 1$ GHz, $\delta_R = 0.15$ m, $T_r = 1$ ms, $M = 32$.

This phenomenon – improved detection of fast moving targets – can be well explained by the diversity of clutter, obtained by coherent integration of the target response during its migration over a few range cells. The faster the target, the more it migrates, the lower the probability to miss the target due to a possible clutter spike in one range cell, so the higher the probability of detection: target range-walk along non-Gaussian clutter thus provides a new way to exploit clutter diversity. This behavior is similar to detection of range-extended targets in CG clutter, where the detection performance depends on the target extent (see e.g. [7]). The observed diversity gain is not linear and saturates as the number of the range cells increases: we have observed that the major improvement is obtained by the first 3 range cells migration, and fully saturates for the range-walk over 5 range cells.

Generally speaking, the radar range resolution should be selected such, that the target of interest (given its expected dimensions and radial velocity) is spread over 5-10 range cells as a result of its range extent and its range migration. So meter resolution of the surveillance radar is sufficient for detection of typical air targets with a spatial diversity

4 Wide beam surveillance

Standard digital beamforming provides wide angular sector instantaneous coverage with a wide beam illumination on transmit by transmitting through one relatively small subarray, or through multiple sub-arrays with appropriate phase coefficients for widening the beam. In this technique, also known as “beam spoiling”, the multiple directive beams are simultaneously formed on receive through coherent summations of signals received on different subarrays, in parallel for each aiming direction.

Digital beamforming generally does not essentially change the power budget, compared to standard focused exploration, since the lower gain on transmit (due to wider illumination) is traded against a longer coherent integration time (made possible by the simultaneous observation of different directions). In fact, the main benefit provided by digital beamforming is an improved velocity resolution obtained through this longer integration time, especially useful for target identification purposes, or for detection of slow targets in clutter.

However, the improved velocity resolution of this wide beam exploration comes at a cost: i.e. the non-directive beam on transmit, which induces a poorer rejection of echoes coming from adjacent directions. For ground or surface applications, this means that detection of small targets in presence of strong clutter will become more difficult: the clutter echoes from different directions, which were cancelled not only through the Doppler rejection, but also through the angular separation on transmit and on receive, are now less easily rejected.

Moreover, the use of a wide beam on transmit implies that the angular resolution – and accuracy – is only obtained on receive; the angular resolution is thus poorer – approximately by a factor $\sqrt{2}$ – compared to the standard pencil beam solution.

In order to recover this angular separation on transmit (which was basic to standard focused beam techniques), it is necessary to code the transmitted signals (space-time coding), such that the signals transmitted in the different directions be different – and then become separable on receive as shown in Figure 5.

Past

- Send maximum energy on target, for maximization of the radar power budget
 - Pencil beam, scanning the panorama

Past: Pencil beam, scanning the panorama



Present

- Multiple simultaneous receive beams for faster search performance and more time to analyse the detected objects, providing additional information.
 - Limitation : the same waveform is sent in every direction. This severely limits systems discrimination performance in complex environments

Present: Widening on transmit, multiple receivers



Future

- Ubiquitous observation: simultaneously, different signals are sent in different directions, with reception on multiple parallel channels.
 - Combines wide coverage with selectivity on transmit

Future: multiple transmitters, multiple receivers



Figure 5: Pencil beam, Widebeam digital beamforming, Space-time coding (from [3]).

This method uses different signals simultaneously transmitted in different directions, thus jointly coding space and time, and then coherently processed in parallel on receive. Such concepts, first proposed and demonstrated by Drabowitch and Dorey [8,9], should now be considered as mature techniques to be implemented in operational systems. Basically, the main advantages to be gained are a better extraction of targets — especially slow targets — from clutter, multipath, and noise, and a better identification of targets obtained through longer observation times, and possibly wider bandwidths.

The generic configuration is shown on Figure 6: different codes (preferably with constant amplitude, for better efficiency of the amplifiers) are transmitted through the different antenna elements, or sub-arrays (e.g. lines, or columns of a 2D array); The resulting signal is transmitted by the antenna, resulting in different modulated signal in the different directions.

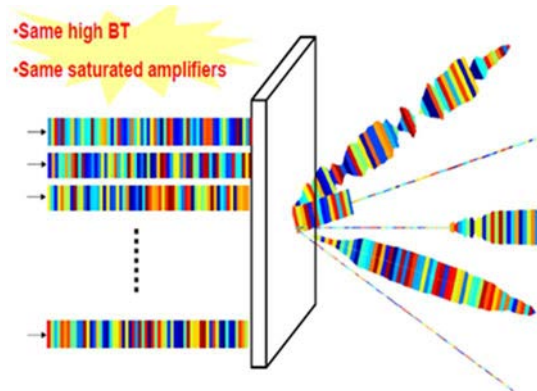


Figure 6: Space-time coding, generic case (from [5])(courtesy J.P. Guyvarch).

The main properties of space-time coding – as far as we are concerned here – are summarized in their range-angle ambiguity function. Indeed, due to the fact that different signals are transmitted through the different antenna elements, the result is a coupling between the range and angle information, so the angular diagrams cannot be analyzed without looking simultaneously at the range domain (the Doppler domain is not affected, since these coding are inside each pulse, and repetitive from pulse to pulse: the Doppler selectivity and Doppler sidelobes are then just as usual).

Let's consider these ambiguity function for "Circulating Codes", which are a good example of such space-time coding, with nice properties as detailed in [5] and [10]. Circulating Codes are generated, as shown on Figure 7, by

the same waveform (e.g. a chirp), “circulating” with a relative time shift δt through N transmitter channels. The relative time shift δt between adjacent circulating signals is equal to 1-time sample, $\Delta t = 1/\Delta F$, where ΔF is the signal bandwidth. As detailed in [5] and [11], it is possible to get different properties with space-time codes based on those circulating codes (or analogously on Frequency Diverse Arrays, which are very similar when the circulating signal is a chirp, since a time shift is then equivalent to a frequency shift).

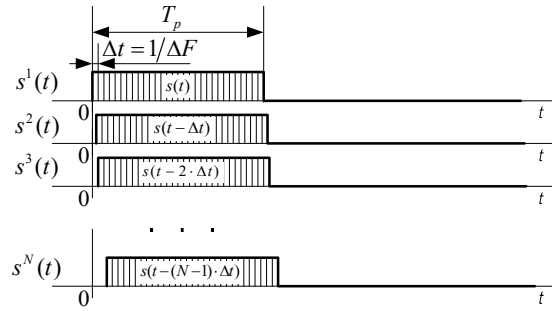


Figure 7: Circulating codes for Space-time coding(from [10]).

$$s^i(t) = s[t - (i-1)\Delta t] \text{ is the signal sent through the } i^{\text{th}} \text{ transmitting element.}$$

For this analysis, the following characteristics are chosen:

Tx/Rx Array: 15 elements, spacing $\lambda/2$; Tx: Wide beam, or space-time coding; Rx: Taylor weighting, 30dB
Signal: Circulating chirps, pulse duration $t=100\mu\text{s}$, $BT=256$, time delay between adjacent chirps: $\Delta t = 1/B = 0.34 \mu\text{s}$; Time weighting (Hamming) on receive. Space-time processing: mismatched « optimal » filtering for Delft codes [12].

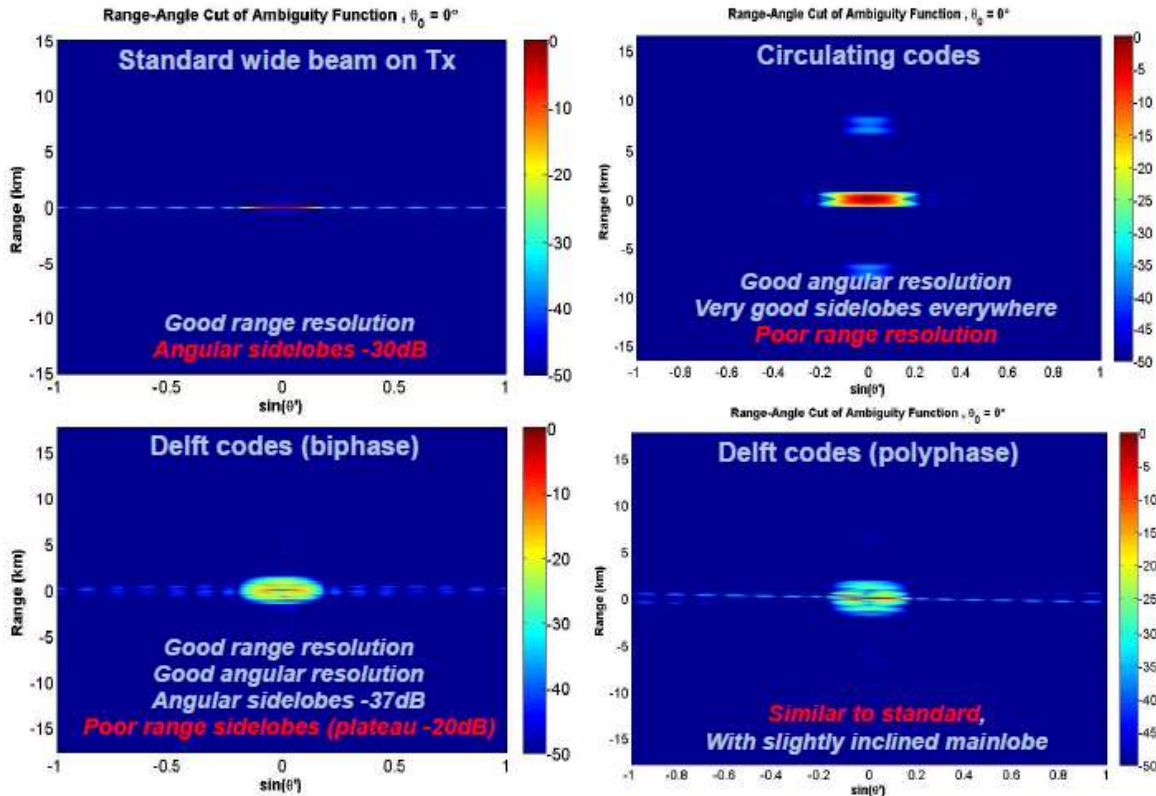


Figure 8: Ambiguity functions for Space-time coding with circulating codes and their variations (from [3]; Delft codes are circulating codes combined with a fixed phase code applied along the antenna array).

The main results can be outlined by looking at the range-angle ambiguity functions (output of the processing for every range and angle, when one target is present at range 0 – arbitrary value – and angle 0), shown on Figure 8:

- The angular resolution is improved with space-time coding, by a factor $\sqrt{2}$, (or 2 for a bidimensional coding of a rectangular array); this is not obvious on the figure, but careful analysis indeed gives this

result (a natural result, actually, since there is now directivity both on transmit and receive, compared with the directivity on receive only for the standard digital beam forming with wide beam on transmit);

- The position and level of the sidelobes are different with the different codings: in angle only for standard beam forming, almost no sidelobes but a degradation in range resolution for pure circulating codes, at different levels and positions for Delft codes [5].

As a result, the increased degrees of freedom provided by space-time coding on transmit open the way to adaptive systems where range and angle resolutions can be traded, depending on the mission and the actual environment (knowledge aided system). This trade-off can be operated differently for each burst, allowing some diversity on clutter to be obtained when several bursts are used for ambiguity removal, for instance. And of course, space-time coding also provides the improvement (mentioned above) in both accuracy and resolution higher than 2, for 2-dimensional antennas, compared with modern wide beam DBF Systems.

5 Discussion

A question then arises: how to combine diversity effects when using agile waveforms? Let us take a baseline example, with a typical modern radar using digital beamforming in elevation only, and a chirp waveform with pulse length 100 μ s, pulse repetition frequency 1 kHz. Any designer would like to benefit from:

1. High Doppler resolution, for visibility of slow and weak targets;
2. High angular resolution, in elevation (for altitude measurement) and azimuth (for tracking);
3. Diversity on the target, for improved detection in noise;
4. Diversity on clutter, for improved detection in clutter.

The first point requires long coherent integration time – but anyway this coherent integration time is limited by the fluctuations of the aspect angle of the target, typically to less than 100ms⁽²⁾.

The second point requires accurate angular measurements (monopulse) with narrow beams on transmit and receive.

The third point requires observations at different carrier frequencies, or from different aspect angles (multistatic system), or integration along a high resolution range profile.

The fourth point requires the target to be superposed to different patches of clutter, either through range migration or range extent of the target, or through multi-bursts with different range ambiguities (so that the target folds over different clutter patches).

These requirements tend to eliminate standard solutions, such as pencil beam with low range resolution (limited velocity resolution due to a short time on target), or standard digital beam forming with no ambiguity in range (limited angular resolution due to the wide beam, limited diversity on clutter)

Facing those trade-offs, several baseline solutions can be sketched:

- a. Pencil beam, high range resolution, unambiguous in range (low PRF): this satisfies all requirements if the coherent integration time is sufficient – and also provides valuable target analysis capabilities, with high resolution range-Doppler signatures.
- b. Space-time coding, low range resolution, ambiguous in range (high/medium PRF): this also satisfies all requirements; pure circulating codes could be a preferred solution in strong clutter environments, providing very low sidelobes everywhere.
- c. Space-time coding with high range resolution, unambiguous in range (low PRF): low sidelobes, high diversity, combined with valuable target analysis capabilities, with high resolution range-Doppler signatures.

These baseline descriptions should of course not be considered as definitive solutions to the very complex task of defining a multifunction radar (for instance, multistatic solutions could also make sense, possibly combined with space-time coding for solving the “rendez-vous” issue, cf [4]). The objective was rather, as outlined in introduction, to highlight and clarify some specificities of diversity effects which have to be considered when designing future systems. Many other aspects, from complexity and cost to multifunction requirements, have also to be taken into consideration – and should also bring out different advantages of high resolution and space-time coding for surveillance radars.

² There may also exist rapid fluctuations of the target due to moving parts (jet engines, rotating blades, wheels) ; however, for moving targets detection, these parts have lower radar cross-sections than the cell of the target, so they can be ignored in this general assessment.

References

- [1] F. Le Chevalier: *Principles of Radar and Sonar Signal Processing*. Norwood, MA, USA: Artech House, 2002.
- [2] M. A. Richards, J. A. Scheer, W. A. Holm, and W. L. Melvin: *Principles of Modern Radar*. Raleigh, NC, USA: SciTech Pub., 2010.
- [3] Cherniakov, M. (ed), 2007. *Bistatic Radar: Principles and Practice*; Wiley, Chichester, UK, 2007
- [4] Chernyak, V., 1998. *Fundamentals of Multisite Radar Systems* by CRC Press.
- [5] F. Le Chevalier: “Wideband Wide Beam Motion Sensing”, Chapter 12 in : *Advanced Ultrawideband Radar: Targets, Signals and Applications*, Ed. J. Taylor, CRC Press, 2016.
- [6] N. Petrov, F. Le Chevalier, A. Yarovoy: “Adaptive Detection of Range Migrating Targets in Correlated Compound-Gaussian Clutter” IET Journal Radar Sonar and Navigation, submitted 2018
- [7] K. Gerlach, “Spatially distributed target detection in non-Gaussian clutter,” IEEE Transactions on Aerospace and Electronic Systems, vol. 35, no. 3, pp. 926–934, 1999.
- [8] Drabowitch, S. and Aubry, C., 1969. “Pattern Compression by Space-Time Binary Coding of an Array Antenna”, Proceedings of the AGARD CP 66, Advanced Radar Systems.
- [9] Dorey, J., Blanchard, Y., Christophe, F., and Garnier, G., 1978. “Le Projet RIAS, Une Approche Nouvelle du Radar de Surveillance Aérienne”, L'Onde Electrique, vol. 64, no. 4.
- [10] Babur, G., Aubry, P. and Le Chevalier, F. 2013. Space-time radar waveforms: circulating codes, Journal of Electrical and Computer Engineering, Vol. 2013, Article ID 809691, 8 pages; Hindawi Publishing Corp.
- [11] Babur, G., Aubry, P. and Le Chevalier, F., 2013. Research Disclosure: Delft Codes: Space-Time Circulating Codes Combined With Pure Spatial Coding For High Purity Active Antenna Radar Systems; Research Disclosure N° 589037, May
- [12] Faucon, T., Pinaud, G. and Le Chevalier, F. 2015. Mismatched filtering for space-time circulating codes, Proceedings of IET International Radar Conference, Hangzhou, PR China, October.
- [13] A. De Maio and M. Greco: *Modern Radar Detection Theory*. Raleigh, NC, USA: SciTech Pub., 2016.

Potentialities of reduced BFN antennas for spatial Ku-band applications

Potentialités des antennes à formateur de faisceaux réduit pour applications spatiales en bande Ku

Phased array, mutual couplings, parasitic elements
Antenne réseau, couplages mutuels, éléments parasites

Jorick MILBRANDT¹, Cyrille MENUJER¹, Marc THEVENOT¹, Thierry Monédière¹

¹ Université de Limoges, XLIM-CNRS UMR 7252, 8706 Limoges, France, jorick.milbrandt@xlim.fr

Abstract/Résumé

To reduce the complexity of radiating arrays, we propose an architecture based on periodic lattice, parasitic elements and multi-excitations. The aim is to reduce the complexity of the beam forming network (BFN) while maintaining enough beam shaping controls. In this work, we detail the principle of this architecture with a simple example based of microstrip antennas within a 10x10 square grid whose only 30% of the elements are fed. Then, we explore the potentiality of reduced BFN antennas for spatial Ku-band applications by considering an architecture with an improved bandwidth, and a triangular grid of 57 elements with only 7 elements fed.

Pour réduire la complexité des antennes réseau, nous proposons une architecture utilisant une distribution périodique d'éléments rayonnants et d'éléments parasites, afin de créer un système multi-excitations. L'objectif est de diminuer la complexité du réseau formateur de faisceau tout en maintenant un nombre de contrôles suffisant pour la reconfiguration de l'antenne. Au cours de ce travail, nous détaillons en premier lieu les potentialités du concept avec un exemple d'une antenne réseau à maille carré de 10x10 éléments, dont seulement 30% des éléments sont excités. Puis nous testerons l'applicabilité de cette architecture pour une application spatiale en bande Ku, utilisant une antenne à 57 éléments disposés suivant une maille triangulaire, et dont seuls 7 accès sont excités.

1. Introduction

Phased array antennas have attracted a lot of research interest due to their capability to provide simultaneously a high gain and beam scanning. However, this technology is still expensive, complex and restricted to military and spatial applications. Indeed, phased array antennas need complex Beam Forming Networks (BFN) which are difficult to design, bulky and expensive [1]. Moreover, couplings between elements can affect the management of the antenna active SWR. That is why having a low-priced and efficient technology is an important challenge, particularly for high gain antennas. Many technologies were put forward to overcome these limitations [2] : thinned array [3], where some radiating elements are off or connected to dummy loads, sparse array [4] with removed elements, or also clustered array based on sub-arrays [5]. Nevertheless, designing irregular arrays is complicated because classical approximations based on periodicity are inapplicable. To overcome these latches, we propose an alternative concept to simplify BFN of antennas by using parasitic elements on a regular grid. The idea is based on the hybridization of concepts coming from antennas with parasitic elements and from lacunar elements antenna arrays, in order to maximize the effective area of the antenna, while minimizing the number of elements connected to the BFN. This architecture offers an interesting tradeoff between performances, cost and complexity [6].

In this paper, we appraise the potentialities of reduced BFN antennas in a specific frequency band. At first, the general architecture of reduced BFN antennas and its working principle will be presented. Next, a basic example to show some potentialities of the concept validity will be given. Finally, the results obtained with a reduced BFN antenna in Ku band will be shown.

2. Presentation of the reduced BFN antenna concept

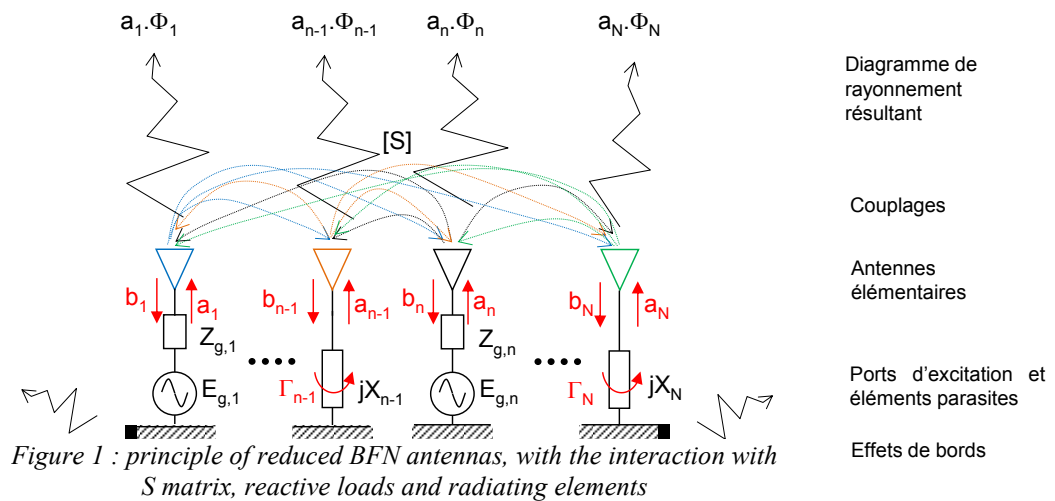
Présentation des antennes à formateur de faisceaux réduit

Reduced BFN antennas are based on regular arrays where all the radiating elements are the same. The originality of this concept is based on the parasitic elements which are loaded on reactive impedances, passive or tunable, and which contribute to the radiation pattern objectives, fixed or reconfigurable. The loads associated with these parasitic elements are synthesized so that they simultaneously contribute to both the optimization of the radiation pattern objective and to an active VSWR constraint. A basic schematic is done Figure 1. This architecture preserves the periodicity in order to simplify the computation of the EM problem.

It can be seen as an application and an extension of the work proposed initially by Harrington approach [7]. Contrary to most applications of Harrington's researches, as electronically steerable parasitic array radiator (ESPAR), our approach is not limited in terms of radiating elements and the antenna can include one or more excited elements to optimize the aperture efficiency of the antenna. One of the main objective in the design is to ensure strong couplings between elements and to spread their contributions on the antenna panel. The loads connected to parasitic elements act as a degree of freedom to tune their reflection coefficient, in order to satisfy simultaneously a radiation objective and a matching constraint. Compared to a classical array, the architecture has a good potentiality to manage the active VSWR of the ports connected to the BFN thanks to the parasitic elements and their loads that help to manage couplings contributions.

A reduced BFN antenna is designed in three steps :

- Design of the unit element and the array geometry,
- Determination of the S-matrix and extraction of radiation patterns Φ_i for each element thanks to an EM simulation software like CST MWS,
- Synthesis process with Matlab to determine the value of each complex load (ideally reactive) and the weightings of ports connected to the BFN.



3. Typical performances on a basic example of reduced BFN antenna

Performances typiques sur une antenne à formateur de faisceau réduit

3.1. Presentation of the antenna

Présentation de l'antenne

This example is based on a panel of classical square microstrip patches designed to operate at 11.8 GHz. The patches are excited by coaxial connectors as shown in Figure 2-a. These elements are arranged in a periodic square lattice of 10x10 elements. In order to show the potentialities of the reduced BFN antennas, we excite only 30 elements among 100. The excited elements are randomly chosen on the grid as shown Figure 2-b. This choice is arbitrary and may lead to some difficulties to realize the BFN. It is only chosen to demonstrate that it is possible to compute the loads that are able to satisfy the objective and constraints of the specifications, whatever the location of active ports. In these cases, it consists in maximizing the directivity in a specified direction and to satisfy a constraint on the frequency range ($|S_{ii}|_{\text{active}} \leq -10$ dB between 11.5 and 12 GHz)

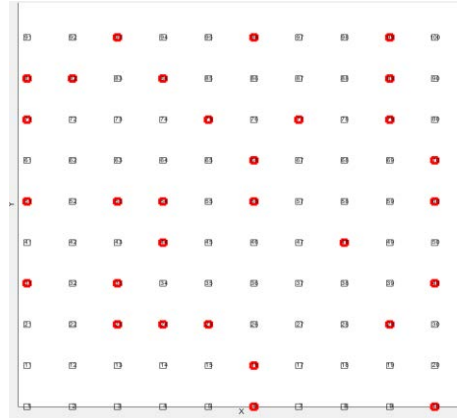
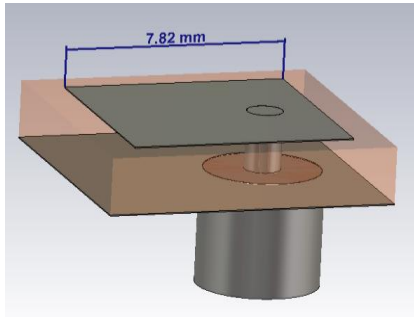


Figure 2 : a) element of the antenna array b) rectangular grid with excited elements (red) and parasitic ones

3.2. Results of the synthesis

Résultats des synthèses

To show the potentiality of the antenna, two directions are chosen : axial radiation $\{\theta = 0^\circ ; \varphi = 0^\circ\}$ and a beam steering $\{\theta = 30^\circ ; \varphi = 0^\circ\}$. Figure 3-b and figure 5-b show a cartography of the coupled waves. Red squares corresponds to elements connected to the BFN. The color of parasitic elements is then linked to couplings. The more the elements tends to red, the more the couplings are stimulated, resulting in an improved contribution to the overall efficiency. We can notice a significant contribution of parasitic in the two configurations tested. For elements connected to the BFN, the color is related to the coupled waves. Therefore, a green or blue level relates to an active matching better than -10 dB.

Figure 4 and figure 6 show the active matching of the fed elements.

Notice that in these results, loads are considered as reactive ones, without any losses. In a practical case, the gain will be affected by the technology chosen for the loads. As example, reflection-type phase shifter can be used with varactors diodes. For better performances, MEMS can be a good solution.

3.2.1. Objective pattern 1 : $\{\theta = 0^\circ ; \varphi = 0^\circ\}$

Directivité : 22.9089

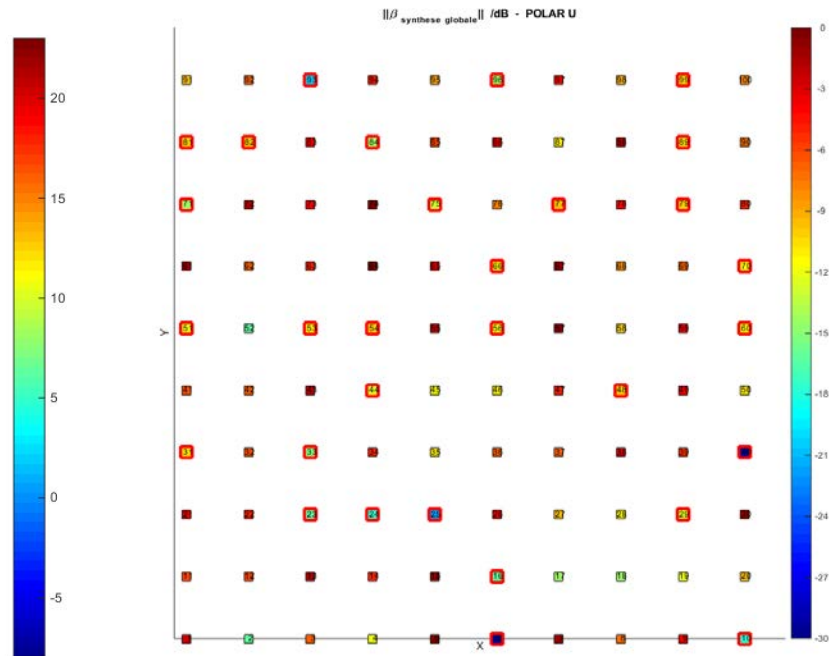
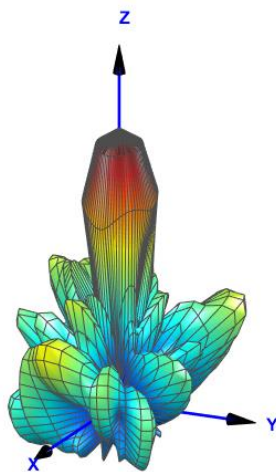


Figure 3: a) radiation pattern obtained with optimization 1 b) excitation weighting (modulus). Red squares correspond to elements connected to the BFN.

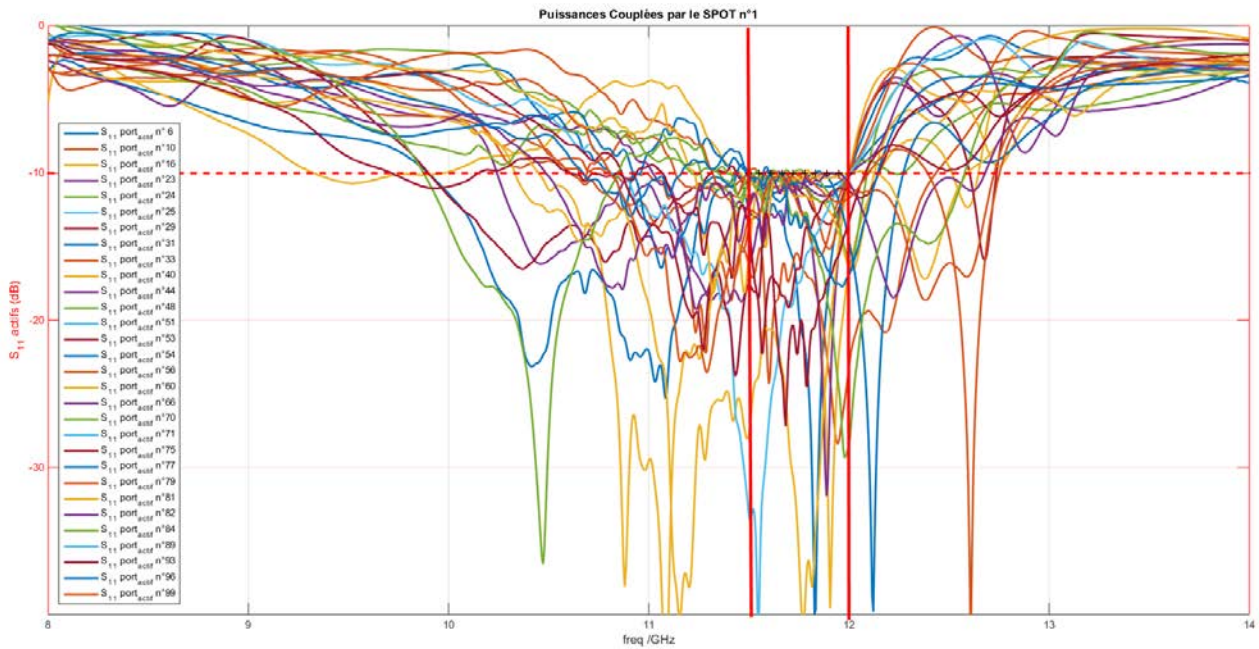


Figure 4: active S_{ii} for objective 1, the optimization bandwidth is delimited by the two red lines, with an constraint set to -10 dB

3.2.2. Objective pattern 2 : $\{\theta = 30^\circ ; \varphi = 0^\circ\}$

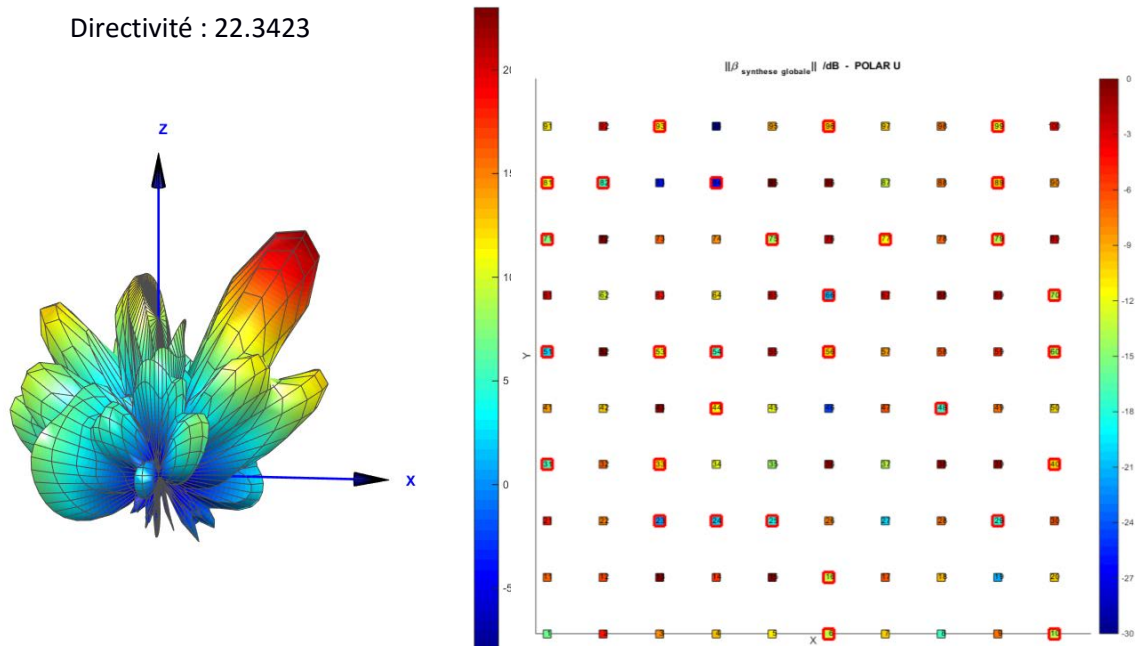


Figure 5: a) radiation pattern obtained with optimization 2 b) excitation weighting (modulus)

For each case, Figure 4 and 6 show that the active matching on the bandwidth is satisfactory (i. e. the reflection coefficients are smaller than -10dB between 11,5 and 12GHz). Directivity of a reduced BFN antenna whose only 30% of the elements are excited is quite similar to the directivity of a classical array, as mentioned in the Table. In this case, this reference directivity is calculated from the one of an equiamplitude and equiphase aperture area. But we have to put into perspective this point as we considered the reactive loads as ideal in these simulations. In Figure 3-b and 5-b it clearly appears that parasitic elements are strongly coupled and contribute to the final radiation pattern. We can conclude that, with appropriate loads, it is possible to stimulate the effective aperture of the panel even if only 30% of the elements are excited.

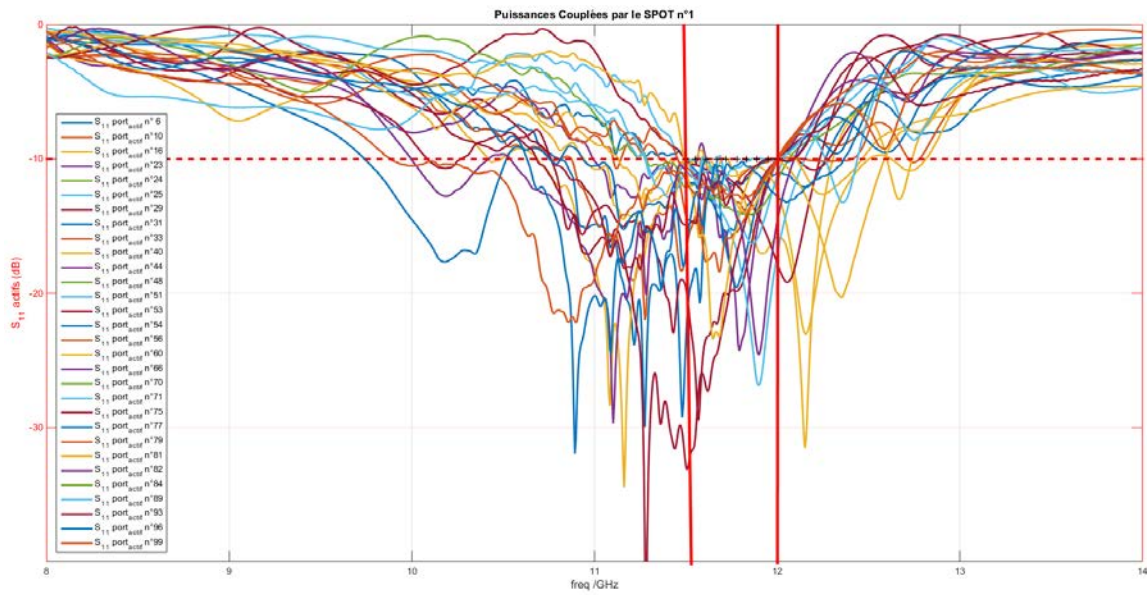


Figure 6: active S_{ii} for objective 2, the optimization bandwidth is delimited by the two red lines

Objective pattern	Directivity max for an ideal array of the same surface	Directivity obtained
1	23.0 dBi	22.9dBi
2	22.4 dBi	22.3dBi

4. Application to a bi-band antenna

Application à une antenne bi-bande

4.1. Design of the antenna

Présentation de l'antenne

The goal of this part is to show the potentiality reduced BFN antennas for spatial applications in Ku-band, where it is necessary to realize a transmit and receive link (Rx : 11.7-12.2 GHz ; Tx : 14-14.5 GHz). An element is designed to work on this band. As Figure 7-a shows, the element is composed of a square patch and a stacked bow-tie to improve the matching bandwidth. An aperture excites the element through a microstrip line. Figure 7-b shows the element reflection coefficient that is below -10 dB on the whole bandwidth.

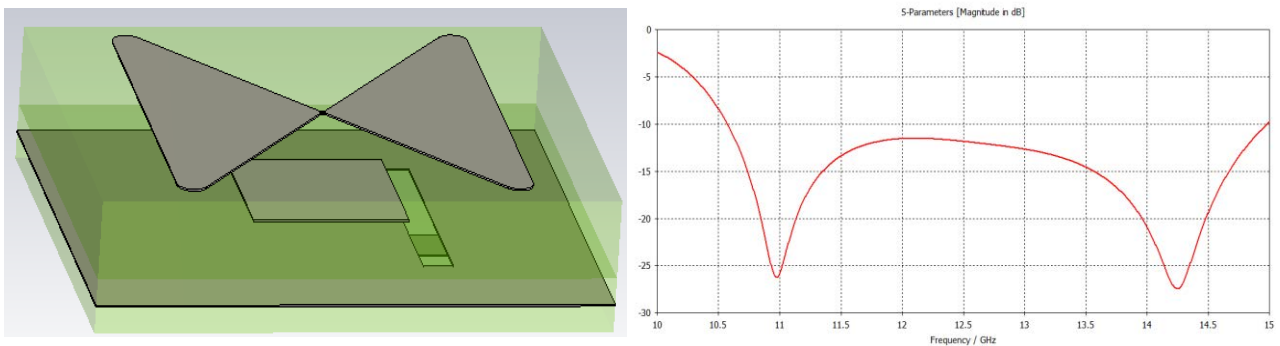


Figure 7: a) antenna element for a Ku-band application b) S_{11} of the element

We choose a triangular grid in order to optimize couplings between elements. Indeed, tests on bow-tie elements show better couplings between elements diagonally placed. Moreover a stabilization of couplings is observed on the full bandwidth thanks to the triangular grid. The grid is displayed in Figure 8.

Contrary to the previous example, a more conventional distribution of the excited elements has been chosen to take into account design features of the BFN. So seven elements are excited for the hexagonal distribution and the center element.

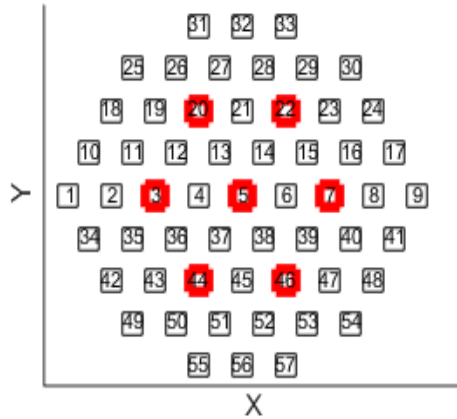


Figure 8: antenna array grid and excited elements (in red)

4.2. Results

Résultats

We have synthesized two configurations to realize a beam steered in the direction $\{\theta = 30^\circ ; \varphi = 0^\circ\}$. The first one correspond to the low part of Ku-band (11.7-12.2 GHz) and the second one to the high part (14-14.5 GHz).

4.2.1. Objective 1 : $\{f = 12\text{GHz} ; \theta = 30^\circ ; \varphi = 0^\circ\}$

This first synthesis is realized on the lower band of the Ku-band. We obtain a directivity of 18.7dBi (Figure 9-a) and all the active $|S_{ii}|$ satisfy the -10dB constraint (Figure 9-b).

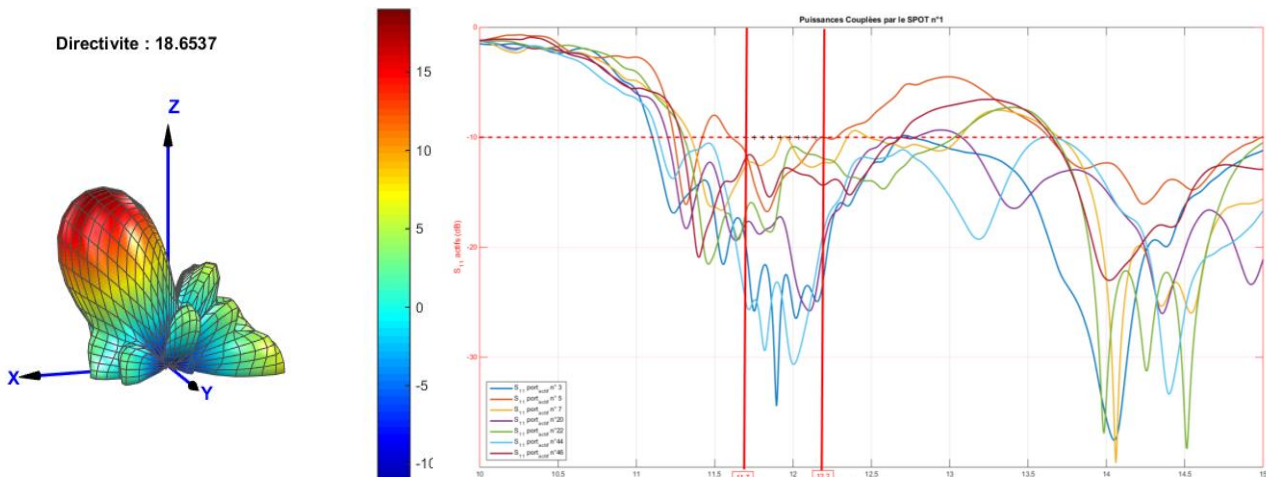


Figure 9: a) Radiation pattern at 12GHz b) Active S_{ii}

4.2.2. Objective 2 : $\{f = 14\text{GHz} ; \theta = 30^\circ ; \varphi = 0^\circ\}$

The synthesis of the reduced BFN antenna on the higher part of the Ku-band shows a directivity slightly lower than the one obtained at 12GHz, even if it reaches 18 dBi (Figure 10-a). The active matching is correct on the full bandwidth between 14-14.5 GHz (Figure 10-b) i.e. the reflection coefficient is lower than -10dB.

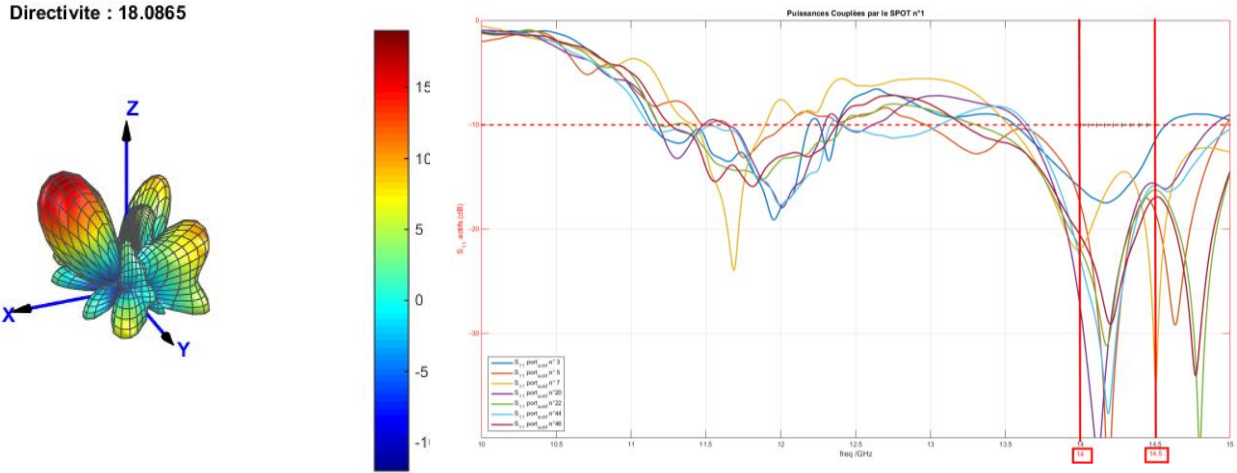


Figure 10: a) Radiation pattern at 14GHz b) Active Sii

A comparison with the directivity of an ideal aperture is performed in the Table below. A difference lower than 3 dB is observed, confirming the potentialities of this solution.

Objective pattern	Directivity max for a classical array of the same surface	Directivity obtained
1	21.3 dB	18.7dB
2	21.3 dB	18.1dB

We note a correct beam steering and good active reflection coefficients on each bandwidth. A slight degradation of directivity appears which is explained by the small number of excited elements (12%). The difference of directivity between the lower and the higher band is caused by higher couplings at lower frequencies. Nevertheless, these performances are really interesting considering the number of excited elements which permits a great simplification of the BFN.

5. Conclusion

An interesting concept of reduced BFN antenna is presented in this paper. We demonstrate the interest of the combination of excited elements and parasitic ones. We proved that it is possible to manage both active matching and radiation patterns with this concept. We showed the capabilities of this concept for Tx/Rx Ku-band applications. For such applications, we estimated the performance numerically on a triangular grid array of 57 elements. Great possibilities are offered by reduced BFN antennas to adapt phased array antennas to mass market by simplifying their design and reducing their cost.

6. Acknowledgment

Remerciements

This work is funded by DGA (French Defense Agency) and CNES (French Space Agency).

7. References

Références

- [1] J. S. Herd and M. D. Conway, "The Evolution to Modern Phased Array Architectures," *Proc. IEEE*, vol. 104, no. 3, pp. 519–529, Mar. 2016.
- [2] P. Rocca, G. Oliveri, R. J. Mailloux, and A. Massa, "Unconventional Phased Array Architectures and Design Methodologies—A Review," *Proc. IEEE*, vol. 104, no. 3, pp. 544–560, Mar. 2016.

- [3] A. Ishimaru and Y.-S. Chen, "Thinning and broadbanding antenna arrays by unequal spacings," *IEEE Trans. Antennas Propag.*, vol. 13, no. 1, pp. 34–42, Jan. 1965.
- [4] C. Coman, I. Lager, and L. Ligthart, "Design considerations in sparse array antennas," 2006, pp. 72–75.
- [5] R. Haupt, "Reducing grating lobes due to subarray amplitude tapering," *IEEE Trans. Antennas Propag.*, vol. 33, no. 8, pp. 846–850, Aug. 1985.
- [6] A. Oueslati, C. Menudier, M. Thevenot, and T. Monediere, "Potentialities of hybrid arrays with parasitic elements," 2014, pp. 1829–1830.
- [7] R. Harrington, "Reactively controlled directive arrays," *IEEE Trans. Antennas Propag.*, vol. 26, no. 3, pp. 390–395, May 1978.

Système radar FMCW pour l'identification de transpondeurs *FMCW Radar System for Transponder Identification*

Nizar BOUHLEL A¹, Stéphane MERIC B¹, Claude MOULLEC C³, and Christian BROUSSEAU D⁴

¹ Univ Rennes, INSA Rennes, CNRS, IETR, UMR 6164, nizar.bouhleb@insa-rennes.fr, stephane.meric@insa-rennes.fr

³ Advanten, claude.moullec@advanten.fr

⁴ Univ Rennes, CNRS, IETR, UMR 6164, christian.brousseau@univ-rennes1.fr

Mots-clés: Onde continue à fréquence modulée, radio-identification, traitement de signal, mesure de distance et de décalage fréquentiel, carte distance-décalage fréquentiel.

Keywords: Frequency-modulated continuous wave (FMCW), RFID transponder, signal processing, distance and shift frequency measurements, range-frequency map.

Résumé:

Dans cet article, nous traitons d'un système radar à onde continue à modulation de fréquence (FMCW) utilisé pour localiser et suivre des cibles par l'évaluation de la fréquence du signal de battement radar reçu. Nous considérons un transpondeur linéaire comme cible et nous réalisons une identification de cible grâce à la fréquence de décalage induite par cette cible active. Les performances de localisation des radars FMCW basés sur la radio-identification dans la gamme (SHF) sont présentées et étudiées. Des simulations et des mesures viennent compléter le papier.

Abstract: In this paper, we deal with a frequency modulated continuous wave (FMCW) radar used for localizing and tracking targets by frequency evaluation of the received radar beat signal. We consider the transponder as target and we achieve an identification application through frequency identification thanks to the shift frequency induced by the active target. The localization performance of FMCW radars based on radio-identification in the range (SHF) are presented and studied. Simulations and measurements complete the paper.

1 Introduction

Au cours de la dernière décennie, les systèmes de radio-identification (RFID) ont reçu une grande attention en raison de leur architecture simple, de leurs dimensions réduites, de leur faible coût, de leur faible puissance et de leur facilité de déploiement. La technologie RFID est principalement utilisée dans l'identification et le suivi des personnes ou des objets. Grâce aux avancées technologiques, il est possible de travailler avec une gamme UHF et SHF permettant une plus grande distance de fonctionnement [1]. Deux types de systèmes RFID sont utilisés : passif et actif. Les transpondeurs UHF passifs peuvent fonctionner jusqu'à une distance de 10 mètres. Les transpondeurs actifs avec leur propre batterie permettent des détections sur des distances plus élevées. Dans le cas de transpondeur actif, nous parlons de systèmes radars primaire et secondaire. Le système radar basé sur le principe du radar à onde continue modulée linéairement en fréquence (FMCW pour *Frequency Modulated Continuous Wave*) a des propriétés intéressantes. En effet, il émet et reçoit simultanément et sans interruption. Il est largement utilisé pour les applications industrielles et académiques avec de nombreuses applications [2]. Les avantages des radars FMCW sont multiples : détection pour les applications à courte distance, résolution en distance dépendant de la bande émise, pas de distance aveugle, facile mise en oeuvre technologique, faible consommation d'énergie, conception de systèmes miniaturisés, simplicité et faible coût. En particulier, pour la détection des cibles proches, le radar FMCW est une alternative au radar à impulsions car il n'est pas limité par la largeur d'impulsion d'émission et par le temps de commutation entre la transmission et la réception. Cette communication porte sur le système GESTAR (GEnerator System for TArget Radar identification). Ce système permet de détecter et de localiser des cibles coopératives dans des environnements complexes (zones urbaines, bâtiments, forêts, neige). Il est composé d'un radar interrogateur (radar primaire) de faible dimension et d'un transpondeur d'identification (radar secondaire). La détection se fait grâce à un décalage fréquentiel produit par le transpondeur, de valeurs supérieure à celui des mobiles environnants. Le transpondeur ne répond à l'interrogateur que lorsque ce dernier produit un signal d'interrogation sous la forme d'un signal FMCW. Ce signal FMCW est caractérisé par une fréquence porteuse RF, une puissance émise et un balayage en fréquence.

Le transpondeur retransmet ce signal en y introduisant sa propre fréquence de décalage. La puissance transmise par le transpondeur est capturée par le radar primaire. Le signal reçu est démodulé et le signal en bande de base est traité numériquement dans le domaine de Fourier pour donner lieu à une carte distance-décalage fréquentiel. La cible (transpondeur) est alors localisée avec précision en utilisant une détection par taux constant de fausses alertes (CFAR pour *Constant False Alarm Rate*). Le radar affiche alors la position en distance radiale, la position angulaire et la fréquence de décalage des cibles. Les performances de ce système sont caractérisées par plusieurs paramètres à savoir : la résolution en distance, la résolution de la fréquence de décalage (identification) [3], la résolution en angle et la capacité de détection (portée maximale).

2 Localisation basée sur un signal FMCW

Les radars FMCW sont des systèmes radar émettant une forme d'onde continue et modulée en fréquence. De plus, ces radars émettent et reçoivent simultanément. Nous présentons ici la forme d'onde FMCW linéaire et le modèle du signal envoyé par un transpondeur (i). Nous nous plaçons dans le cas d'un signal bande étroite où la bande émise est petite devant la fréquence porteuse qui se situe dans la gamme SHF.

2.1 Modélisation du signal reçu et modulé

Le principe de la radio-identification utilisant les radars FMCW est présenté dans la figure 1. Le principe est résumé ainsi : émission et propagation d'une onde électromagnétique, rétrodiffusion de cette onde par le transpondeur (i) en injectant sa propre fréquence de décalage f_D , réception et analyse du signal reçu par le radar. La forme d'onde FMCW linéaire consiste à émettre un *chirp* de fréquence de départ f_0 et de bande

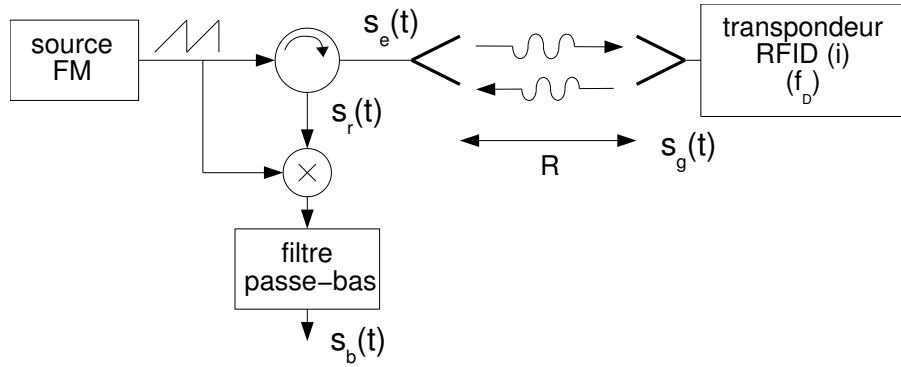


FIGURE 1 – Configuration du système de localisation RFID basé sur un radar FMCW

B avec une période de répétition T . Ce *chirp* est une onde sinusoïdale dont la fréquence instantanée évolue linéairement au cours du temps. Le signal $s_e(t)$ émis par le radar primaire FMCW s'écrit au cours du temps $0 < t < T$

$$s_e(t) = A_e \exp j2\pi \left(f_0 t + \frac{B}{2T} t^2 \right) \quad (1)$$

avec A_e , f_0 , B et T représentant respectivement l'amplitude, la fréquence porteuse, la largeur de bande, et la durée de balayage ou temps de modulation. Le transpondeur RFID situé à une distance R du radar, module la fréquence porteuse par une fréquence de décalage f_D . Le transpondeur RFID est considéré comme un radar secondaire qui transmet le signal $s_g(t)$ décrit avec :

$$s_g(t) = K A_e \exp j2\pi \left(f_0 \left(t - \frac{R}{c} \right) + \frac{B}{2T} \left(t - \frac{R}{c} \right)^2 - f_D t \right) \quad (2)$$

où le facteur K représente un coefficient d'atténuation et directement lié à l'équation du radar. Le signal reçu par le radar $s_r(t)$ est une version atténuée et retardée du signal transmis avec un retard de $\tau = 2R/c$ où c est la vitesse de l'onde. Le signal reçu s'exprime alors avec :

$$s_r(t) = K' A_e \exp j2\pi \left(f_0 (t - \tau) + \frac{B}{2T} (t - \tau)^2 - f_D \left(t - \frac{\tau}{2} \right) \right) \quad (3)$$

où K' prend en compte le trajet retour de l'onde. Le principe d'une démodulation homodyne du signal reçu consiste à mélanger le signal $s_e(t)$ avec le signal reçu $s_r(t)$. L'expression du signal de battement à la sortie du mélangeur est $s_b(t) = s_e(t)s_r^*(t)$ où $(*)$ désigne l'opérateur conjugué. Le retard entre le signal transmis et le

signal reçu se traduit par une fréquence de battement constante appelée f_b qui permet de mesurer la distance entre la cible (transpondeur) et le radar (interrogateur). Le signal de battement s'exprime comme suit

$$s_b(t) = K'' A_e \exp j2\pi \left(\phi_0 + \frac{B}{T}\tau t + f_D t \right) \quad (4)$$

où la fréquence de battement f_b est définie par

$$f_b = \frac{B}{T}\tau + f_D = \frac{2BR}{cT} + f_D$$

et ϕ_0 est une phase constante (K'' correspond à l'influence de l'opération de démodulation). Si la valeur de f_D est connue et compensée, la connaissance de la valeur de f_b permet de déterminer R (figure 2(a)). Dans le cas d'un décalage de fréquence induit dans le signal reçu, un radar FMCW peut également mesurer ce décalage de fréquence f_D en intégrant des données de plusieurs cycles de réplication d'onde comme le montre la figure 2(b). Pour ce faire, on définit

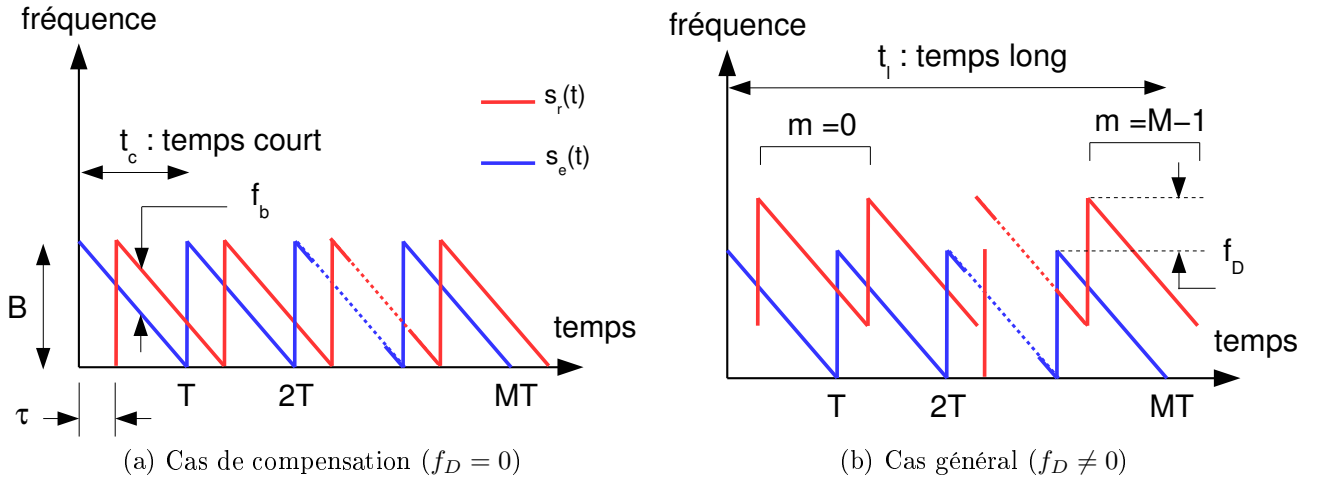


FIGURE 2 – Signal de battement sur plusieurs durées de modulation T par rapport à la fréquence de décalage f_D

- le temps court t_c pendant la durée d'un chirp ($0 < t_c < T$),
- le temps long t_l qui définit le numéro m de chirp $t_l = mT$ où $m = 0, 1, \dots, M - 1$ avec M est le nombre total de durée de modulation,
- le temps courant $t = t_c + t_l$.

Par conséquent, le signal de battement $s_b(t)$ est décrit à la fois par rapport à l'axe des temps t_c et t_l [3]

$$s_b(t_c, mT) = K'' A_e e^{j2\pi\phi_0} e^{j2\pi f_b t_c} e^{j2\pi f_D mT}. \quad (5)$$

Pour établir la carte distance-décalage fréquentiel (permettant d'identifier des cibles par leurs distances et leur fréquence de décalage), l'analyse du signal de battement est effectuée dans le domaine de Fourier. Deux transformées de Fourier doivent être appliquées au signal. En cas de temps continu, la première transformée de Fourier notée $S_b^c(f, t_l) = TF\{s_b^c(t_c, t_l)\}(f)$ est calculée pour $\tau_{max} < t_c < T$ où l'exposant c signifie le cas continu et τ_{max} correspond au retard d'une distance maximale de visualisation (ou d'analyse). L'expression de $S_b^c(f, t_l)$ est donnée comme suit

$$S_b^c(f, t_l) = K A_e e^{j2\pi\phi_0} e^{j2\pi f_D t_l} (T - \tau_{max}) \times e^{-j\pi(f-f_b)(T+\tau_{max})} \text{sinc}(\pi(f-f_b)(T-\tau_{max})) \quad (6)$$

Il est clair que le maximum de $S_b^c(f, t_l)$ est atteint quand $f = f_b$. L'application de la deuxième transformée de Fourier pour le traitement de fréquence de décalage où $0 < t_l < (M - 1)T$ est donné comme suit

$$S_b^c(f, f') = K A_e e^{j2\pi\phi_0} (T - \tau_{max}) e^{-j\pi(f-f_b)(T+\tau_{max})} \text{sinc}(\pi(f-f_b)(T-\tau_{max})) \times (M-1)T e^{-j\pi(f'-f_D)(M-1)T} \text{sinc}(\pi(f'-f_D)(M-1)T) \quad (7)$$

L'expression donnée par (7) qui est la transformée de Fourier bidimensionnelle du signal de battement permet d'établir une carte distance-fréquence de décalage. Dans le cas d'une description de signal à temps discret,

le signal de battement traité est maintenant discret. Le temps court t_c est échantillonné selon la fréquence d'échantillonnage F_e , et le temps long t_l selon la fréquence de répétition de période $F_s = 1/T$. La transformée de Fourier en temps discret qui est liée à la transformée de Fourier dans le cas continu par la formule de Poisson est donnée comme suit

$$\begin{aligned}
S_b(f, f') &= F_e F_s \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} S_b^c(f - nF_e, f' - kF_s) \\
&= F_e F_s (T - \tau_{max})(M - 1) T e^{j2\pi\phi_0} \sum_{n=-\infty}^{\infty} e^{-j\pi(f - nF_e - f_b)(T + \tau_{max})} \text{sinc}(\pi(f - nF_e - f_b)(T - \tau_{max})) \\
&\times \sum_{k=-\infty}^{\infty} e^{-j\pi(f' - kF_s - f_D)(M - 1)T} \text{sinc}(\pi(f' - kF_s - f_D)(M - 1)T). \tag{8}
\end{aligned}$$

2.2 Analyse angulaire

Dans le cas du système GESTAR, le radar primaire dispose de 2 antennes de réception. Ainsi, il est possible de mesurer la position angulaire des cibles. Une méthode monopulse de comparaison de phase [4] est utilisée pour estimer la position angulaire des cibles dans le plan horizontal. Cette méthode consiste à mesurer la différence de phase ϕ observée entre les deux canaux récepteurs séparés d'une distance d :

$$\frac{\Delta}{\Sigma} = -j \tan\left(\frac{\phi}{2}\right) \tag{9}$$

où Δ et Σ sont respectivement la différence et la somme des réponses de la cible sur les deux antennes de réception. La phase mesurée ϕ est liée à la séparation horizontale d des deux antennes et à l'angle d'arrivée θ du signal rétrodiffusé par la cible.

$$\phi = 2\pi \frac{d}{\lambda} \sin(\theta). \tag{10}$$

La détermination de ϕ nous permet d'estimer la position angulaire θ de la cible.

3 Performances du système radar FMCW

Dans notre cas, les performances radar sont définies par la résolution en distance et la résolution en fréquence de décalage. Ces résolutions dépendent de la durée du balayage, de la bande passante du balayage et du pré-traitement des données.

3.1 Résolution distance

Pour un radar homodyne FMCW, la résolution de distance (capacité du radar à discriminer 2 cibles dans l'axe distance) peut être représenté par [5]

$$\Delta R = \frac{cT}{2B} \Delta f_b \tag{11}$$

où Δf_b est la résolution de fréquence de battement qui est proportionnelle au temps d'observation du signal traité, et qui est donnée par $\Delta f_b = \frac{1}{T - \tau_{max}}$. Ainsi, la résolution peut être écrite $\Delta R = c/(2B)$ [6]. Il est clair que le temps de modulation T n'affecte pas la résolution en distance tant qu'elle reste beaucoup plus grande que τ_{max} .

3.2 Résolution de la fréquence décalage

La fréquence de décalage f_D induite par le transpondeur RFID est mesurée en prenant en compte le signal de battement sur M durées de balayage [7]. En supposant que le déplacement du transpondeur RFID soit inférieur à la résolution de distance, une procédure de transformée de Fourier discrète (TFD) est exécutée sur l'axe des temps t_l pour estimer la fréquence de décalage f_D . Ainsi, la résolution en fréquence de décalage Δf_D est inversement proportionnelle au temps d'observation du signal traité et est donnée comme suit

$$\Delta f_D = \frac{1}{MT}. \tag{12}$$

En augmentant la modulation de temps T , la résolution de fréquence de décalage diminue, ce qui signifie que cette résolution est améliorée.

3.3 Paramètres radar

Les paramètres spécifiques qui caractérisent notre système radar sont fournis dans le tableau 1. Il s'agit des principaux paramètres de la transmission radio, de la modulation, du traitement du signal, du transpondeur et de la détection CFAR. Pour étudier la performance du radar et argumenter la portée visible maximum demandée (285 m), une étude de bilan de puissance est nécessaire. Pour cela, nous avons besoin de la puissance théorique

TABLE 1 – Paramètres de simulation

Radio	Fréquence porteuse f_c	5.8 GHz
	Puissance émission P_e	25 mW PIRE
	Gains antenne G_e, G_r	12 dBi
	Facteur de bruit récepteur F	5 dB
Antenne	Type	Antenne 3D
	Distance entre antenne (d)	0.8λ
Modulation	Type	Dent de scie
	Largeur bande B	60 MHz
	Temps modulation T	80 μ s
	Nombre total M	1600
Traitement signal	Fréquence échantillonnage F_e	20 MHz
	Nb. points (distance)	1530 points
	Nb. points (fréquence décalage)	1600 points
	Temps total MT	128 ms
Transpondeur	Section efficace radar σ	-10 dBm ²
	Fréquence décalage f_D	[19 kHz - 21 kHz]
CFAR	Type bruit	Gaussien
	SNR détection	13 dB
	Probabilité de fausse alerte P_{fa}	$1e - 3$
	Probabilité de détection P_d	0.9

reçue permettant une détection fiable des cibles dans un bruit gaussien. En effet, la puissance réfléchiée du transpondeur RFID, notée P_r à l'emplacement R de l'antenne radar primaire est liée à la puissance transmise P_e et est donnée par l'équation radar suivante [4] :

$$P_r = P_e \frac{G_e G_r \lambda^2 \sigma}{(4\pi^3) R^4 L} \quad (13)$$

où G_e et G_r représentent le gain des antennes d'émission et de réception du radar primaire FMCW, $\lambda = c/f_0$ la longueur d'onde du système en espace libre, R la distance entre l'interrogateur (radar primaire) et le transpondeur (radar secondaire), σ la section équivalente radar du transpondeur considéré comme cible et L le facteur de perte. Comme le signal acquis peut être extrêmement faible, des étapes d'amplification de réception sont nécessaires pour exploiter les données reçues par le radar primaire. La chaîne d'amplification introduit du bruit qui est ajouté au signal reçu. La puissance moyenne du bruit thermique en réception est définie par :

$$P_n = K_b T_0 B_n F \quad (14)$$

où K_b est la constante de Boltzman, T_0 la température de la chaîne réceptrice, F le facteur de bruit et B_n la bande passante reçue du radar. Dans notre cas, la puissance de bruit moyenne P_n est égale à -96 dBm ($B_n = 20$ MHz, $T_0 = 290$ K et $F = 5$ dB). En présence de bruit gaussien, un rapport signal sur bruit RSB de 13 dB est requis pour obtenir une probabilité de détection P_d de 0,9 avec une probabilité de fausse alarme P_{fa} de 0,001. Le RSB à 13 dB est la valeur utilisée pour estimer la capacité de détection radar [8] dans des conditions très favorables (espace libre). Le gain de traitement du système GESTAR est fonction du nombre de points analysés N . De plus, nous appliquons une fenêtre Nuttall comme fonction de pondération sur le signal de battement défini par (5). Par conséquent, le gain de traitement est également pondéré par la bande passante de bruit équivalente normalisée (NENBW) [9], qui est égale à 1,9761 pour la fenêtre Nuttall. Ainsi, le gain de traitement est $P_{GR} = 10 \log_{10}(1530/\text{NENBW}) = 28,89$ dB pour une analyse en distance et est $P_{GD} = 10 \log_{10}(1600/\text{NENBW}) = 29,08$ dB pour l'analyse de fréquence de décalage. Par conséquent, le traitement global P_G est donné par

$$P_G = P_{GR} + P_{GD} \quad (15)$$

Enfin, la puissance reçue P_r du transpondeur RFID en dessous de laquelle la valeur de probabilité de détection est trop faible, est obtenue par

$$P_r = 13 - P_G + P_n = -140,8 \text{ dBm}. \quad (16)$$

TABLE 2 – Performances réalisées par le radar

Capacité de détection	
SNR	≥ 13 dB
Portée visible maximale	285 m
Temps détection	400 ms
Estimation distance	
Plage de mesure	[0-500 m]
Résolution	2,5 m
Résolution à -6 dB	6,56 m
Estimation de la fréquence de décalage	
Plage de mesure	[19 kHz-21 kHz]
Résolution	7,8 Hz
Résolution à -6 dB	20,52 Hz
Estimation angulaire	
Plage de mesure	de -15 deg. to 15 deg.

De (13) on peut aussi déduire la distance visible maximum correspondante R_{mvr} . Basé sur (12) la résolution de fréquence de décalage est égale à $\Delta f_D = 7,8$ Hz avec 1600 points répartis sur l'intervalle [0, 12,5 kHz] ou [12,5 kHz, 25 kHz]. La résolution de l'analyse de la fréquence de décalage est dégradée par la fenêtre de Nutall utilisée dans cette étude. Ainsi, la résolution à -6 dB est donnée par $2,627/MT = 20,52$ Hz où 2,627 correspond à la largeur normalisée du lobe principal de la fenêtre Nutall. Pour la résolution en distance, il est donné par $\Delta R = c/2B = 2,5$ m. Comme mentionné précédemment, cette résolution est dégradée par l'utilisation de fenêtres Nutall, et par conséquent, la résolution à -6 dB est donnée par $2,627 \times \Delta R = 6,567$ m. Le tableau 2 montre les performances fournies par le radar en fonction des paramètres du tableau 1.

4 Développement du système complet

4.1 Description des éléments du système GESTAR

Le radar utilise une antenne d'émission et deux antennes de réception pour permettre l'estimation de la position angulaire des cibles. Puisque la transmission et la réception sont simultanées, nous avons maintenu une isolation élevée pour atteindre de bonnes performances de détection. Les différents modules qui composent le système radar sont : un module d'alimentation, un module de traitement du signal, de modules d'émission et de réception et d'un système d'antenne. Le module de traitement de signal est implémenté dans une carte FPGA, chargée de générer le signal radar transmis par le module radio, de traiter le signal reçu par les deux chaînes de réception radio pour détecter la présence des transpondeurs, d'identifier leur fréquence de décalage, de les localiser en estimant leur distance radiale et leur position angulaire et d'afficher les informations correspondantes sur l'écran de la diode électroluminescente organique (OLED) du radar (voir figure 3). La mise en œuvre de l'analyse



FIGURE 3 – Le radar conçu équipé d'un écran OLED.

spectrale du signal de battement est effectuée via une transformée de type *Chirp Z-transform* (CZT) qui est la transformée en Z d'un signal le long d'un contour en spirale [10, 11]. L'avantage de la CZT est d'évaluer la fréquence du signal de battement uniquement dans une zone d'intérêt $[F_{min}, F_{max}]$. L'utilisateur peut définir le nombre de points d'analyse permettant d'améliorer la résolution en fréquence. Nous illustrons sur la figure 4 les différentes étapes pour obtenir une cartographie de fréquence à distance de distance à partir des échantillons du signal de battement donné par 5. Ces échantillons sont collectés et positionnés dans une matrice. La première

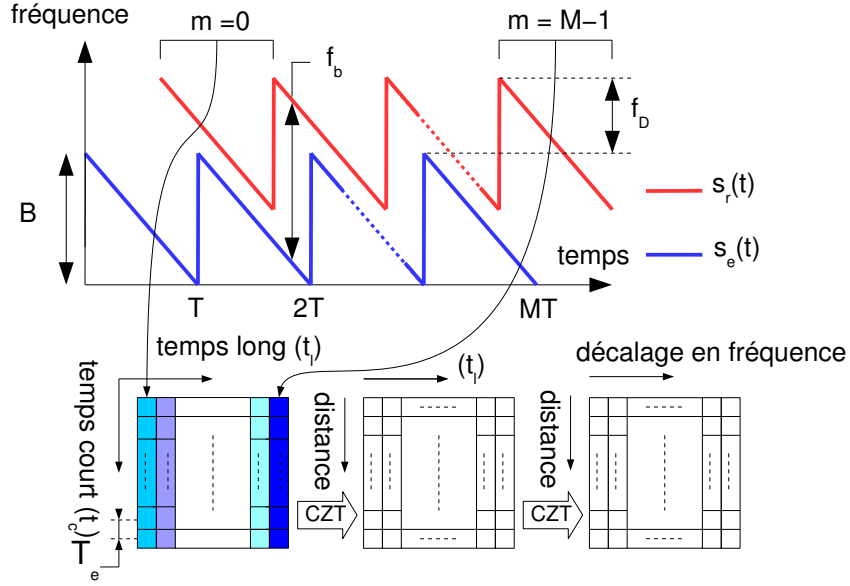


FIGURE 4 – Obtention d’une carte distance-décalage fréquence sur M rampes.

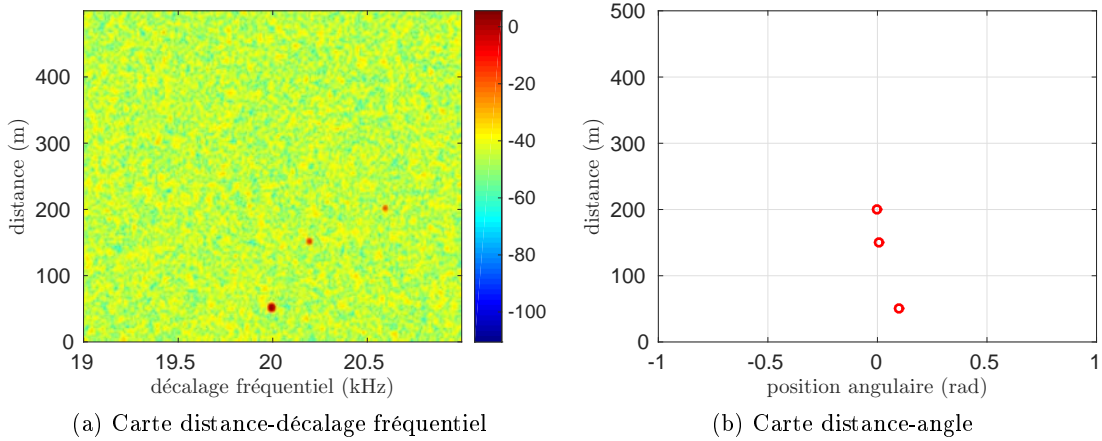


FIGURE 5 – Simulation : exemple de détection de 3 cibles avec des distances $[50, 150, 200]$ m avec des fréquences décalage respectives $[20, 20,2; 20,6]$ kHz. Les cibles sont orientées à des angles mesurés dans le plan horizontal $\theta = [\pi/30; \pi/20; 0]$.

CZT est appliquée pour chaque colonne pour former une matrice de profil de distance. Le numéro de colonne k de cette nouvelle matrice représente le profil de distance du nombre de chirp k . La seconde CZT est appliquée à chaque ligne pour le traitement de la fréquence de décalage. La sortie de cette seconde CZT fournit à l’opérateur radar une cartographie distance-fréquence de décalage de la scène.

5 Résultats de simulations et de mesures

Des simulations de carte distance-décalage fréquence sont réalisées et sont en accord avec les résultats attendus en théorie. La figure 5 montre un exemple de simulation. De plus, la réalisation du radar interrogateur et du transpondeur permet des mesures réelles en espace libre (indoor et outdoor). Elles sont menées avec différentes distances et angles d’orientation pour le transpondeur. Les résultats obtenus respectent les performances exigées du système. La figure 6 montre un exemple de mesure réelle d’une cible distante de 25 m du radar principal.

6 Conclusion

Dans cette communication, un système radar FMCW pour l’identification RFID est présenté. La carte distance-décalage fréquence et la carte distance-angle sont réalisées à partir de l’analyse spectrale du signal de battement. Nous définissons un transpondeur linéaire lorsque la fréquence de décalage est induite linéairement par le signal transmis au système de lecture. Le projet est validé à partir de données radar FMCW simulées et réelles.

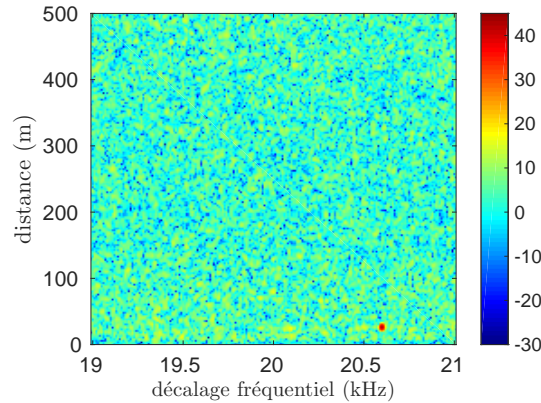


FIGURE 6 – Mesure réelle : carte distance-décalage fréquence établie pour une seule cible à une distance $R = 25$ m du radar avec une fréquence de décalage du transpondeur $f_D = 20,6$ kHz et avec un angle $\theta = 0$.

En perspective, nous projetons l'utilisation d'autres formes d'onde émises comme onde continue à fréquence de pas (SFCW) afin d'améliorer le rapport signal sur bruit. Une autre interrogation demeure concernant un comportement non-linéaire du transpondeur.

7 Références

- [1] Desa Hazry, Rosbi Mohd Sofian Muhammad, and Shaiful Zairi, "Study of integration 2.4GHz and 5.8GHz in RFID tag," *Proceedings of the International Conference on Man-Machine Systems (ICoMMS)*, 10 2009.
- [2] J. W. Ting, D. Oloumi, and K. Rambabu, "FMCW SAR System for Near-Distance Imaging Applications—Practical Considerations and Calibrations," *IEEE Transactions on Microwave Theory and Techniques*, vol. PP, no. 99, pp. 1–12, 2017.
- [3] L. Ghattas, S. Méric, and C. Brousseau, "Performance assessment of FMCW radar processing for transponder identification," in *2016 European Radar Conference (EuRAD)*, Oct 2016, pp. 258–261.
- [4] M. Skolnik, *Radar Handbook*, McGraw-Hill, third edition,, 2008.
- [5] S.O. Piper, "Receiver frequency resolution for range resolution in homodyne FMCW radar," in *Conference Proceedings National Telesystems Conference 1993*, Jun 1993, pp. 169–173.
- [6] F. Le Chevalier, *Principles of Radar and Sonar Signal Processing*, Artech House, Norwood, MA,, 2002.
- [7] A.G. Stove, "Linear FMCW radar techniques," *IEE Proceedings F (Radar and Signal Processing)*, vol. 139, pp. 343–350(7), October 1992.
- [8] M. Skolnik, *Introduction to Radar Systems*, McGraw-Hill, New-York,, 1980.
- [9] F. J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform," *Proceedings of the IEEE*, vol. 66, no. 1, pp. 51–83, Jan 1978.
- [10] L. Rabiner, "The chirp z-transform algorithm—a lesson in serendipity," *IEEE Signal Processing Magazine*, vol. 21, no. 2, pp. 118–119, Mar 2004.
- [11] L. Rabiner, R. Schafer, and C. Rader, "The chirp z-transform algorithm," *IEEE Transactions on Audio and Electroacoustics*, vol. 17, no. 2, pp. 86–92, June 1969.

Amélioration de la résolution angulaire d'un radar FMCW panoramique

R.Fouettilloux^{1,2}, O.Lacrouts¹, G.Motyl¹, L.Ferro-Famil², S.Méric²

¹DIADES MARINE, 70 rue Jean Doucet Pôle d'activité du Grand Girac 16470 SAINT MICHEL

²Université de RENNES 1, IETR Campus de Beaulieu 263 avenue du Général Leclerc 35042 Rennes CEDEX

Mots-clés: RADAR, technologie FMCW, déconvolution, problème inverse, Tikhonov, Richardson-Lucy, ISRA, BBSMA (Blind BeamSpace Music Algoritm) . . .

Keywords: RADAR, FMCW technology, deconvolution, inverse problem, Tikhonov, Richardson-Lucy, ISRA, BBSMA (Blind BeamSpace Music Algoritm) . . .

Résumé:

Pour des capteurs d'imagerie, la résolution angulaire θ est limitée par la dimension d'ouverture D de l'antenne à la valeur $\theta \propto \lambda/D$ où λ est la longueur d'onde. En effet, l'image radar acquise étant le résultat de la convolution de la densité de réflectivité de la scène radar par le diagramme de rayonnement de l'antenne ayant des dimensions finies devient alors floue. Néanmoins, la résolution de cette image peut être améliorée en déconvoluant celle-ci par le diagramme d'antenne. Cette déconvolution, ou filtrage inverse, est intrinsèquement numériquement instable. Notre contribution consiste à adapter des méthodes pour obtenir des résultats stables présentant une résolution et discrimination angulaire améliorées. Dans la version étendue, des exemples de simulation radar en onde centimétrique ainsi que des mesures de terrain sont présentés. Pour terminer un nouvel algorithme BBSMA (Blind BeamSpace Music Algorithm) sera démontré.

Abstract: For imaging sensors, the angular resolution θ is limited by the aperture dimension D of the antenna to the value $\theta \propto \lambda/D$ where λ is the wavelength. Indeed, the acquired radar image being the result of the convolution of the reflectivity density of the radar scene by the radiation pattern of the antenna having finite dimensions then becomes fuzzy. Nevertheless, the resolution of this image can be improved by deconvoluting it by the antenna pattern. This deconvolution, or inverse filtering, is intrinsically numerically unstable. Our contribution is to adapt methods to obtain stable results with improved resolution and angular discrimination. Examples of centimeter wave radar simulation as well as field measurements are presented. To complete a new BBSMA (Blind BeamSpace Music Algorithm) algorithm will be demonstrated.

1 Introduction

Cet article se place dans un contexte de conception d'un radar à onde continue modulée en fréquence FMCW pour améliorer la détection, la discrimination et le pistage de petites cibles dans le fouillis de mer. Les portées instrumentées de ces types de radar sont en général de l'ordre 48 NM (89 Km). Ce radar devra posséder une antenne dont son angle d'ouverture devra être inférieur à 2 degrés et devra être capable de discriminer angulairement deux bouées séparées de 2.5 degrés à 1 NM (radar certifiable IMO). Pour respecter ces critères tout en gardant un capteur compact et léger, une solution pour améliorer la résolution et la discrimination angulaire de ce radar par post-traitement sur l'image radar est présentée. Dans un premier temps, cet article reprend le principe de convolution, puis présente des méthodes de déconvolution.

2 Principe de convolution

L'écho en azimut peut être modélisé, pour une position en distance donnée, comme la convolution de la densité de réflectivité de la scène observée $x(\theta)$, par le diagramme de rayonnement de l'antenne utilisée, $h(\theta)$ (voir fig. 1) et nous avons :

$$y(\theta) = h(\theta) * x(\theta) + n(\theta) \quad (1)$$

où $n(\theta)$ représente le bruit de mesure indépendant de $x(\theta)$ et $y(\theta)$ est le signal mesuré par le radar à la position angulaire θ .

Les données enregistrées à la sortie du système radar sont une version filtrée passe-bas de la scène d'origine en raison de la taille finie de l'antenne [1]. Les données filtrées correspondent aux composantes spectrales hautes

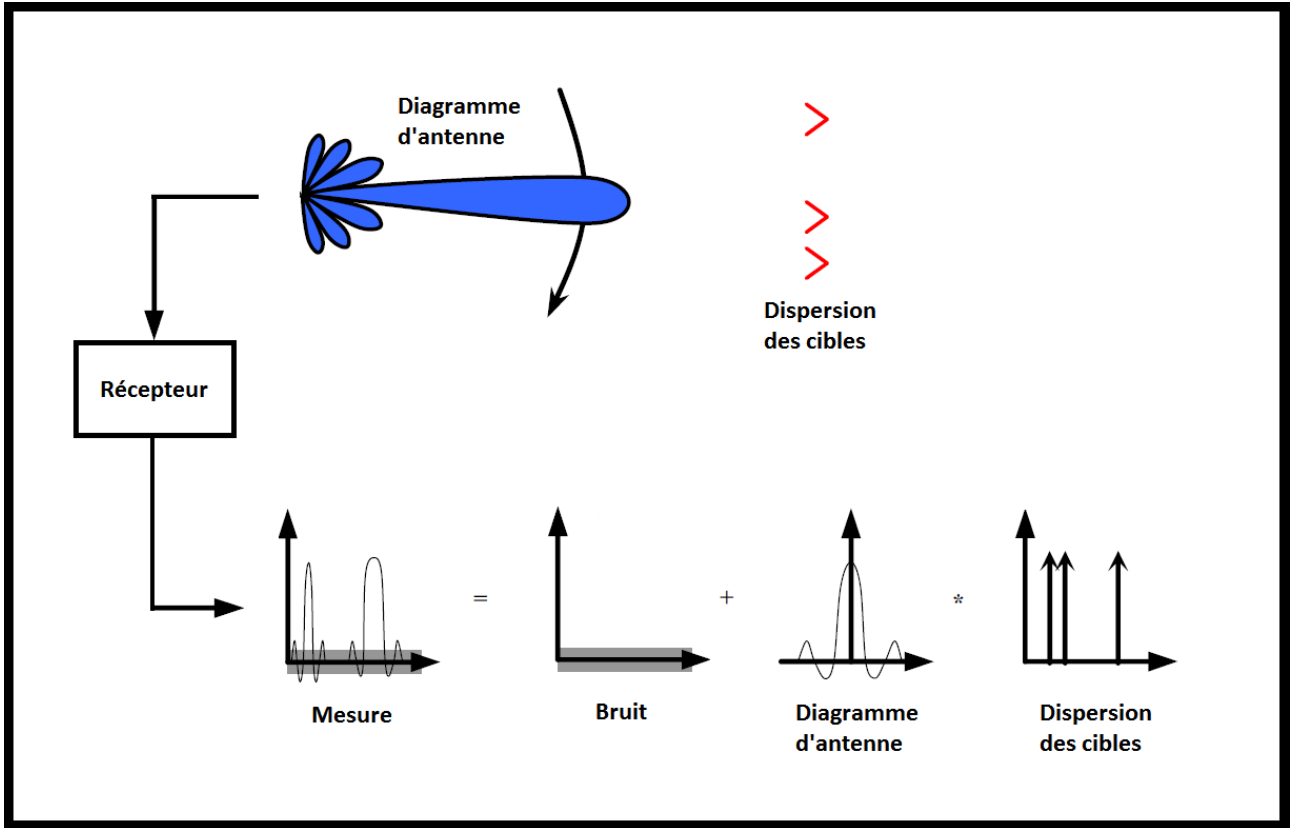


FIGURE 1 – Configuration de mesure.

fréquences ce qui génère un flou sur l'image radar. Une opération de déconvolution peut alors être effectuée pour réduire ce flou et améliorer la résolution.

3 Opération de déconvolution

En l'absence de bruit ($n(\theta) = 0$), l'estimation de x à partir de y peut être effectuée à l'aide d'une simple division dans le domaine spectral. Néanmoins, en présence de bruit, ce problème inverse mal conditionné conduit à une variance de bruit en sortie qui dépend du diagramme de rayonnement et qui peut atteindre des valeurs importantes. La déconvolution par filtrage linéaire peut ainsi être reformulée comme :

$$\hat{x}(\theta) = f(\theta) * y(\theta) \quad (2)$$

Le but ici est de trouver une fonction $f(\theta)$ permettant de donner la meilleure estimation possible $\hat{x}(\theta)$ à partir du signal mesuré $y(\theta)$. On note $[\mathbf{x}]_j$, la réflectivité de la scène à l'angle azimutal θ_j et $[\mathbf{y}]_i$ représente la réponse bruitée de la scène complète, mesurée pour la position du radar θ_i . L'expression du signal mesuré devient :

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \text{ avec } [\mathbf{H}]_{ij} = h(\theta_i - \theta_j) \quad (3)$$

Ce papier présente 2 méthodes de déconvolution permettant de préciser la position et le nombre de cibles radar.

3.1 Méthode de la régularisation de Tikhonov

Une première méthode est la régularisation de Tikhonov [2] qui consiste à régulariser l'approche des moindres carrés et permet de mieux conditionner la technique de déconvolution, en maintenant la variance de $\hat{\mathbf{x}}$ à un faible niveau.

$$J_T(\mathbf{x}) = \|\mathbf{H}\mathbf{x} - \mathbf{y}\|_2 + \lambda\|\mathbf{x}\|_2 \quad (4)$$

et

$$\hat{\mathbf{x}}_T = (\mathbf{H}^T\mathbf{H} + \lambda I)^{-1}\mathbf{H}^T\mathbf{y} \quad (5)$$

avec \mathbf{H}^T , la transposée de \mathbf{H} . Néanmoins la régularisation de Tikhonov présente un paramètre de régularisation λ dont le réglage efficace requiert un fort rapport signal à bruit. Un certain nombre de méthodes permettent d'estimer automatiquement la valeur que doit prendre ce paramètre pour réaliser le meilleur compromis entre le niveau du bruit et la fidélité des données (méthode L curves) [3].

3.2 Méthode de Richardson-Lucy

La méthode de Richardson-Lucy est une approche de super-résolution estimant \mathbf{x} selon le principe du maximum de vraisemblance basée sur la règle de Bayes. Il repose sur le principe de l'estimateur MAP (maximum a posteriori) qui résout le problème suivant [4] :

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) = \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \quad (6)$$

où $p(\mathbf{x})$ et $p(\mathbf{x}|\mathbf{y})$ représentent la densité de probabilité a priori et a posteriori de la réflectivité des cibles et $p(\mathbf{y}|\mathbf{x})$ est la fonction de vraisemblance associée.

L'algorithme Richardson-Lucy est basé sur l'hypothèse que le bruit de mesure est indépendant dans les cellules observées et obéit à la distribution de Poisson. La distribution a priori des réflectivités étant inconnue, on la suppose uniforme et l'on remplace l'estimation selon le maximum a posteriori (MAP) de (6) par une approche utilisant le maximum de vraisemblance (ML) :

$$\hat{\mathbf{x}} = \arg \max_{\mathbf{x}} p(\mathbf{x}|\mathbf{y}) = \arg \max_{\mathbf{x}} p(\mathbf{y}|\mathbf{x}) \quad (7)$$

Or :

$$p(y/x) = \prod_{i=1}^M \frac{(Hx)_i^{y_i} * e^{-(Hx)_i}}{y_i!} \quad (8)$$

où i est la cellule en azimuth et avec $(Hx)_i = \sum_{j=1}^M h_{ij}x_j$

Le critère devient donc :

$$\begin{aligned} J_{RL}(x) &= -\log p(y/x) \\ &= \sum_{i=1}^M (Hx)_i + \log(y_i!) - y_i \log[(Hx)_i] \end{aligned} \quad (9)$$

La déconvolution, ou optimisation de (7), revient donc à déterminer $\hat{\mathbf{x}}$, qui minimise la log-vraisemblance $J_{RL}(\mathbf{x}) = -\log p(\mathbf{y}|\mathbf{x})$. L'annulation du gradient de $J_{RL}(\mathbf{x})$ conduit à une expression récursive que l'on résout simplement de manière itérative :

$$\hat{\mathbf{x}}_{i+1} = \text{diag}(\mathbf{H}^T \text{diag}^{-1}(\mathbf{H}\hat{\mathbf{x}}_i)\mathbf{y}) \hat{\mathbf{x}}_i \quad (10)$$

4 Résultats de simulation

Les paramètres de la scène et du radar simulé sont répertoriés dans le tableau 1. Ces 2 méthodes sont testées sur

Paramètres du radar et de la scène observée	
vitesse de rotation du radar	120 degrés/s (20 t/min)
PRF	2000Hz
Nombre d'antennes (réseau)	50 ($\theta_{-3dB} = 2$ degrés)
Largeur angulaire des cibles	0.3 degrés
Scène observée	+/-12 degrés
Rapport signal à bruit	25 dB

TABLE 1 – Paramètres du radar et scène

une scène radar simulée pour déterminer deux critères : la résolution et la discrimination angulaire. Deux cibles espacées 1.26 degrés (fig. 2) ainsi qu'une troisième cible seule forment cette scène. La courbe bleu représente la scène radar idéale et la courbe verte représente le signal de sortie du radar. Les courbes rouge et jaune sont les signaux de sortie après application des algorithmes.

Ces résultats montrent qu'il est totalement impossible de discriminer deux cibles sans déconvolution (courbe verte). La méthode de la régularisation de Tikhonov permet de discriminer les deux cibles seulement avec un paramètre de régularisation faible. La méthode R-L permet de mieux discriminer les deux cibles et un ratio de super-résolution de $2/1.26 = 1.59$ est obtenu.

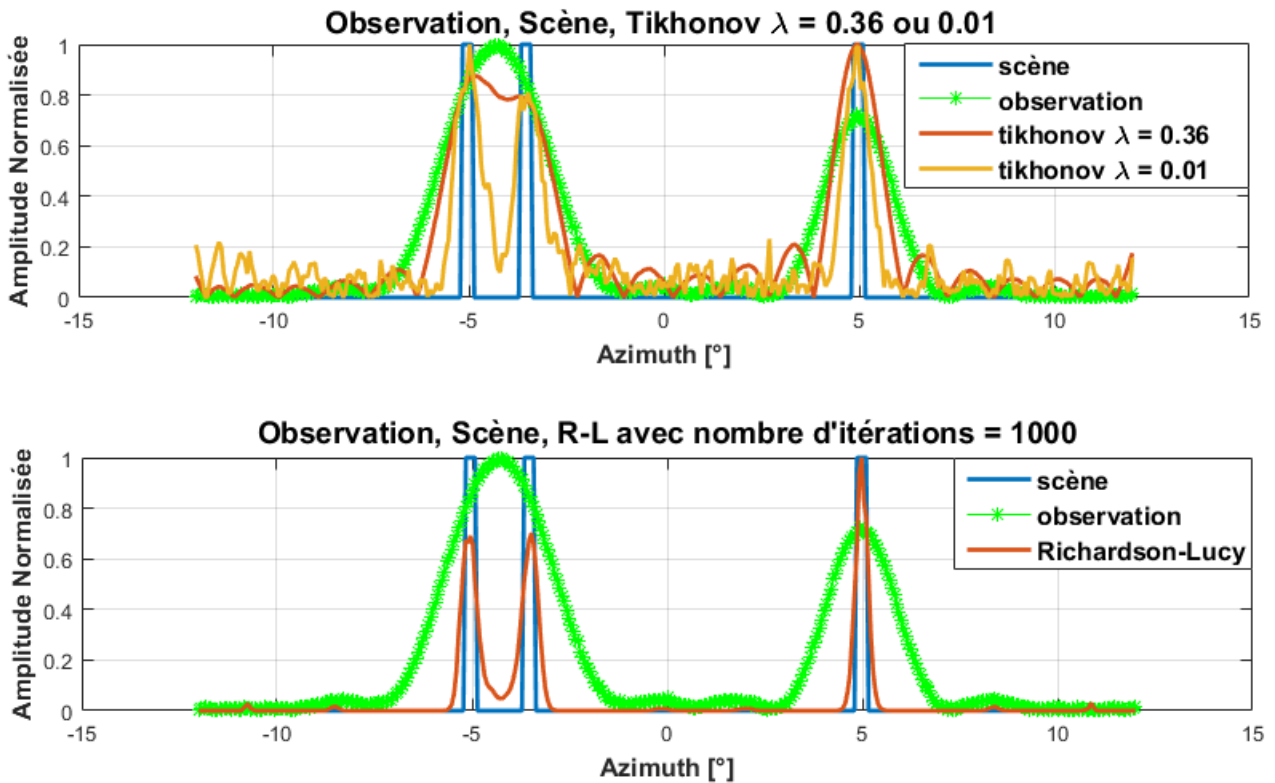


FIGURE 2 – Cibles espacées de 1.26 degrés

5 Conclusion

Dans la partie étendue, il serait intéressant de présenter un nouvel algorithme nommé BBSMA qui permet, sans connaître le diagramme de rayonnement d’un réseau d’antennes panoramique et en travaillant dans le domaine doppler, d’appliquer la décomposition en espace signal et bruit d’une scène radar Music). Cet algorithme est plus rapide à exécuter que la méthode R-L.

6 Remerciements

Les auteurs tiennent à remercier la DGA qui participe à ce projet par l’intermédiaire du financement de la thèse CIFRE Défense numéro 009/2015/DGA.

7 Références

- [1] M. A. Richards, “Iterative noncoherent angular superresolution [radar],” in *Radar Conference, 1988., Proceedings of the 1988 IEEE National.* IEEE, 1988, pp. 100–105.
- [2] A. N. Tikhonov, “Regularization of incorrectly posed problems.” SOVIET MATHEMATICS DOKLADY, 1963.
- [3] P. C. Hansen and D. P. O Leary, “The use of the l-curve in the regularization of discrete ill-posed problems,” *SIAM Journal on Scientific Computing*, vol. 14, no. 6, pp. 1487–1503, 1993.
- [4] Y. Zha, Y. Huang, Z. Sun, Y. Wang, and J. Yang, “Bayesian deconvolution for angular super-resolution in forward-looking scanning radar,” *Sensors*, vol. 15, no. 3, pp. 6924–6946, 2015.

GNSS

GNSS : une révolution pour le géopositionnement précis *GNSS: a revolution for geopositioning*

Félix Perosanz

GET/CNES, Felix.Perosanz@cnes.fr

Mots clés : GNSS, IGS, SSR, OSR

Key words: GNSS, IGS, SSR, OSR

Résumé

Les systèmes globaux de navigation par satellite (GNSS) n'avaient pas à l'origine de vocation à répondre aux besoins de positionnement pour des applications scientifiques précises. Mais un usage détourné des signaux a permis d'accéder à des mesures de distance satellite-récepteur d'une précision millimétrique. Si la technique différentielle s'est imposée comme une méthode incontournable, le mode PPP est aujourd'hui une alternative de plus en plus populaire. Mais quelle que soit l'approche, les GNSS ont révolutionné la réalisation des repères d'espace et de temps ainsi que les nombreux usages nécessitant en temps différé ou en temps réel un positionnement précis.

Abstract

Global Navigation Satellite Systems (GNSS) were not initially designed to be used for precise scientific applications. However a smart signal processing can provide a satellite-receiver range measurement with millimeter accuracy. If the differential technique has become an essential method, PPP mode is today an increasingly popular alternative. But whatever the approach, GNSS have revolutionized the realization of space and time frames and the many uses requiring post-processed or real-time precise positioning.

1 Les GNSS : une technique géodésique d'opportunité

Les systèmes globaux de navigation par satellite (GNSS) tel le populaire GPS des USA opérationnel depuis 1995 n'avaient pas à l'origine de vocation à être exploités pour des usages scientifiques précis. Leurs signaux radioélectriques sont basés sur la modulation de la phase d'une onde porteuse en bande L (1-2 GHz) par un code PRN d'une fréquence proche de 1 MHz. La démodulation de ce code par un récepteur permet de mesurer la distance entre les satellites émetteurs et l'utilisateur avec une précision de quelques mètres. Les récepteurs étant multicanaux, l'utilisateur peut déduire des mesures simultanées sur plusieurs satellites sa propre position par « tri-latération » avec cette même précision. Mais rapidement l'idée d'un usage détourné de ces signaux en mesurant directement la phase de l'onde porteuse est apparue. Avec une fréquence 1000 fois supérieure, c'est à une mesure d'une précision millimétrique que les utilisateurs équipés de récepteurs spécifiques ont eu accès. Les scientifiques ont depuis largement tiré profit de ces mesures providentielles en provenance d'un nombre sans cesse croissant de constellations de satellites comme GLONASS (Russie), Galileo (Europe), ou Beidou (Chine). Les mesures GNSS sont devenues de façon opportune une technique géodésique précise et largement exploitées pour de nombreux usages scientifiques. Dans un premier temps cet article compare les techniques différentielles et PPP du point de vue de la corrélation des paramètres. L'apport essentiel des GNSS pour la réalisation des repères d'espace et de temps est ensuite rappelé ainsi que le rôle moteur du service International IGS qui a fait évoluer les techniques GNSS vers plus de précision et moins de délais.

2 Le GNSS précis : une question de (dé)corrélation

L'exploitation de la précision millimétrique des mesures de phase n'est pas immédiate car elle doit passer par une modélisation très fine de chacun des termes qui la composent :

- la position du centre de masse du satellite émetteur à l'instant d'émission
- le vecteur centre de masse du satellite centre de phase de l'antenne émettrice qui est fonction de la géométrie et de l'attitude (orientation dans l'espace) du satellite
- le biais d'horloge du satellite par rapport à l'échelle de temps de la constellation
- la position du centre de phase de l'antenne réceptrice à l'instant de réception

- le vecteur « excentricité » entre le centre de phase récepteur et le point à positionner
- les délais de propagation dans l'ionosphère et la troposphère
- les termes relativistes
- l'effet de « phase wind-up » dû à la rotation relative entre l'émetteur et le récepteur
- le biais de mesure de phase correspondant au nombre entier de cycles initial inconnu au démarrage de l'acquisition des données d'un satellite. Ce biais appelé « ambiguïté » est commun à toutes les mesures de phase d'un même passage d'un satellite (entre son lever et son coucher) sauf en cas d'interruption du signal
- des biais « hardware » au niveau du satellite et du récepteur qui font perdre leur nature entière à l'ambiguïté de phase

La clé d'un usage « précis » des données GNSS passe par une maîtrise de ces différents termes. Le problème mathématique à résoudre est généralement surdimensionné c'est-à-dire que le nombre d'observations en provenance des satellites en visibilité est supérieur au nombre de paramètres à déterminer. Cependant une solution « fiable » ne sera obtenue que si un niveau de décorrélation suffisant entre les paramètres est atteint. Concrètement cela signifie qu'un nombre suffisant d'équations (c'est-à-dire d'époques d'observations) devra être accumulé pour séparer les contributions des termes d'horloge récepteur, coordonnées, allongement troposphérique et ambiguïté de phase qui sont au départ corrélés (ou anti-corrélés) à plus de 90%. Naturellement la situation est sensiblement plus favorable dans le cas d'un post-traitement où l'ensemble du système d'équations peut être résolu globalement a posteriori. Mais que ce soit pour des applications en temps réel ou en temps différé deux types de stratégies ont été adoptées au cours du temps pour résoudre ce problème: l'approche différentielle et l'approche PPP.

3 L'approche différentielle

L'idée de base est de créer de nouvelles observables à partir de la différence de mesures de phase issues de deux récepteurs proches et/ou de deux satellites en visibilité simultanée. Les termes affectant de façon sensiblement identique les mesures sont ainsi éliminés. La différentiation des mesures de phase GNSS est une forme de décorrélation « ultime ». En effet, plusieurs termes du modèle de mesure comme les biais d'horloge et hardware ne sont plus des inconnues à résoudre. Sur ce principe des réseaux denses de stations GNSS de référence (de position connues) ont été déployés dans de nombreux pays pour fournir un service de « corrections différentielles » destiné aux usagers nécessitant une précision centimétrique (géomètres, topographes, agriculteurs,...). En France, Teria, Orpheon ou Satinfo par exemple, fournissent ce type de service. On parle dans ce cas d'une représentation OSR (Observation State Representation) pour laquelle un seul terme correspondant à la somme des corrections à appliquer aux observations est transmis aux usagers. Dans le cas du post-traitement et de récepteurs « immobiles » le rapport nombre d'observations sur nombre d'inconnues devient très favorable et la technique différentielle permet d'atteindre des précisions millimétriques. Elle s'est imposée dès les années 1990 comme méthode incontournable pour tous les usages précis des GNSS.

4 Le mode PPP

Une approche alternative à la technique différentielle consiste à modéliser ou estimer chaque terme composant les mesures de phase plutôt que de chercher à les éliminer. On parle dans ce cas de représentation SSR (State Space Representation). Le PPP (Positionnement Ponctuel Précis) d'un récepteur GNSS sans nécessiter de récepteur de référence à proximité devient alors possible. Cependant l'accès à une source de corrections précises d'orbite et d'horloge des satellites par exemple est nécessaire. De plus, d'un point de vue des corrélations la situation est moins favorable car le nombre de paramètres à estimer (un biais d'horloge par époque de mesures par exemple) est nettement plus important. Ainsi, le positionnement d'un mobile en temps réel va nécessiter un temps de convergence plus long que dans le cas d'une approche différentielle. [1] et [2] détaillent les avantages et les inconvénients des approches différentielles et PPP ainsi que des représentations OSR et SSR.

5 La fixation des ambiguïtés de phase

L'exploitation ultime des mesures de phase GNSS passe par l'estimation du biais d'ambiguïté entière. La mesure GNSS devient alors précise et exacte, le nombre de paramètres dans le problème chute et ceux restants sont mieux décorrélés. Comme les biais « hardware » satellite et récepteur font perdre leur nature entière aux ambiguïtés de phase, la seule solution envisagée pendant de nombreuses années consistait à les éliminer par double-différentiation. En 2007 [3] observe que les biais satellites étant stables (constants sur une journée), ils peuvent être décorrélés du biais d'un récepteur pourvu que celui-ci observe simultanément un nombre suffisant

de satellites. La résolution des ambiguïtés entières en mode PPP devient alors possible. Ce mode de traitement appelé « Integer-PPP » (IPPP) ou « PPP-AR » permet de positionner un récepteur GNSS « isolé » avec une précision améliorée et devient une technique concurrente du mode différentiel. L'idée que le système Galileo dispose d'un service libre de diffusion de corrections SSR de ses propres signaux est à l'étude.

6 Le service international IGS

La décision de l'Association Internationale de Géodésie de créer en 1994 l'International GNSS Service est un acte fondateur. Ce service est basé sur un effort international de déploiement d'un réseau de stations géodésiques permanentes. La contribution française via le réseau REGINA [4] composé d'une quarantaine de stations du CNES et de l'IGN réparties sur toute la planète, répond à ce besoin et à la recommandation 69/L.53 des Nations Unies sur le « repère de référence géodésiques mondial pour le développement durable ». L'IGS dispose aussi de centres d'archivage de données et de centres d'analyses qui les traitent (calcul des orbites précises par exemple). Plus de 200 organismes contribuent volontairement à l'IGS avec l'objectif de fournir librement les produits GNSS les plus précis (IGS 2017). De par la « compétition amicale » que se livrent les différents centres d'analyses et le dynamisme de ses « Working Groups », l'IGS a progressivement imposé les GNSS comme une technique géodésique majeure en fournissant des produits de plus en plus précis, en diffusant des logiciels de traitement (GAMIT, Bernese, GIPSY, GINS,...), en imposant des standards (RINEX, SINEX, RTCM,...), en publiant des centaines d'articles. Si l'objectif premier était et reste de contribuer à l'amélioration de la réalisation des repères d'espace et de temps (cf. paragraphe suivant) et de mesurer la cinématique terrestre (cf. article de Pierre Briole) l'héritage de l'IGS est considérable et les usages précis des GNSS se sont étendus à des domaines aussi variés que l'océanographie, l'hydrologie, la météorologie terrestre et spatiale par exemple. Depuis 2010, dans le cadre d'un projet pilote qui a pour vocation à devenir un service « opérationnel », des produits sont calculés et délivrés en temps réel. Ainsi l'IGS a été le terreau de nombreuse innovations et développements qui ont considérablement amélioré la qualité des traitements et ouvert le champ à de multiples usages des GNSS.

7 Les GNSS incontournables pour la réalisation des références d'espace et de temps

L'ITRF (International Terrestrial Reference Frame) 2014 est aujourd'hui le standard international pour exprimer des coordonnées dans un repère lié à [la croûte ?] terrestre. L'ITRF est défini par un jeu de coordonnées et de vitesse d'un réseau de stations géodésiques de références qui d'une certaine façon matérialisent le repère. De la qualité de la détermination de ces coordonnées et vitesse dépend la qualité de la réalisation du système de référence lui-même. Celle-ci est critique par exemple lorsqu'il s'agit de surveiller le niveau des océans qui s'élèvent en moyenne de 3 mm/an. Il faut aussi noter que la précision d'un ITRF se dégrade avec le temps principalement du fait des incertitudes sur les vitesses des stations de référence. La technique GNSS (plus économique et facile à mettre en œuvre en comparaison aux techniques SLR, DORIS ou VLBI) représente la grande majorité des 1000 sites qui composent ce réseau (figure 1).

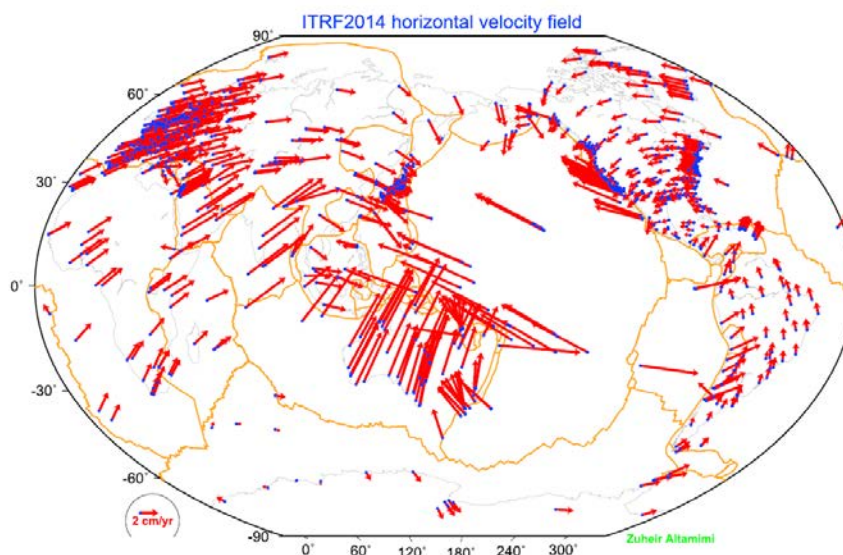


Figure 1 : Champ des vitesses horizontales de l'ITRF2014 [5]

Le passage du système de référence céleste (dans lequel sont calculées les trajectoires des satellites) au repère terrestre (dans lequel sont exprimées les coordonnées des stations) est paramétré par 5 quantités : les angles de précession et de nutation (issus de modèles), les coordonnées X_p et Y_p du pôle et l'écart UT1-UTC. Ce dernier

(qui représente les irrégularités de la vitesse de rotation de la Terre) ne peut être observé directement que par la technique VLBI qui a pour cible non pas des satellites mais des objets célestes (quasars). Le mouvement du pôle quant à lui est estimé par méthode inverse. Tout écart de XP et YP produit une oscillation journalière des coordonnées des stations dont l'amplitude et la phase dépendent de la position de la station. Ainsi les stations géodésiques GNSS de par leur nombre permettent aujourd'hui de mesurer Xp et Yp avec une précision remarquable de 0.2 milliseconde de degré (mas).

La réalisation des échelles de temps nécessite la comparaison et la combinaison des observations d'un réseau mondial de plus de 200 horloges atomiques (<https://www.bipm.org/en/bipm/tai/>). Cette tâche fondamentale repose aujourd'hui de façon essentielle sur la technique GNSS car elle permet d'estimer avec une grande précision l'écart à une référence commune de ces horloges lorsqu'elles sont connectées à un récepteur GNSS de position connue. La technique PPP qui ne nécessite pas que ces stations soient en vue commune des satellites, est largement utilisée par les laboratoires contribuant à la réalisation des échelles de temps. Récemment, [6] a montré que les solutions IPPP, en assurant la continuité des ambiguïtés entières entre passages de satellites successifs, permet un gain de précision sensible. Ainsi que ce soit pour la réalisation des références d'espace et de temps ou la détermination du mouvement du pôle, les GNSS ont permis des avancées considérables et sont devenus incontournables.

8 De la géodésie à la navigation centimétrique

Si à la création de l'IGS, 3 ordres de grandeur de précision séparaient le monde des géosciences (positionnement statique en temps différé) du monde de la navigation (positionnement en temps réel) cette situation a depuis considérablement évolué. La densification des réseaux, les modernisations à la fois des systèmes eux-mêmes, des équipements et des moyens de communication et de calcul, ont fait que les délais pour positionner un récepteur GNSS se sont réduits et que la précision s'est améliorée. Les coordonnées GNSS des stations géodésiques étaient déterminées de façon hebdomadaire pour la réalisation de l'ITRF2008 par exemple. Aujourd'hui la cadence est journalière et diverses études sont menées sur la modélisation et la mesure des mouvements sub-diurnes de la Terre. [7] et [8] par exemple ont pu observer respectivement les déformations de charge océaniques à haute fréquence et des modes propres de vibration d'un glacier grâce au positionnement précis sub-diurne de stations GPS. Pour le calcul d'orbite précis de satellites embarquant un récepteur GNSS l'approche différenciée n'est pas adaptée car la vue commune entre une station terrestre et un satellite en orbite basse est très limitée. [9] avait démontré l'intérêt de la technique IPPP pour ce type de problème en déterminant avec une précision centimétrique la trajectoire des satellites GRACE et JASON-1 en temps différé. En utilisant cette même technique Fund [10] a montré que l'on pouvait calculer avec une précision similaire la trajectoire d'une bouée océanographique (en post-traitement). Au CNES, le projet PPP-Wizard (<http://www.ppp-wizard.net/>) a pour objectif de démontrer la faisabilité d'une navigation (en temps réel) en mode IPPP (centimétrique). Ce service est basé sur le calcul en temps réel des orbites précises des satellites GNSS, puis des corrections SSR et enfin de la position corrigée du récepteur (figure 2).

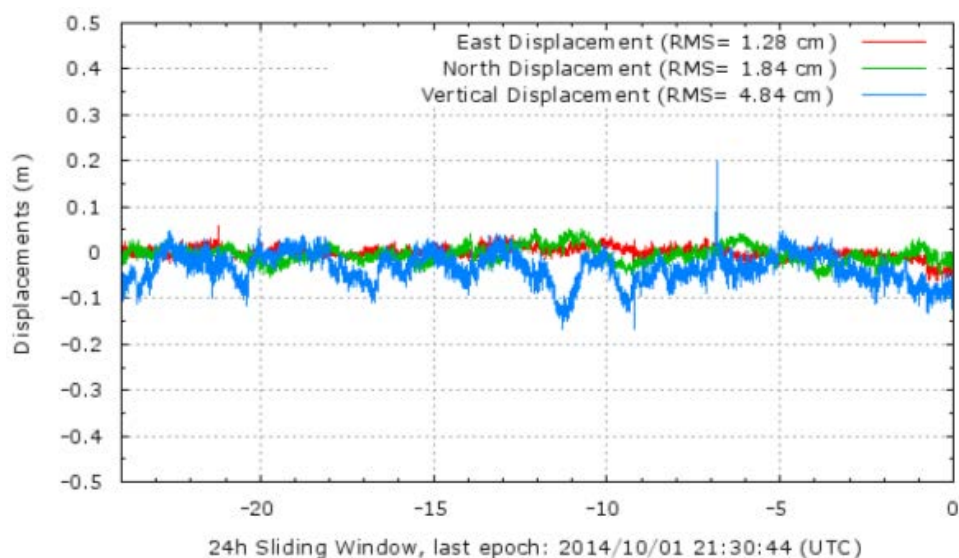


Figure 2 : Coordonnées temps réel du récepteur (fixe) KRGG (démonstrateur PPP-Wizard)

C'est sur ce principe que par exemple l'application pour smartphones « PPP-Wizlite » développée par le CNES [11] devrait permettre prochainement de se localiser instantanément avec une précision 10 à 100 fois meilleure que celle pour lesquelles les GNSS ont été spécifiés initialement.

Ainsi au-delà des enjeux économiques, technologiques ou de souveraineté les GNSS en révolutionnant le géopositionnement précis en temps différé et en temps réel ont un impact scientifique et sociétal majeur.

Références bibliographiques

- [1] G. Wubben, M. Schmitz, A. Bagge, PPP-RTK: Precise Point Positioning Using State-Space Representation in RTK Networks, ION GNSS 18th International Technical Meeting of the Satellite Division, 13-16 September 2005, Long Beach, CA
- [2] S. Choy, S. Bisnath, C. Rizos, Uncovering Common Misconceptions in GNSS Precise Point Positioning and its Future Prospect, International Federation of Surveyors, Article of the Month – September 2016
- [3] D. Laurichesse D, F. Mercier, Integer ambiguity resolution on undifferenced GPS phase measurements and its application to PPP. ION GNSS 2007 20th International Technical Meeting of the Satellite Division, 25-28 September 2007, Fort Worth, TX
- [4] C. Velut, L. Jolivet, B. Garayt, T. Donal, J.P. Cardaliaguet. Regina, A CNES/IGN worldwide GNSS receivers network for IGS and navigation: overview and perspectives. 6th International Colloquium – Scientific and fundamental aspects of the Galileo programme, Technical University of Valencia, Spain, 25-27 October 2017
- [5] Z. Altamimi, P. Rebischung, L. Métivier, and X. Collilieux (2016), ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions, *J. Geophys. Res. Solid Earth*, 121, 6109–6131, doi:10.1002/2016JB013098.
- [6] G. Petit, A. Kanj, S. Loyer, J. Delporte, F. Mercier and F. Perosanz. 1×10^{-16} frequency transfer by GPS PPP with integer ambiguity resolution. *Metrologia* 52 (2015) 301–309, doi:10.1088/0026-1394/52/2/301
- [7] S. Melachroinos, Biancale R., Llubes M., Perosanz F., Lyard F., Bouin M.N., Masson F., Nicolas J., Morel L., and Durand S., 2007, Ocean tide loading (OTL) displacements from global and local grids : comparisons to GPS estimates over the shelf of Brittany, France, doi: 10.1007/s00190-007-0185-6: *J. Geodesy*, v. 82, p. 357-371.
- [8] L. Lescarmontier, Legresy B., Coleman R., Perosanz F., Mayet C. and Testut L. (2012). Vibrations of Mertz Glacier ice tongue, East Antarctica. *Journal of Glaciology* 58(210): 665-676.
- [9] D. Laurichesse, F. Mercier, J.P. Berthias, P. Broca, L. Cerri, “Integer Ambiguity Resolution on Undifferenced GPS Phase Measurements and its Application to PPP and Satellite Precise Orbit Determination”, *Navigation, Journal of the institute of Navigation*, Vol. 56, N° 2, Summer 2009
- [10] F. Fund, Perosanz F., Testut L. and Loyer S. (2013). An Integer Precise Point Positioning technique for sea surface observations using a GPS buoy. *Advances in Space Research* 51(8): 1311-1322.
- [11] D. Laurichesse, C. Rouch, F.X. Marmet, M. Pascaud. Smartphone Applications for Precise Point Positioning. ION GNSS+ September 24-28, 2018 Miami, Florida.

Positionnement GNSS statique et cinématique appliqué à l'étude des déformations tectoniques de la Terre

Static and kinematic GNSS positioning for the study of the tectonic deformations of the Earth

Pierre Briole¹

¹ *Laboratoire de Géologie (LG) CNRS : UMR8538, École Normale Supérieure de Paris, Université de recherche Paris Sciences Lettres (PSL), 45 Rue d'Ulm 75005 Paris - France*

Mots clés (en français et en anglais) :

Présentation en session plénière / Plenary session communication

Résumé / Abstract

Au cours des années 1960 l'accumulation de décennies d'observations variées permit aux physiciens du globe de s'accorder sur deux théories simples, l'une permettant de modéliser les mouvements des plaques tectoniques, l'autre permettant de modéliser la rupture sismique. Les développements de la géodésie allaient alors permettre d'apporter des validations supplémentaires et des raffinements à ces modèles. Les données de nivellement, extrêmement précises, furent fondamentales pour l'étude des séismes. Dès leur apparition, les mesures de géodésie spatiale basées sur le VLBI et le SLR apportèrent les premières mesures géodésiques directes des mouvements des plaques.

L'arrivée, à l'aube des années 1990, de deux nouveaux outils, le GPS et l'interférométrie radar à synthèse d'ouverture, constitua un tournant majeur qui allait donner peu à peu un rôle fondamental à la géodésie dans l'étude de la dynamique de la Terre de l'échelle globale à l'échelle locale. La précision et la densité croissante d'observations GPS (puis GNSS), la disponibilité de séries temporelles plus longues, couvrant désormais trois décennies, a permis de faire progresser considérablement les modèles initiaux. Le caractère non purement élastique de la croûte et du manteau sous-jacent est clairement démontré par de nombreux enregistrements de mouvements distribués dans le temps, comme par exemple les relaxations post-sismiques qui peuvent être observées pendant des années après les plus gros séismes.

Ces observations apportent des contraintes sur la rhéologie de l'intérieur de la Terre dans les zones où les séismes nucléent et par conséquent sur les temps nécessaires pour que des contraintes migrent dans le milieu d'une zone aux zones adjacentes qui pourraient à leur tour rompre parce que d'autres failles s'y trouvent. Il existe aussi des observations de plus en plus nombreuses montrant l'existence de glissements asismiques sur des failles, ceci à toutes les échelles, depuis des petites failles très superficielles jusqu'à de grandes interfaces de subduction. L'existence de ce type de phénomène fut envisagée dès le grand séisme du Chili de 1960 au cours duquel des enregistrements de sismomètres longue période et de gravimètres suggéraient que près de la moitié de l'énergie sismique pourraient avoir été libérée de manière « silencieuse » dans les minutes précédant le choc principal. Mais ce n'est qu'à partir de l'arrivée du GPS que ces phénomènes ont pu être observés de mieux en mieux, puis utilisés quantitativement pour modéliser les processus.

Les grands séismes sont souvent les plus destructeurs et ce sont aussi ceux qui génèrent les plus gros signaux. Ce sont ceux pour lesquels le GPS a permis de faire les plus grandes avancées théoriques ces dernières années. Cependant ces grands séismes sont rares et ne sont pas le reflet de la diversité de la déformation tectonique de la Terre qui possède aussi des zones de déformation plus faible, plus distribuée, et dans lesquelles, a priori, l'aléa

est moins élevé. C'est le cas de l'Europe et de la Méditerranée par exemple. Si dans ces régions l'aléa est plus faible, le risque sismique n'en est pas moins négligeable, en raison de la densité de population, de la possible existence d'infrastructures à risques, et de la possibilité d'événements rares mais très forts et très superficiels et potentiellement très meurtriers s'ils se produisent sous une ville. Les failles responsables de ces séismes peuvent être de dimensions assez modestes, typiquement 10 à 20 km. La surveillance géodésique de telles failles est un challenge à la fois pour l'interférométrie radar et pour le GNSS. Pour l'interférométrie radar parce que la rareté des événements impose de s'intéresser à de longues séries temporelles de données, bien plus longues que la durée de vie d'une mission spatiale. Pour le GNSS parce que la taille de grille nécessaire au sol est du kilomètre typiquement, donc à plusieurs ordres de grandeur de la densité des réseaux GNSS permanents existants ou envisageables dans n'importe quel pays.

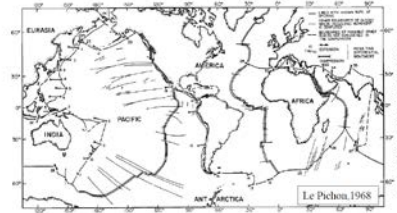
En conséquence la capacité future des scientifiques et des pouvoirs publics à appréhender les déformations liées à des séismes modérés à forts, jusqu'à la magnitude 6, dépend de leur capacité à mettre en place des méthodes nouvelles permettant de mesurer non plus seulement des ensembles peu nombreux de points GNSS épars, mais des lignes continues d'observations par exemple en utilisant tous types de réseaux (réseaux de communications ferroviaires, routiers, ... ; réseaux électriques, informatiques, ...) continus

Positionnement GNSS statique et cinématique appliqué à l'étude des déformations tectoniques de la Terre

Pierre Briole - CNRS/ENS-PSL
28 mars 2018

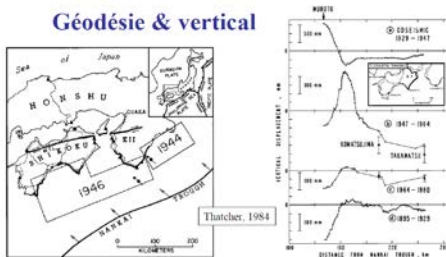


Plaques tectoniques



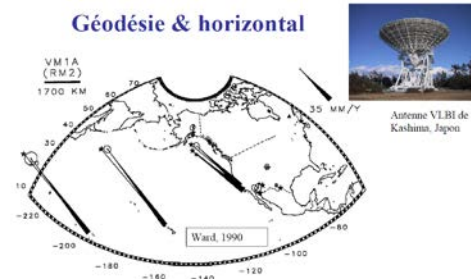
Au cours des années 1960 l'accumulation de décennies d'observations variées permit aux physiciens du globe de s'accorder sur deux théories simples, l'une permettant de modéliser les mouvements des plaques tectoniques, l'autre permettant de modéliser la rupture sismique.

Géodésie & vertical



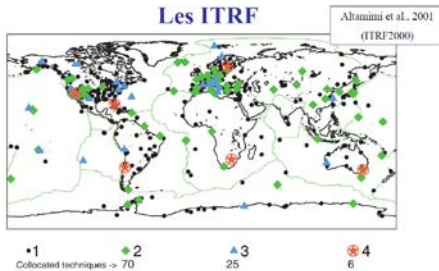
Les développements de la géodésie allaient alors permettre d'apporter des validations supplémentaires et des raffinements à ces modèles. Les données de nivellement, extrêmement précises, furent fondamentales pour l'étude des séismes.

Géodésie & horizontal



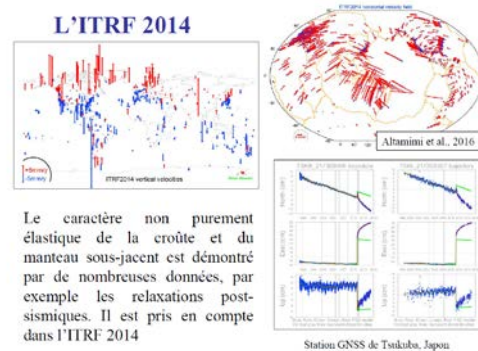
Dès leur apparition, les mesures de géodésie spatiale basées sur le VLBI et le SLR apportèrent les premières mesures géodésiques directes des mouvements des plaques.

Les ITRF



L'arrivée du GPS constitua un tournant majeur qui allait donner peu à peu un rôle fondamental à la géodésie dans l'étude de la dynamique de la Terre de l'échelle globale à l'échelle locale.

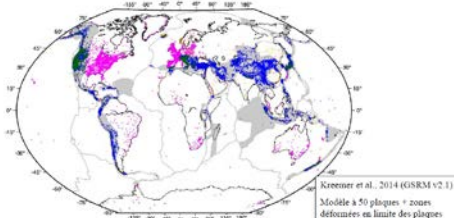
L'ITRF 2014



Le caractère non purement élastique de la croûte et du manteau sous-jacent est démontré par de nombreuses données, par exemple les relaxations post-sismiques. Il est pris en compte dans l'ITRF 2014

Station GNSS de Tsukuba, Japon

Les modèles de mouvement des plaques



La précision et la densité croissante d'observations GPS (puis GNSS), la disponibilité de séries temporelles plus longues, couvrant désormais trois décennies, a permis de faire progresser considérablement les modèles initiaux.

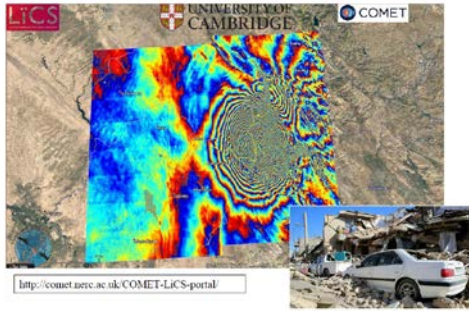
Cartographie des déformations du sol par interférométrie radar



L'interférométrie radar a donné un élan extraordinaire à l'étude des déformations liées aux séismes et autres phénomènes telluriques

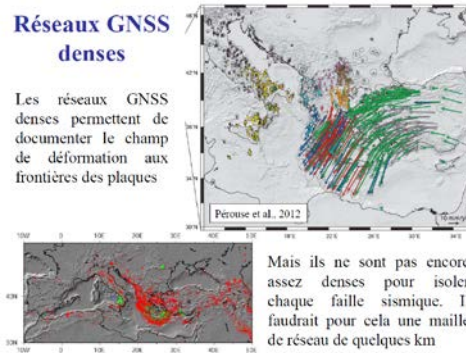
Massonnet et al., 1993

Séisme du Kurdistan 12/11/17 $M_w = 7.3$

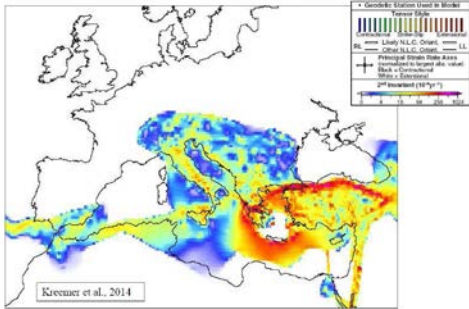


Réseaux GNSS denses

Les réseaux GNSS denses permettent de documenter le champ de déformation aux frontières des plaques



Déformation en Méditerranée d'après le GPS



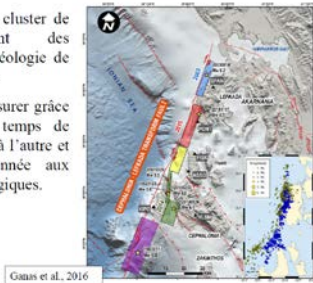
En verts, séismes méditerranéens étudiés par géodésie (nivellement, trilateration, GNSS, InSAR) depuis 1900



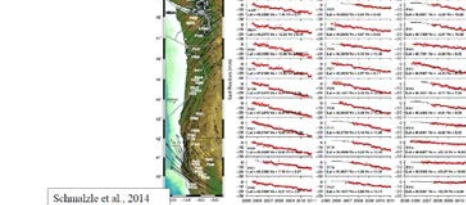
Les séismes (et déformations associées) peuvent se produire en cascade

Les observations de cluster de séismes apportent des contraintes sur la rhéologie de l'intérieur de la Terre.

Il est possible de mesurer grâce à la géodésie les temps de transfert d'une zone à l'autre et de relier cette donnée aux observations sismologiques.

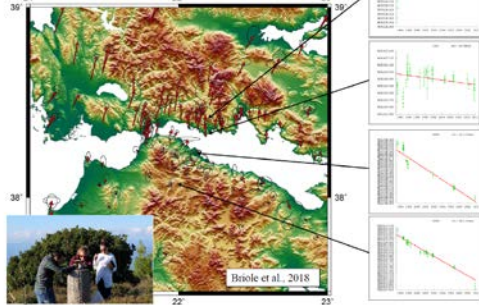


Mouvements transitoires aux frontières de plaques

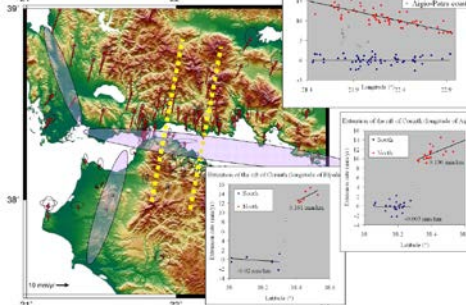


Il existe des observations de plus en plus nombreuses montrant l'existence de glissements aiséismiques sur des failles, ceci à toutes les échelles, depuis des petites failles très superficielles jusqu'à de grandes interfaces de subduction.

Campagnes GNSS et séries temporelle de coordonnées



Localisation de la déformation

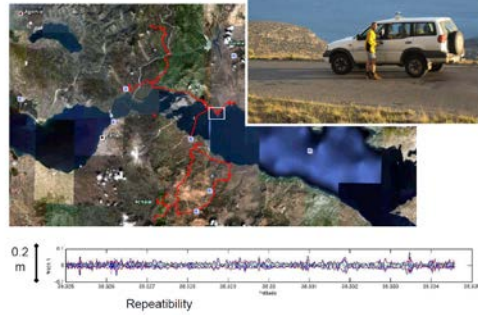


Combiner les avantages du GNSS et de l'InSAR pour le suivi des déformations de la Terre

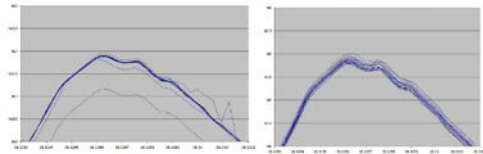
Il y a deux enjeux pour aller plus loin avec le GNSS et l'InSAR

- (1) Coté GNSS: Avoir des levés GNSS plus denses permettant de s'approcher localement de la densité de l'InSAR
- (2) Coté InSAR: Réussir à avoir des séries temporelles longues et cohérentes, au moins sur un certain nombre de pixels

GPS cinématique

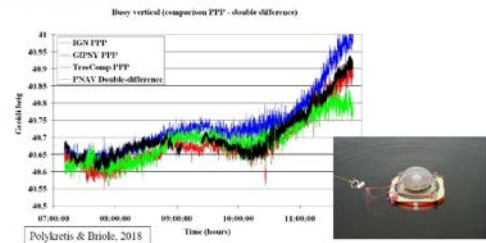


Il faut améliorer les logiciels de calculs GNSS cinématique PPP (+ hybridation)



Observations Septembre 2016, <http://school2016.crlab.eu/>

Les logiciels PPP fonctionnent assez bien dans des cas simples (bouées) mais assez mal dans les applications réelles embarquées sur des véhicules



Polykretis & Briole, 2018

Pour un maillage d'observations GNSS de haute précision

Pour appréhender les déformations liées à des séismes modérés (M<6), les scientifiques et les pouvoirs publics ont besoin d'observations très denses.

L'InSAR est efficace sur des temps courts, principalement pour le co-sismique. Sa perte de cohérence rapide (mois, ans) le rend inefficace pour mesurer l'accumulation séculaire de la déformation.

Les réseaux GNSS permanent et même ceux de campagnes n'ont pas la résolution spatiale suffisante pour imager les failles une à une.

Des réseaux cinématiques continus et maillés, utilisant de manière opportuniste les infrastructures existantes (ferroviaires, routiers, électriques, informatiques, ...) permettraient de résoudre une partie de ce problème.

Il faut des logiciels PPP performants



Antennes compactes en cavité pour applications GNSS *Compact Antennas in Cavities for GNSS Applications*

L. García-Gámez^{1,2}, L. Bernard², R. Sauleau¹, S. Collardey¹, K. Mahdjoubi¹, P. Pouliguen³, P. Potier³

¹ IETR Institut d'Electronique et de Télécommunications de Rennes, Rennes, France, laura.garcia-gamez@isl.eu

² ISL French-German Research Institute of Saint-Louis, Saint-Louis, France

³ DGA Direction Générale de l'Armement, France

Mots clés : antenne, cavité, métamatériaux, GNSS.

Key words: antenna, cavity, metamaterials, GNSS.

Abstract

L'objectif de ce travail est le développement d'une antenne miniature intégrée dans une cavité métallique rectangulaire. Cette antenne doit couvrir simultanément trois systèmes GNSS (Galileo E1, GLONASS G1 et GPS L1). De ce fait, la bande passante doit être au moins de 60 MHz. Un modèle à base de métasurfaces sera présenté, ainsi que les résultats obtenus en simulation.

The objective of this work is the development of a miniature antenna integrated inside a rectangular metal cavity. This antenna must cover three of the GNSS systems at the same time (Galileo E1, GLONASS G1 and GPS L1). To this end, the bandwidth must be at least 60 MHz. A model based on metasurfaces will be presented, as well as the results obtained in simulation.

1 Introduction

Geolocalization of flight systems, like projectiles, often need small antennas embedded in metal cavities [1]. In this case, a small square or nearly square antenna, which has the ability to cover three GNSS systems simultaneously, is needed. In order to cover the frequencies used by Galileo E1, GLONASS G1 and GPS L1 with a small single antenna, a device with at least a 60-MHz bandwidth is required with a central frequency of 1.578 GHz. Due to metallic environment, the bandwidth (BW) reduction is the immediate and main drawback [2].

First, it has been demonstrated that it is not possible to obtain the above-mentioned requirements with a standard microstrip patch antenna because of the reduced dimensions of the cavity (maximum $0.21 \lambda_0$, 40 mm). Another option to achieve the desired bandwidth consists in using stacked patches, although, for very small aperture size. This configuration can offer a substantial increase of bandwidth compared to a single-layer patch. The use of a metasurface is also an attractive solution, it is possible to increase the bandwidth without increasing the dimensions of the cavity. Therefore, the goal is to develop the metasurface-based design of a compact antenna with the widest possible bandwidth. This bandwidth must be higher than 60 MHz and the width of the antenna can not exceed 40 mm.

In previous studies a square antenna in cavity based on metasurfaces has been presented [3]. The metasurface consists of small patches of equal size printed on the same surface. This concept leads to a decrease in the inductance and to a purely capacitive system. In the study mentioned above, stacking these metasurfaces has also been studied. This approach helps to increase even more the capacitive effect of the cavity antenna. This concept was introduced in [4]. Here, due to the fixed width of the antenna carrier, $0.21 \lambda_0$ (40 mm), it was not possible to use a square metasurface antenna to reach the desired bandwidth. Because of this, a rectangular antenna of size $0.21 \times 0.24 \times 0.11 \lambda_0$ has been selected in order to design an antenna with an aperture size as close as possible to a square. As we are targeting a platform flying in altitude, radio propagation does not suffer from multipath propagation, and the antenna polarization is chosen to be linear.

The presentation of this study is organized as follows. Firstly, the design and optimization results of the stacked patch antenna are presented in Section 2. Second, the metasurface-inspired antenna is exposed in Section 3. Finally, the two previous designs are compared and discussed in the Conclusion.

2 Stacked patch antenna

To verify if it is possible or not to meet the requirements set out above, a stacked patch design has been studied. Figure 2 shows the cross-section view of a square cavity antenna based on two stacked patches of different size with a cavity dimensions $a \times a \times h$, where a is the aperture size and h the depth. A parametric study has been carried out by varying the variable a . The depth of the cavity is a fixed value, $h = 20$ mm and the patches (lengths l_1 and l_2) are printed on a dielectric material. The gap between both patches is filled by air. Finally, the feed point is displaced a distance fp respect of the center of the cavity.

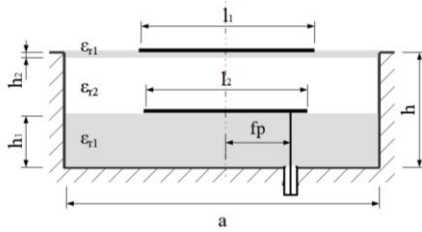


Figure 2 : Cross-section view of the stacked patch antenna in a metallic cavity.

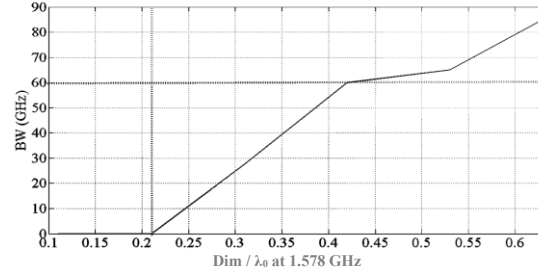


Figure 1 : Antenna bandwidth as a function of the aperture size of a square cavity for stacked patches.

Figure 1 shows the maximum achievable bandwidth depending on the aperture-size of the cavity. A comparative study has been performed using the commercial software CST Studio Suite [5]. As the limits of the requirements show, it can be concluded that it is not possible to use a stack-based design to reach the desired 60-MHz of bandwidth at 1.578 GHz with aperture dimensions lower than $0.43 \lambda_0$ ($a = 81$ mm), moreover, it is not possible to match this antenna to the frequency of 1.578 GHz for dimensions lower than $0.21 \lambda_0$.

3 Metasurface geometry

A schematic view of the rectangular metasurface-inspired antenna is shown in Figure 3(a). The design is based on a metallic cavity with a depth $h = 20$ mm and with a rectangular aperture of 40×45 ($a \times b$) mm^2 ($0.21 \times 0.24 \lambda_0$). Most of the cavity, 90%, can be filled with air ($h_1 = 18.08$ mm), or with a solid material with low permittivity, as for example polypropylene ($\epsilon_r = 2.3$) [6], which will be finally used in the manufacturing process for mechanical issues. The upper part of the cavity is composed by the metasurface arranged in two parallel layers, as shown in Figure 3(b). The metasurface provides a capacitive effect to the structure of the antenna to balance the inductance provided by the cavity, and to make the antenna resonate. Each layer is composed of three rectangular patches of equal size ($l_1 \times l_2$), with the aim of obtaining the maximum capacity without reaching the coupling, and they are printed on a thin substrate of high permittivity ($\epsilon_r = 10.2$). Finally, the feeding of the antenna is made by a coaxial probe located 11 mm from the edge of the cavity and 9 mm from the center, on one of the side patches.

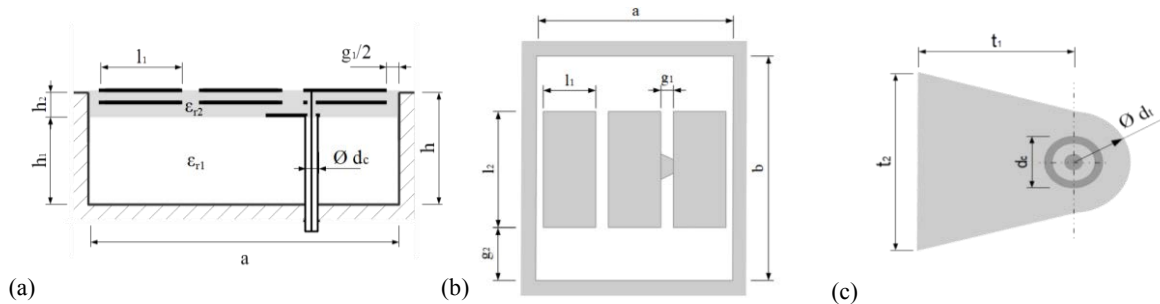


Figure 3 : (a) Cavity antenna cross section. (b) Top view of the metasurface-inspired cavity antenna. (c) Detail of the metal part welded to the feeding point.

Additionally, a small piece of metal is added to the bottom of the high permittivity substrate to match the antenna and facilitate soldering process, Figure 3(c).

Figure 4(a) shows the reflection coefficient S_{11} obtained in simulation. The BW is 68 MHz (FBW = 4.1% at the center of 1.578 GHz) around the central frequency which gives us a slightly higher bandwidth than the minimum required. The variables used in to optimize the antenna performance are the dimensions of the small metal part (t_2 , t_1), the position of the feed point (fp) and the gap between patches (g).

This antenna has a high efficiency shown in Figure 4(b). Simulated gain is 4.54 dBi at the central frequency 1.578 GHz with an efficiency of 96.5%.

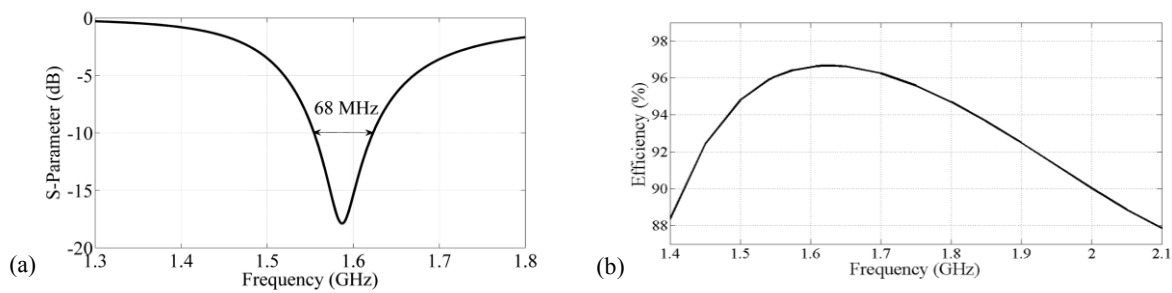


Figure 4 : (a) Simulated S_{11} for metasurface-inspired cavity antenna of dimensions $0.21 \times 0.24 \times 0.11 \lambda_0$ at 1.578 GHz ($40 \times 45 \times 20$ mm). (b) Antenna efficiency for metasurface-inspired rectangular cavity antenna at maximum directivity.

4 Conclusion

In conclusion, this paper shows two different studies to realize an antenna that covers the bands used by Galileo E1, GLONASS G1 and GPS L1 (60-MHz bandwidth) with a single antenna embedded in small metallic cavity. With a constraint on dimensions, the only possible solution to fulfil the bandwidth condition consists of a metasurface-inspired patch antenna. The antenna based on stacked patches is a well-known solution to increase the bandwidth of a certain system. For the specific case of metallic cavities of small size, the stack up of two patches do not give satisfying results to reach a sufficient BW (60 MHz).

On the other hand, the use of metasurfaces allows to increase the bandwidth due to the capacitive effect that they introduce. In other words, it is possible to minimize the size of the antenna for a required BW. The numerical results have shown that a square cavity of $0.21 \times 0.21 \lambda_0$ did not reach the specified BW of 60 MHz, but a rectangular cavity of $0.21 \times 0.24 \lambda_0$ is sufficient to reach a BW of 68 MHz. In comparison to the stacked patch configuration, the aperture reduction is of 72%. This result illustrates the advantage of a metasurface-inspired antenna for integration into a metallic cavity. Prototypes are currently under fabrication.

References

- [1] N. C. Karmakar, "On miniaturisation and bandwidth enhancement of a cavity backed circular microstrip patch antenna," *Ant. and Propag. Society International Symposium*, vol. 3, pp. 596-599, 2002.
- [2] M. Martinis, K. Mahdjoubi, R. Sauleau, S. Collardey and L. Bernard, "Circuit models explaining the bandwidth behavior of small circular cavity-backed patch antennas," in *EuCAP 2014 - 8th European Conference on Antennas and Propagation*, The Hague, The Netherlands, 2014.
- [3] M. Martinis, "Development and characterization of metamaterials in cavities. Applications to the design of compact antennas.," Université de Rennes I, Rennes, 2014.
- [4] M. Martinis, K. Mahdjoubi, R. Sauleau, S. Collardey and L. Bernard, "Bandwidth behavior and improvement of miniature cavity antennas with broadside radiation pattern using a metasurface," *IEEE Trans. Ant. Prop.*, vol. 63, no. 5, pp. 1899-1908, 2015.
- [5] *CST Studio Suite*, <http://www.cst.com>, 2016.
- [6] M. E. Cos and F. Las-Heras, "Polypropylene-based dual band CPW-fed monopole antenna," *IEEE Antennas Propag. Mag.*, vol. 55, no. 3, pp. 264-273, 2013.

Titre (*français*) : Les destins liés des théories et des systèmes radioélectriques.

Title (*English*): The linked destiny of theory and radio-electric systems.

Auteur : Pierre Fuerxer

Ingénieur de l'armement, Ancien adjoint militaire au directeur de CNET.

Mots clés : *Ondes, gravitation, théories, expériences*
Waves, gravitation, theories, experiments

Résumé:

Dans le passé, les scientifiques avaient construit leurs théories sur les résultats d'expériences rustiques. Ceci apparaît clairement à la lecture de la plupart des articles décrivant les expériences cruciales réalisées au 20^e siècle.

En rêvant à de nouveaux systèmes, des chercheurs ont joué un rôle décisif dans le développement de la radioélectricité. Ainsi, les besoins de l'industrie des télécommunications ont largement contribué au développement des théories et des technologies électroniques. Très souvent, les travaux théoriques ont été conduits par des entrepreneurs espérant conquérir un marché solvable.

Aujourd'hui, les GNSS ont atteint un marché de masse. Est-il possible d'utiliser leurs constellations de satellites pour réaliser des expériences susceptibles d'améliorer nos connaissances en géophysique et en cosmologie ?

Abstract:

In the past, scientists based their theories on behalf of very crude experiments. This is clearly attested by the analysis of the papers describing most of the 20th century decisive experiments.

Dreaming to new systems, scientists played a decisive role in the development of the radio electric science. The needs of the telecommunication industry contributed widely to the development of theories and electronic technologies. Very often, and especially in the radio electric field, theoretical works have been driven by managers expecting a solvable market.

Now, GNSS is a mass market. Is it possible to use their satellites constellations to build new experiments able to enhance our knowledge in geophysics and cosmology?

Le rôle important du hasard :

Généralement, les scientifiques conçoivent les expériences dans le but de confirmer leurs options théoriques. La lecture du mémoire de Miller publié en 1933 illustre parfaitement cette attitude [1]. L'expérience était très rustique : la source lumineuse était une lampe à acétylène, les relevés faits au dixième de frange par des opérateurs. Bien que ses résultats aient été 400 fois inférieurs à ses prévisions, Miller a pensé avoir confirmé l'existence de l'Ether, et le déplacement de la Terre par rapport à ce milieu hypothétique.

De même, le 15 avril 1928, Esclangon [2] publie un article dans lequel il décrit les déviations optiques observées à Strasbourg avec une lunette astronomique. Ne disposant pas des moyens de calcul actuels, il pensait avoir observé une période d'un jour sidéral. Ces résultats n'ayant pas été confirmés à Paris avec un télescope, cette observation a été attribuée à des erreurs expérimentales. Une analyse "spectrale" des 154 mesures publiées par Esclangon conduit à des résultats surprenants.

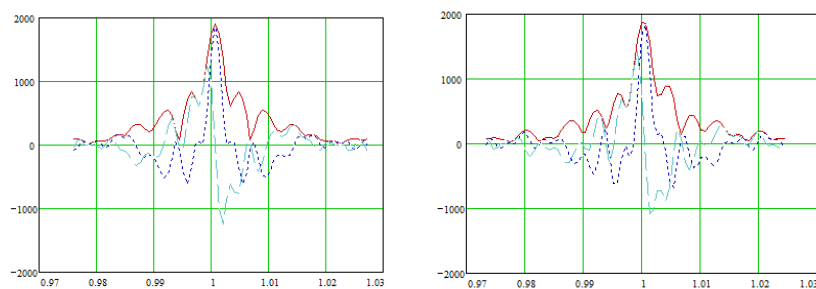


Figure 1 : A gauche : analyse des résultats. A droite : test sur la période d'un jour solaire.

Il est aujourd'hui facile de faire une analyse de ses mesures (figure 1). Il existe bien une périodicité mais sa valeur était d'un jour solaire. Enfin, le traitement étant ambigu, la période des fluctuations pourrait tout autant être de 12 heures que de 24 heures.

Est-on certain que les anomalies rapportées par Miller et Esclangon résultent de causes fortuites, en particulier des variations de température ou de pression atmosphérique ? Toutes les théories ayant un domaine de validité limité, est-on certain que la contraction de Lorentz est strictement celle qui assure la conservation de la vitesse de la lumière ? La théorie de la relativité générale prédit une courbure des rayons lumineux. Est-on certain que l'influence gravitationnelle du Soleil et la Lune ne peuvent pas expliquer leurs résultats ? Enfin, l'entraînement des ondes par les milieux diélectriques ne pourrait-il pas correspondre à la formule de Fresnel ?

Souvent, les scientifiques pensent qu'une validation expérimentale est la preuve d'une théorie. En fait, les mêmes faits peuvent souvent être décrits par plusieurs théories. L'optique est un domaine dans lequel nous acceptons d'utiliser, au cas par cas, les théories les plus aptes à l'étude d'un dispositif particulier. Qu'un phénomène puisse être décrit par une théorie n'invalide pas les autres approches, lorsqu'elles conduisent à des résultats identiques, indiscernables dans la pratique, ou même d'une précision suffisante.

Considérons l'expérience de William Bertozzi [3]. Son but était de vérifier la formule donnant l'énergie d'un électron en fonction de sa vitesse. Comme dans les cas précédents, cette expérience a été faite alors que la technologie électronique était encore très rudimentaire. Les mesures de vitesse se faisaient à l'aide de photos d'un écran d'oscilloscope !

Les résultats de William Bertozzi tiennent dans une liste de cinq mesures. Sur la figure 2, les croix correspondent aux quatre premières mesures de la vitesse d'un électron en fonction du rapport k de son énergie cinétique à son énergie au repos.

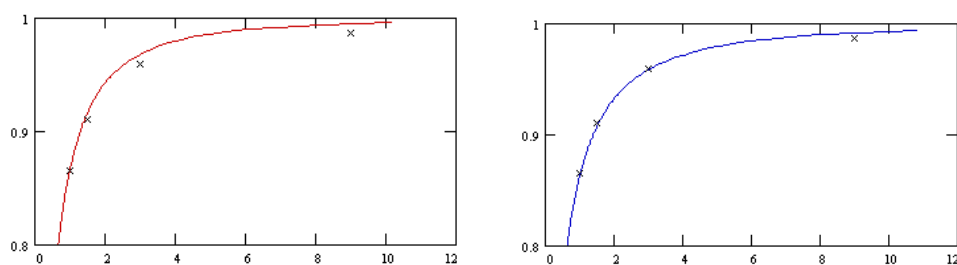


Figure 2 : A gauche : Formule relativiste. A droite : variante possible.

Les résultats obtenus par William Bertozzi peuvent être calculés par la formule relativiste, mais également en faisant les hypothèses suivantes:

- L'énergie de l'électron au repos se partage en parties égales entre sa masse mécanique et l'énergie de son champ électromagnétique calculée selon les formules de l'électromagnétisme classique.
- Lorsqu'il est mobile, l'énergie de son champ électromagnétique est calculée en faisant appel à une représentation ondulatoire.

Nous obtenons ainsi une autre expression de l'énergie cinétique en fonction de la vitesse de l'électron (figure 2). Quel aurait été le choix de William Bertozzi s'il n'avait pas connu la formule relativiste ?

Les centrales de navigation inertielles utilisent des gyromètres LASER. L'expérience a montré que ces dispositifs présentent une plage aveugle dans laquelle ils ne détectent pas leur rotation. Cette anomalie a été prise en compte en introduisant artificiellement une rotation périodique de ces dispositifs. Nous avons compris que ce défaut provient de la rétrodiffusion des miroirs qui introduit un couplage entre les deux ondes contrarotatives. Que se serait-il passé si l'expérience de Michelson et Gale [4] avait échoué en 1925 ? Aurait-on tenté de réaliser des gyromètres LASER ?

Une nécessité : conserver des archives précises :

La crédibilité de la physique suppose que nous soyons certains de ses bases expérimentales. Il est donc nécessaire que les données brutes soient conservées de telle sorte que nous puissions vérifier les protocoles expérimentaux et les traitements de données associés. Les physiciens de la première moitié du 20^e siècle l'ont fait. Aujourd'hui, les possibilités offertes par les réseaux informatiques devraient faciliter grandement le travail important que représente la constitution de ces archives.

Grâce à la technologie moderne, aux LASERS et aux nouveaux moyens de calcul, les étudiants devraient pouvoir refaire les expériences qui sont à la base de nos théories, en réaliser de nouvelles qui n'étaient pas faisables au 20^e siècle, en diffuser largement les résultats bruts à la communauté scientifique et enfin leurs conclusions. Ceci devrait être fait pour l'ensemble des programmes de recherche nationaux et internationaux, mais aussi pour les données collectées de façon non intentionnelle par ces grands programmes.

Aujourd'hui, des géolocalisations d'une précision encore inimaginable lors du lancement des grands programmes de localisation par satellites ont été obtenues.

L'analyse des irrégularités apparaissant dans les orbites des satellites pourrait-elle conduire à des découvertes en géophysique? La géolocalisation pourra-t-elle apporter une révolution dans notre perception de l'espace-temps, puis de la cosmologie et finalement améliorer radicalement notre compréhension de l'univers?

Les capacités de traitement numérique des données dont nous disposons aujourd'hui sont sans commune mesure avec celle dont disposaient les physiciens au 20^e siècle. Elles nous permettent même de traiter des données échantillonnées de façon irrégulière ou aléatoire.

L'accès aux données internes des grands programmes internationaux permettrait à notre communauté scientifique d'exploiter les informations inconsciemment collectées par les systèmes de géolocalisation. Pourquoi les responsables de ces systèmes ne rendraient-ils pas publiques ces données, y compris les anomalies éventuellement observées dans les trajectoires des satellites ?

Deux exemples significatifs :

Conserver des archives est utile, à la condition qu'elles puissent être exploitées par nos successeurs. Considérons, sans prendre parti, la polémique qui a suivi la publication des résultats de l'expérience OPERA et le document publié en 2011 par le CERN, toujours présent sur INTERNET.

Ce document de synthèse sur l'expérience OPERA donne de nombreux détails sur le dispositif expérimental, et sur les équipements électroniques utilisés. Il était difficile de mesurer le temps d'arrivée d'une impulsion de 10 μ s avec une précision de l'ordre de 1 ns. En revanche, il manque un exposé permettant à un étudiant de l'an 2100 de comprendre réellement ce qui a été fait : Un schéma de principe clair, les hypothèses implicitement retenues... Cet étudiant saura-t-il si les durées sont exprimées en temps légal actuel ou en temps atomique, et dans quel repère sachant que le synchronisme dépend de celui-ci...

Malgré leurs points obscurs, les documents relatifs aux résultats de VIRGO sont, de ce point de vue, bien plus utiles. La détection quasi miraculeuse d'un événement cosmique rare, confirme la nature ondulatoire de la gravitation. Ceci est un providentiel coup de pouce de la nature. Pour autant, l'analyse de cet interféromètre complexe sera-t-elle convaincante pour les étudiants des générations futures ?

Le rôle fondamental des applications opérationnelles :

La physique des radioélectriciens s'est développée sur la base de leurs seuls besoins. Ils ont mis au point des amplificateurs paramétriques sans avoir recours aux notions quantiques. Pour eux, théorie et technologie n'ont pas été des fins en soi, mais les moyens de construire les briques nécessaires à la réalisation des systèmes qu'ils avaient imaginés. Ils ont parfois fait des erreurs théoriques, par exemple dans l'étude de la modulation, mais l'expérience a eu le dernier mot.

Les multi-trajets ont longtemps semblé devoir limiter de façon radicale les performances des systèmes radioélectriques. Les radioélectriciens ont donc développé des traitements de signaux radicalement nouveaux rendus possibles par les progrès de la technologie numérique. Les multi-trajets sont devenus une richesse et ont parfois été introduits volontairement (par exemple dans le système de radiodiffusion DAB). Là encore, les radioélectriciens ont apporté une contribution originale.

Principe de relativité ou théorie de la relativité ?

Les théories physiques que nous utilisons ont été construites sur la géométrie euclidienne, base mathématique de tous nos raisonnements. Nous parlons de corrections relativistes et considérons implicitement la position, la vitesse et l'accélération des mobiles par rapport à un repère "absolu", choisi en fonction de critères propres à nos systèmes. Le meilleur repère n'est-il pas celui dans lequel nous travaillons ? Notre approche a été d'étudier séparément chaque éléments en tenant compte de leur environnement (Mouvement par rapport au repère "privilegié" choisi, champ de gravitation...). Ensuite, nous assemblons ces modules.

Cette démarche pragmatique n'est-elle pas totalement conforme au principe de relativité ? Comment pourrions-nous construire une théorie prenant en compte toutes les perturbations des systèmes de localisation si nous ne réhabilitons pas, faute de repère et de temps absolus, la notion de repère privilégié ?

Les concepteurs des interféromètres LIGO et VIRGO supposent implicitement leur appareils indéformables et immobiles par rapport au milieu de propagation des ondes gravitationnelles. C'est bien ce postulat qui leur permet d'imaginer pouvoir détecter des fluctuations gravitationnelles. Par ailleurs, les très grandes dimensions de ces instruments, choisies de façon à privilégier la sensibilité, imposent la réalisation d'asservissements qui éliminent les vibrations du sol, et en même temps l'effet de toute variation gravitationnelle basse fréquence.

Il est naturel de penser que tous nos systèmes terrestres et satellitaires sont soumis à un champ de gravitation lentement variable, modifié par les mouvements relatifs du Soleil, de la Lune et de la Terre.

Quelle que soit l'approche théorique retenue, la courbure gravitationnelle des rayons lumineux et la vitesse des mobiles par rapport au repère géocentrique ne peuvent qu'imposer des corrections locales.

Pendant plus d'un siècle, des scientifiques n'ont cessé de réaliser, souvent avec des moyens très limités, des expériences qui leur semblaient contredire notre usage actuel du principe de relativité. Je citerai, à titre d'exemples, les travaux d'Héctor Múnera [5] et les premiers résultats de Vincent Morin [6]. Qui peut raisonnablement affirmer qu'une loi de la physique valable dans un élément de volume $dx.dy.dz.dt$, valide dans un système d'unités locales, est également valide à l'échelle macroscopique ? Qui peut être certain de raisonner juste dans un espace ad-hoc non-euclidien construit sur le raccordement de ces éléments de volume ?

Le principe de réalité :

L'utilisation des systèmes de géolocalisation est si naturelle que leurs utilisateurs peuvent en oublier les inévitables limitations. La précision exigée pour les applications professionnelles ou scientifiques sera de plus en plus grande. Les systèmes satellitaires prennent déjà en compte les fluctuations de l'ionosphère. Au fur et à mesure de l'évolution des besoins, il faudra prendre en compte les marées terrestres, la forme exacte du géoïde, les inhomogénéités du champ de gravitation terrestre...

Pour prendre en compte toutes les perturbations des systèmes de location, les corrections sont et vont continuer à être faites par différentes approches. Lois empiriques, asservissement, modélisation fines...

Loin des querelles philosophiques, nous adopterons, au cas par cas, les théories les plus adaptées à la résolution des problèmes concrets. Ne faisons-nous pas des approximations lorsqu'elles n'influent pas sensiblement sur le résultat final ? La diversité des théories et des représentations imagées est une richesse. Elle seule permet l'évolution de la science.

Comme au 18^e siècle, la curiosité peut pousser des intellectuels à échafauder des théories. Au cours du 20^e siècle, le développement de l'électromagnétisme a été largement dû à des "amateurs de génie", motivés par la recherche d'applications concrètes, et visant un marché qu'ils pensaient solvable. En cherchant à offrir des services, ils ont été à l'origine du RADAR et d'immenses progrès scientifiques : La théorie des tubes électroniques, les techniques hyperfréquences, les antennes... Plus récemment, les techniques d'imagerie hyperfréquence, développées pour la discrétion RADAR, ont fait progresser notre connaissance des interactions entre les ondes et les objets qui nous entourent.

Aujourd'hui, la géolocalisation est devenu un marché de masse. Des précisions encore inimaginables lors du lancement des grands programmes de localisation par satellites ont été obtenues. Ces systèmes vont-ils être à l'origine de nouveaux progrès scientifiques ?

D'extraordinaires retombées scientifiques potentielles:

En 2011, au cours des journées scientifiques sur " les Radiotélescopes du futur", je posais une question à mes yeux fondamentale : La physique du 21^e siècle sera-t-elle ondulatoire? La confirmation de l'existence d'ondes de gravité, dont les astronomes avaient supposé l'existence il a bien longtemps, vient étayer ma conviction. La physique du 21^e siècle sera ondulatoire !

Grâce à cette approche ondulatoire, les champs de gravitation fixes et mobiles deviendront, comme en électromagnétisme, des combinaisons de solutions particulières d'équations différentielles. Les interactions entre la gravitation et les ondes électromagnétiques pourront alors être décrites en introduisant des non linéarités dans les équations de Maxwell.

Dans un système quasi isolé comme le système Solaire, il sera possible de construire un repère euclidien dans lequel étudier les interactions entre le Soleil, les planètes et leurs satellites. Nous pourrions déterminer avec une précision accrue les trajectoires des satellites appartenant à nos systèmes de géolocalisation.

Nous devons adapter à la gravitation les méthodes ondulatoires adoptées en acoustique aérienne ou sous-marine pour prendre en compte des variations locales de la vitesse des ondes. Evidemment, cette transposition ne sera pas facile à faire. Pour les acousticiens, la propagation des ondes électromagnétiques était instantanée. Ils pouvaient les utiliser pour construire le repère dont ils avaient besoin. Nous ne disposons que d'ondes électromagnétiques dont la propagation dépend du champ de gravitation. Comment ferons-nous pour construire ce nouveau repère ?

Nos systèmes de localisation visent à rendre le meilleur service possible à nos clients. Notre objectif prioritaire est, et doit rester, d'améliorer leurs performances. La modélisation la plus précise possible des erreurs est donc nécessaire. En fin de compte, les théories et les systèmes de radiolocalisation seront jugés sur la base du service opérationnel rendu, mais aussi sur la valeur de leur contribution au progrès scientifique. L'identification des différentes causes d'erreur restera une préoccupation de la communauté scientifique. La modélisation de plus en plus fine des irrégularités apparaissant dans les orbites des satellites pourra probablement conduire à des découvertes en géophysique.

Lorsque j'étais le dernier adjoint militaire au directeur du CNET, et malgré la réalisation par l'équipe de Lannion de sondeurs ionosphériques modernes, la contribution française au réseau mondial de sondeurs ionosphériques a pris fin. Pourquoi ne pas chercher à utiliser les constellations de satellites pour continuer l'étude et l'observation de l'ionosphère ? Des récepteurs GPS modifiés, constitueraient une antenne lacunaire de très grandes dimensions. Associés à INTERNET, ils pourraient collecter des informations très utiles. Nous avons en effet montré que l'analyse des 154 mesures d'Esclangon permet d'obtenir des résultats intéressants. Un réseau de 1000 récepteurs de type GPS collecterait en une année plus de 10^9 mesures qui constitueraient une base de données essentielle à la conduite de nombreuses recherches.

La participation des gestionnaires des grands programmes internationaux à cet effort de recherche serait assurée au titre d'un programme d'amélioration du service des GNSS baptisé par exemple GNSSIP (GNSS Improvement Program).

Notre objectif serait la réalisation d'une antenne lacunaire 3D, de dimensions mille fois plus grandes que l'interféromètre VIRGO. Il pourrait être baptisé GWSA (Gravitation Waves Spatial Antenna).

Nos laboratoires disposeraient alors des données essentielles collectées par les récepteurs fixes, mais aussi des données internes des grands programmes internationaux. Cette antenne 3D, en s'affranchissant des limites de nos systèmes terrestres, permettrait de rêver à des observations aujourd'hui impossibles.

En contribuant au développement de nombreux services, la géolocalisation a apporté une révolution dans les usages. Elle nous permettra certainement d'affiner notre perception de l'espace-temps, puis d'améliorer radicalement notre compréhension de l'univers.

- [1] The Ether-Drift Experiment and Determination of the Absolute Motion of the Earth, *Reviews of Modern physics*, July 1933, volume 3.
- [2] Dissymétries optiques de l'espace, *Journal des observateurs*, avril 1928.
- [3] Article publié par W. Bertozzi dans : *American Journal of Physics*, Volume 32 (1964).
- [4] A. A. Michelson et Henry G. Gale, « *The Effect of the Earth's Rotation on the Velocity of Light, II.* », *Astrophysical Journal*, vol. 61, avril 1925
- [5] Héctor A. Múnera, "Absolute velocity of earth from our stationary Michelson-Morley-Miller experiment at CIF, Bogota, Colombia", poster presented at *Physical Interpretations of Relativity Theory PIRT2017*, Bauman Moscow State Technical University, Julio 3-6, 2017 .
Should the Laws of Gravitation Be Reconsidered? publié à Montréal (2011).
- [6] Déviations lumineuses d'Allais-Esclangon, un indice expérimental concernant les relations de la gravitation et de l'électromagnétisme ? Vincent Morin vincent.morin@univ-brest.fr, Pierre Fuerxer pierre.fuerxer@wanadoo.fr, *Annales de la Fondation Louis de Broglie*, Volume 33, no 3-4, 2008.

Investigations on TLE time series accuracy for GNSS satellites, over short and long-time scales

Florent Deleflie¹, Alexis Petit², Jérôme Daquin³, Felix Perosanz⁴, Michel Capderou⁵

¹IMCCE / Observatoire de Paris / GRGS

²IMCCE / Observatoire de Paris

³RMIT Melbourne,

⁴GET, Observatoire Midi-Pyrénées / GRGS

⁵Michel Capderou, LMD, Ecole Polytechnique

The trajectory of GNSS satellites, and the visibility conditions from the ground, can be carried out from precise GNSS measurements, or from the well-known Two-Line Elements (TLE) provided by NORAD. Depending upon the purpose, the latency of the tracking data, the time available to derive the results, precise orbits (at the level of one-centimeter accuracy) or rough orbits can be propagated and delivered to the users, through a large set of various tools, some of them being publicly available. We propose in this paper to analyse the compatibility of different approaches and different initial conditions, propagated over very short (a couple of days) and very long (two centuries) time scales, applied to some satellites in the various GNSS constellations (GPS, Galileo, GLONASS, BeiDou).

GRGS (Groupe de Recherche de Géodésie Spatiale) is a French research group gathering up French scientists within different institutions focusing their activities to space data analysis, and in particular related to space geodesy and precise orbitography. Hence, GRGS is in charge of the French analysis center of GNSS data providing precise orbits of the GNSS satellites IGS, the International GNSS Service, as a component of the International Association of Geodesy. Using the time series of orbital elements built by GRGS is way to take benefit of the best possible accuracy to the tracking data.

As a first part of this presentation, we analyse over a couple of days the visibility conditions as scheduled from precise orbits and from TLE time series, that have the state of "mean elements », with an accuracy which is quantified though through formulas deriving osculating elements from mean elements and conversely. As a second part of this presentation, the stability of GNSS orbits is analysed over two centuries, to investigate from the TLE times series whether the future storage orbits are likely to be sensitive, or not, to deep resonance effects, by using the French STELA s/w designed in the framework of the French space operation act.

Filtrage de Kalman et estimateurs / Kalman filter and estimators

iltrage de Kalman invariant pour la navigation inertielle hybridée
Invariant Kalman filtering for GPS-aided inertial navigation

Silvère Bonnabel¹

¹ *Centre de Robotique (CAOR) - MINES ParisTech - École nationale supérieure des mines de Paris, PSL
Research University60, boulevard Saint-Michel 75272 Paris cedex 06 - France*

Mots clés (en français et en anglais) :

Présentation en session plénière / Plenary session communication

Résumé / Abstract

Le filtre de Kalman, ou plus précisément filtre de Kalman étendu (EKF), est un outil fondamental de l'ingénieur très utilisé dans le domaine de la navigation. Le domaine récent du filtrage de Kalman dit "invariant", est consacré à l'utilisation de la géométrie sous-jacente de l'espace d'état et de la dynamique pour modifier et améliorer l'EKF classique, notamment en termes de garanties de convergence. Les principales applications de cette méthodologie sont la localisation, la navigation, mais aussi la localisation et cartographie simultanée, connue sous le nom de SLAM en anglais, pour lequel l'EKF invariant résout les problèmes bien connus d'incohérence de l'EKF classique. Bien que la méthodologie soit récente, les remarquables propriétés du filtre de Kalman invariant ont d'ores et déjà motivé une implémentation industrielle dans le domaine de la navigation, par Safran Electronics & Defense, anciennement Sagem. Le but de cet exposé est de proposer une introduction au filtrage de Kalman invariant, et de fournir les grandes intuitions permettant de comprendre sa supériorité par rapport au filtrage de Kalman étendu classique en ce qui concerne le domaine de la navigation haute précision.

Invariant Kalman filtering for GPS-aided inertial navigation

S. Bonnabel (Mines ParisTech)
Joint work with A. Barrau (Safran Tech / SAGEM)

Journées URSI
Meudon
28 mars 2018

Introduction

The **extended Kalman filter (EKF)** has been the most widespread tool for sensor fusion since the 1960s for navigation. **Developed by NASA** in the 1960s in the Apollo program.

SAGEM: pioneered industrial applications to navigation, notably thanks to Pierre Faurre.



Rudi Kalman



Pierre Faurre

Introduction

The **invariant extended Kalman filter (IEKF)** introduced a decade ago is a modification to EKF based on symmetries of the problem.¹

Safran Electronics & Defense (SAGEM) chose to invest through the PhD of A. Barrau (2012-2015). The company has registered three patents for applications of invariant filtering.

This talk: some industrial results, and some insights from a simplified example.

¹For an overview see "Invariant Kalman Filtering", Barrau & Bonnabel, The Annual Reviews, 2018

Industrial results

Safran Electronics & Defense (SAGEM) has released an industrial product using the Invariant extended Kalman filter, the Eurofir 410, embedded in the drone Patroller.



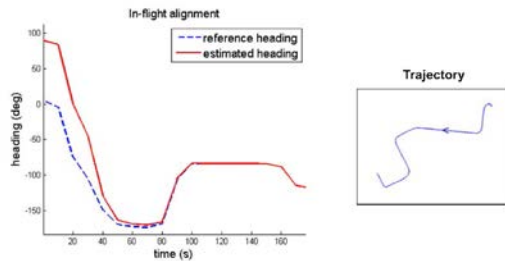
Eurofir 410



Patroller drone

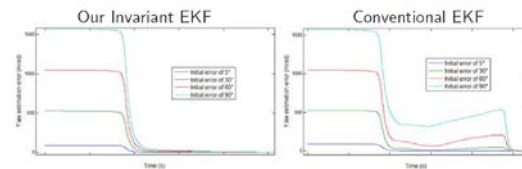
Industrial results

In flight alignment consists in finding the orientation of the vehicle from inertial sensors and GPS.



Industrial results

In flight alignment consists in finding the orientation of the vehicle from inertial sensors and GPS.



Estimation error, with initial error from 5° to 90°

Industrial results

Why is alignment challenging for the EKF ?

- Highly accurate inertial sensors
- Vertical easy to find (\approx known)
- Heading unknown \Rightarrow possible large estimation error

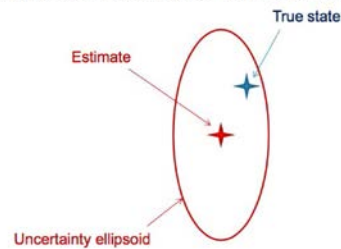
Caveats of the EKF ?

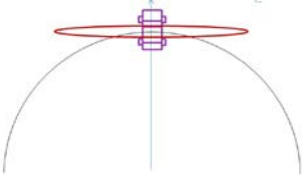
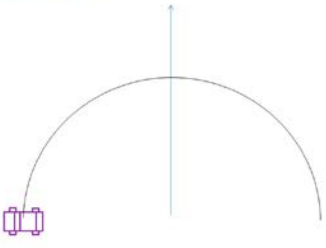
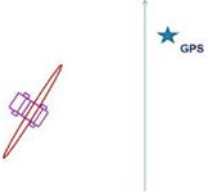

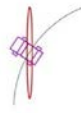
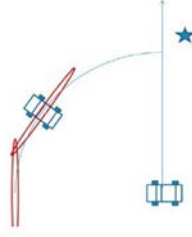
- Designed for small errors
- Here the estimate lives a (constrained) subspace of the state space.

Remainder of the talk devoted to provide insight on the superiority of the IEKF for this type of problems.

Preliminary

Graphical representation: The EKF computes an estimate \hat{X}_t with a covariance matrix P_t . This is the confidence ellipsoid.



<p>A simple terrestrial navigation problem</p> <p>Introduction : Extended Kalman Filter</p> <ul style="list-style-type: none"> Accurate odometer Initial position known Heading unknown 	<p>A simple terrestrial navigation problem</p> <p>Non-linear system : estimate</p> 
<p>A simple terrestrial navigation problem</p> <p>This is what the EKF should do</p> 	<p>Reminder: EKF equations</p> <p>Consider continuous time deterministic dynamics with discrete time measurements t_0, t_1, \dots</p> $\frac{d}{dt} X_t = f(X_t, U_t), \quad Y_n = h(X_t) + V_n$ <p>The EKF is based on a prior distribution with \hat{X}_0, P_0.</p> <p>Propagation step: for $t_{n-1} < t < t_n$</p> $\frac{d}{dt} \hat{X}_t = f(\hat{X}_t, U_t) \quad \frac{d}{dt} P_t = A_t P_t + P_t A_t^T$ <p>Update step: at $t = t_n$,</p> $K_n = P_{t_n} H_n^T (H_n P_{t_n} H_n^T + N)^{-1}$ $\hat{X}_{t_n}^+ = \hat{X}_{t_n} + K_n (Y_n - \hat{X}_{t_n})$ $P_{t_n}^+ = (I - K_n H_n) P_{t_n}$
<p>A simple terrestrial navigation problem</p> <p>And the EKF does this</p> 	<p>A simple terrestrial navigation problem</p>  <p>Obviously the EKF poorly handles small eigenvalues in P_t.</p> <p>Use of symmetries as the remedy.</p> <p>Namely: we use an Invariant (I)-EKF².</p> <p>²See "Invariant Kalman filtering", Annual Reviews, 2018.</p>
<p>Invariant Kalman filtering</p> <p>Main difference the estimation error it linearizes is not</p> $\begin{pmatrix} \hat{\theta} - \theta \\ \hat{x}_t - x_t \end{pmatrix}, \quad \text{where } \theta \in S^1, x \in \mathbb{R}^2$ <p>It is the estimation error</p> $\begin{pmatrix} \hat{\theta} - \theta \\ R(\theta)^T (\hat{x}_t - x_t) \end{pmatrix}$ <p>where $R(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$.</p>	<p>Invariant Kalman filtering</p>  <p>Estimation error</p> $\begin{pmatrix} \hat{\theta} - \theta \\ R(\theta)^T (\hat{x}_t - x_t) \end{pmatrix}$ <p>Theorem: The estimate belongs only to sets "physically" reachable.</p>

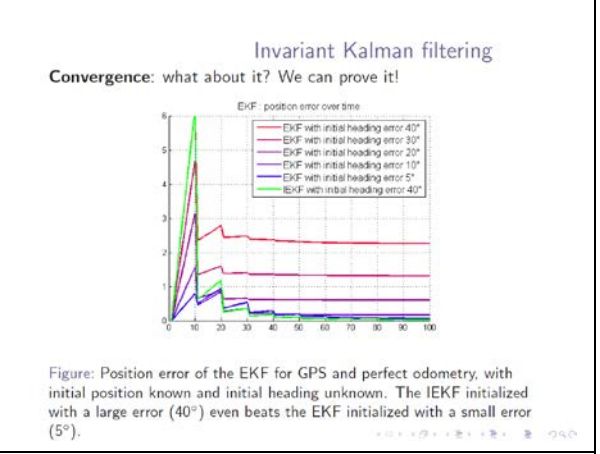
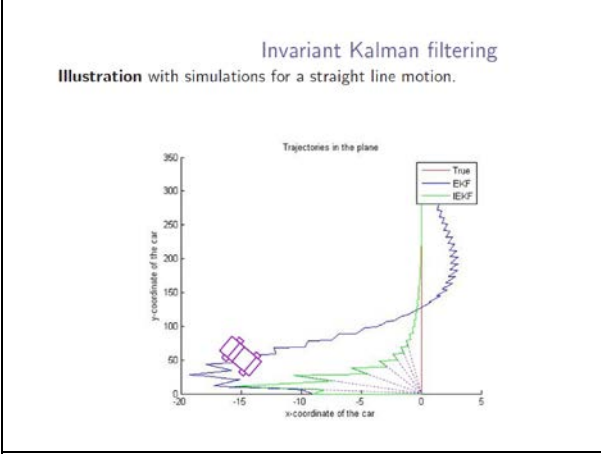
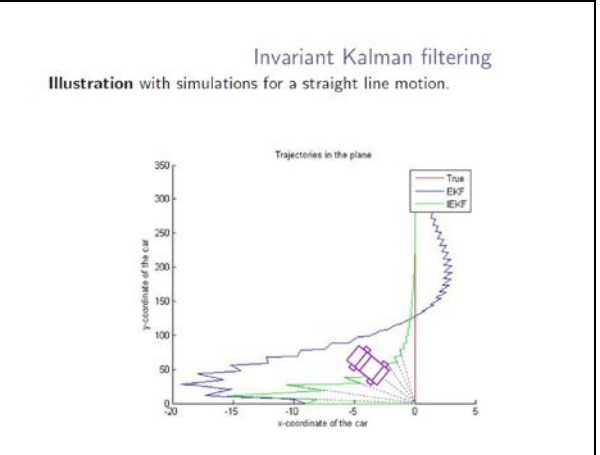
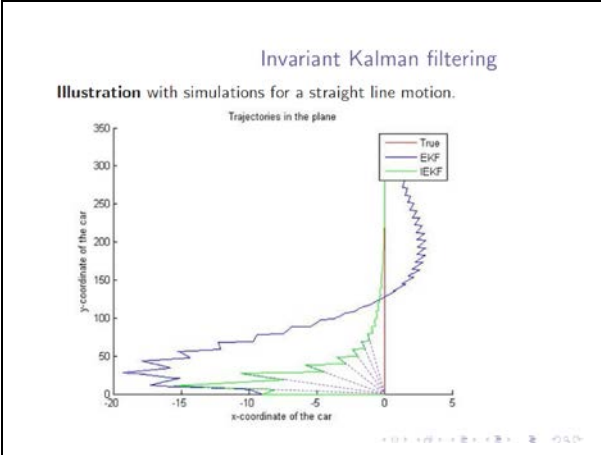


Figure: Position error of the EKF for GPS and perfect odometry, with initial position known and initial heading unknown. The IEKF initialized with a large error (40°) even beats the EKF initialized with a small error (5°).

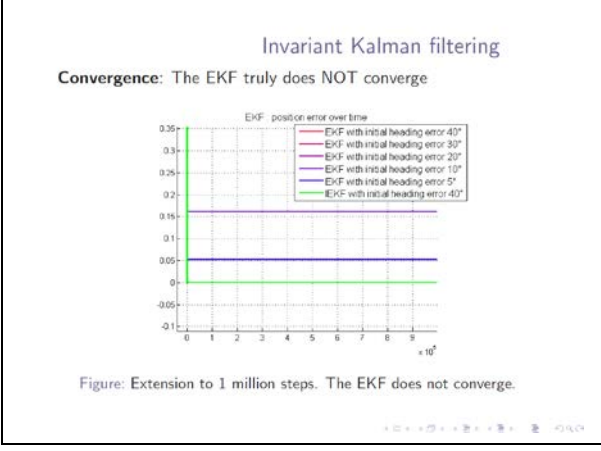


Figure: Extension to 1 million steps. The EKF does not converge.

Conclusion and perspectives

The Invariant consists of using alternative coordinates dictated by symmetries for EKF design.

Industrial application: the problem we considered serves as a simplified in flight alignment problem, an inevitable task for industrial high precision inertial navigation

Other applications (Signal Processing): Object tracking with radars (see next talk by Marion Pilté).

Other applications (Robotics): Visual inertial odometry (see work of Martin Brossard).



Invariant Extended Kalman Filter for Target Tracking *Filtre de Kalman Etendu Invariant pour Pistage de Cibles*

Marion Pilté¹, Silvère Bonnabel², and Frédéric Barbaresco³

¹Mines ParisTech & Thales Air Systems, marion.pilte@mines-paristech.fr

²Mines ParisTech, silvere.bonnabel@mines-paristech.fr

³Thales Air Systems, frederic.barbaresco@thalesgroup.com

Keywords: Kalman, tracking, Lie groups

Mots-clefs: Kalman, pistage, groupes de Lie

Abstract:

A 3D target model expressed in intrinsic coordinates will be developed in this article. The frame used is the Frenet-Serret frame, that is a practical frame to represent the commands a pilot can have on his aircraft for instance. A quite accurate description of the possible motions of an aircraft is to assume the commands are piecewise constant. Once the target model is derived, a filtering algorithm is needed to perform state estimation. As the target model is not expressed in a vectorial space, but rather in a Lie group setting, a novel algorithm, based on results from the inertial navigation field has to be established. This new filter is called the Invariant Extended Kalman Filter (IEKF).

Résumé:

Cet article présente un modèle de cible en 3D et en coordonnées intrinsèques. Le repère de Frenet-Serret est utilisé, il permet de représenter les commandes qu'un pilote peut avoir sur son appareil par exemple. Une description relativement réaliste des mouvements possibles d'un avion est de considérer des commandes constantes par morceaux. Une fois le modèle de cible établi, il faut un algorithme de filtrage pour faire l'estimation de l'état. Comme le modèle de cible n'est pas exprimé dans un espace vectoriel, comme c'est le cas habituellement, mais dans un groupe de Lie, il faut développer un algorithme de filtrage nouveau pour le pistage radar, qui s'inspire de ce qui existe déjà dans le domaine de la navigation inertielle, à savoir le filtre de Kalman étendu invariant.

1 Introduction

One application of radar target tracking is to maintain the tracks initiated within the beam of the radar. In order to be sure the target is not lost, an estimation of the state of the target is necessary, and more specifically, a precise position and velocity estimation is needed. For other radar applications, such as target guiding, a very precise estimation of the velocity vector is essential (among which its direction is of crucial importance). Filtering algorithms are thus very popular among the radar community to perform state estimation.

To perform target tracking, two elements are essential and complementary. The performances of the estimation is due to the target model on the one hand, and to the filtering algorithm on the other hand. First the target model must be accurate enough to describe the possible motions of different classes of targets. Indeed a radar is required to track any aircraft, missiles, boats and all possible types of targets. The accuracy required for each target is set by the client of course, and a single radar cannot track with a high accuracy all types of targets, but it has to keep track of all of them. The target model must thus be loose enough to take into account all these different classes of targets. The other element needed is a filtering algorithm. It is fed with the target model and proceeds in two steps. The first step is merely the propagation of the target model. The second step occurs after the algorithm receives a measurement from the radar, and it can thus adjust the prediction made with the measurement received. The filtering algorithm outputs the state estimation along with its covariance, giving the confidence one can have in the estimation.

In industrial applications, most target models are linear ones, among which we find the famous Singer model [1] or constant velocity models. For the filtering algorithms, the linear Kalman filter [2], or the Interacting Multiple Model (IMM) [3] that runs several filters in parallel are very popular. For nonlinear target or measurement models, the Extended Kalman Filter (EKF) [4] is the most well-known, although it is not very robust. A robustified solution is to use the Castella noise-adaptive algorithm of [5]. Other filtering techniques include particle filters [6], or the Rao-Blackwell particle filter [7], which is a refined particle filter that requires less particles.

However, targets are more maneuvering nowadays, and the linear target models are not efficient enough with the new generation of targets. In this paper we thus propose a target model expressed in intrinsic coordinates, based on the 3D Frenet-Serret frame. Our target model is close to the ones of [8], [9] and [10]. A new Kalman-based filtering algorithm is derived to match specifically this model, as it is not expressed in a vectorial space as usual. This filter is called the Invariant Extended Kalman Filter (IEKF), see for example [11] or [12]. Some specific properties of the IEKF will be highlighted, among which the fact that it is more robust than an EKF.

The paper is organized as follows. In section 2 we derive the target model from the Frenet-Serret frame evolution. In section 3, we detail the construction of an appropriate filter, the IEKF of [13] adapted to the target tracking problem. Finally, in section 4 we show the performances of the designed estimation method.

2 Target model in the Frenet-Serret frame

The idea is to express the target model in intrinsic coordinates, and to model the evolution by assuming the target undergoes constant command motions (as the ones a pilot would apply on an aircraft for instance). The commands are represented by piecewise constant parameters, as it is described below.

2.1 Derivation of the model equations

The target model is based on the Frenet-Serret frame. This is inspired by the model used in [14]. The three vectors of the frame are called the tangential vector, T_t , the normal vector, N_t and the binormal vector, B_t . Their evolution are known with respect to the tangential velocity u_t , the curvature κ_t and the torsion $\tilde{\tau}_t$ of the curve, as described in (1). Then we let $\gamma_t = u_t \kappa_t$ and $\tau_t = u_t \tilde{\tau}_t$ be the curvature and the torsion, with a slight abuse of language. In the simulations of Section 4, we will plot κ_t and $\tilde{\tau}_t$, the usual definitions of the curvature and the torsion.

$$\frac{dT}{dt} = u\kappa N, \quad \frac{dN}{dt} = u(-\kappa T + \tilde{\tau}B), \quad \frac{dB}{dt} = -u\tilde{\tau}N \quad (1)$$

The state of the target is then composed of the rotation matrix $R_t = (T_t \ R_t \ B_t)$, the cartesian position of the target x_t , the curvature of the trajectory γ_t , its torsion τ_t and the tangential velocity of the target u_t . The state is explicitly defined by (2). It does not belong to a vectorial space due to the presence of the rotation matrix R_t .

$$X_t = (R_t \ x_t \ \gamma_t \ \tau_t \ u_t) \quad (2)$$

The evolution of this state can be derived from the Frenet-Serret equations, and from the choice of piecewise constant commands. These commands are represented by the parameters γ_t, τ_t, u_t . In equation (3), we have let $(a)_\times \in \mathbb{R}^{3 \times 3}$ denote the skew-symmetric matrix associated with the cross product with vector $a \in \mathbb{R}^3$. Let us also call $v_t = (u_t \ 0 \ 0)^T$, and $\omega_t = (\tau_t \ 0 \ \gamma_t)^T$. Finally the noises are supposed to be white and gaussian, and are denoted $w_t^\omega, w_t^\gamma, w_t^\tau, w_t^u$; and they account for small changes over time, but also for the jumps in the piecewise constant commands, as it will be explained in Section 4.

$$\frac{dx_t}{dt} = R_t v_t + w_t^x, \quad \frac{dR_t}{dt} = R_t (\omega_t + w_t^\omega)_\times, \quad \frac{d\gamma_t}{dt} = 0 + w_t^\gamma, \quad \frac{d\tau_t}{dt} = 0 + w_t^\tau, \quad \frac{du_t}{dt} = 0 + w_t^u \quad (3)$$

2.2 Particular form of the state

The state is in fact composed of two fairly different parts. We can separate a matricial part and a vectorial part in it. Indeed, the translation and rotation can be considered as an element of $SE(3)$, and we call

$$\chi_t = \begin{pmatrix} R_t & x_t \\ 0_{1,3} & 1 \end{pmatrix}$$

The group $SE(3)$ of rotations and translations in 3D is of dimension 6. The other part of the state is composed of the curvature, the torsion and the tangential velocity:

$$\zeta = \begin{pmatrix} \gamma_t \\ \tau_t \\ u_t \end{pmatrix}$$

Before going further, let us define some basic operations on the Lie group $SE(3)$. A matrix Lie Group is a set of invertible matrices, stable by multiplication and inversion. A Lie group is differentiable. One can thus define a tangent space at the neutral element to the group, called a Lie algebra. A Lie algebra is a vectorial space, equipped with an intern bilinear multiplication. $SE(3)$ describes the possible motions of a point mass in the 3D space. It is defined as follows:

$$SE(3) = \left\{ \begin{pmatrix} R & x \\ 0_{1,3} & 1 \end{pmatrix}, x \in \mathbb{R}^3, R \in \mathcal{M}_3, RR^T = R^T R = I_3, \det(R) = 1 \right\}$$

We can represent the associated Lie algebra as:

$$\mathfrak{se}(3) = \left\{ \left(\begin{array}{cccc} 0 & -c & b & \alpha \\ c & 0 & -a & \beta \\ -b & a & 0 & \gamma \\ 0 & 0 & 0 & 0 \end{array} \right), \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3, \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \in \mathbb{R}^3 \right\}$$

We call $(\cdot)_\times$ the operator

$$(\cdot)_\times : \mathbb{R}^3 \rightarrow \mathcal{M}_3, \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{pmatrix}$$

It is then possible to write the evolution equations (3) in a slightly more compact way, as in (4) with $\omega_t = (\tau_t \ 0 \ \gamma_t)^T$ and $v_t = (u_t \ 0 \ 0)^T$. To do this, one must introduce the corresponding noises $w_t^x = \begin{pmatrix} (w_t^R)_\times & w_t^x \\ 0_{1 \times 3} & 0 \end{pmatrix}$, and $w_t^\zeta = (w_t^\gamma, w_t^\tau, w_t^u)$ and the matrix ν_t as follows:

$$\nu_t = \begin{pmatrix} 0 & -\gamma_t & 0 & u_t \\ \gamma_t & 0 & -\tau_t & 0 \\ 0 & \tau_t & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\frac{d\chi_t}{dt} = \chi_t(\nu_t + w_t^x) = \begin{pmatrix} R(\omega_t)_\times & Rv \\ 0_{1 \times 3} & 0 \end{pmatrix}, \quad \frac{d\zeta_t}{dt} = 0 + w_t^\zeta \quad (4)$$

The measurements are assumed to be in cartesian coordinates, and occur at times $t_n, n \in \mathbb{N}$, and are thus expressed as (5), with $V_n \in \mathbb{R}^3$ a gaussian independant white noise.

$$Y_n = x_{t_n} + V_n \quad (5)$$

The classical Kalman filters cannot be applied to this formulation of the target state. However, if we assume that ζ_t , the vectorial part of the state is known, then χ_t follows a left-invariant differential equation. We thus introduce the Invariant Extended Kalman Filter, as in [11], and extend it to the case where the velocity, the curvature and the torsion are not known, as in [15] or [12].

3 Estimation algorithm

In this section, we derive the filter's equations to perform the estimation.

3.1 Similarities with the Invariant theory

χ_t verifies an equation of the type (6).

$$\frac{d}{dt}\chi_t = f_{i_t}(\chi_t) + \chi_t w_t \quad (6)$$

where $i_t = \zeta_t \in \mathbb{R}^3$ is a known input (we suppose ζ_t known for the moment), w_t is a continuous white gaussian noise, and f satisfies the condition:

$$f_i(ab) = af_i(b) + f_i(a)b - af_i(Id)b \quad (7)$$

for all $(i, a, b) \in \mathbb{R}^3 \times SE(3) \times SE(3)$. An Invariant EKF can thus be designed to estimate χ .

In the case where the input ζ_t is not known, that is of interest in the target tracking problem, the algorithm can be adapted to treat χ as a Lie group part, and ζ as a standard vectorial part.

The system satisfies thus (8).

$$\begin{cases} \frac{d}{dt}\chi_t &= f_i(\chi_t, \zeta_t) + \chi_t w_t^x \\ \frac{d}{dt}\zeta_t &= g(\zeta_t) + w_t^\zeta \end{cases} \quad (8)$$

where in our case

$$f_\zeta : \begin{pmatrix} R & x \\ 0_{1,3} & 1 \end{pmatrix} \rightarrow \begin{pmatrix} R(\omega)_\times & Rv \\ 0_{1,3} & 0 \end{pmatrix}, \quad g(\zeta) = 0_{3,1}$$

We can also write the observations (5) with the help of the Lie group setting:

$$Y_n = \chi_{t_n} \begin{pmatrix} 0_{3,1} \\ 7\mathbf{B} \end{pmatrix} + \begin{pmatrix} V_n \\ 0 \end{pmatrix}$$

The condition (7) for f is easily verified.

3.2 Derivation of the algorithm

3.2.1 Error definition

The classical definition of the error for a vectorial state X_t , $\eta = \hat{X}_t - X_t$ does not hold here. Indeed, if χ_1 and χ_2 belong to the Lie group $SE(3)$, there is no reason why $\chi_1 - \chi_2$ should also belong to this same Lie group. We define the error differently depending on which part of the state we are considering:

$$\begin{aligned}\eta_t^x &= \chi_t^{-1} \hat{\chi}_t \\ \eta_t^\zeta &= \hat{\zeta}_t - \zeta_t\end{aligned}$$

More explicitly the global error $\eta_t = (\eta_t^x, \eta_t^\zeta)$ is defined as:

$$\eta = \begin{pmatrix} \eta_t^R \\ \eta_t^x \\ \eta_t^\gamma \\ \eta_t^\tau \\ \eta_t^u \end{pmatrix} = \begin{pmatrix} R_t^T \hat{R}_t \\ R_t^T (\hat{x}_t - x_t) \\ \hat{\gamma}_t - \gamma_t \\ \hat{\tau}_t - \tau_t \\ \hat{u}_t - u_t \end{pmatrix} \quad (9)$$

3.2.2 Linearization of the error and propagation step

The propagation equations are:

$$\frac{d}{dt} \hat{\chi}_t = f_{\omega_t, u_t}(\hat{\chi}_t) = \hat{\chi}_t \hat{\nu}_t, \quad \frac{d}{dt} \hat{\zeta}_t = g(\hat{\zeta}_t) = 0 \quad (10)$$

Now let us assume once again (for the last time) that ζ_t is known. We can compute the error η_t^x evolution. This gives:

$$\frac{d}{dt} \eta_t^x = \nu_t \eta_t - \eta_t \nu_t - \eta_t w_t^x$$

This equation has the particular property that it does not depend on the predicted state $\hat{\chi}_t$ at all. This is due to the property (7) of the evolution.

As for our radar tracking application ζ_t is not known the evolution of η^x is slightly modified and it writes

$$\frac{d}{dt} \eta_t^x = \hat{\nu}_t \eta_t - \eta_t \nu_t - \eta_t w_t^x \quad (11)$$

indeed, the matrix ν_t depends on the vector ζ_t , so it has to be estimated as well. The evolution of η_t^ζ is more conventional:

$$\frac{d}{dt} \eta_t^\zeta = 0 + w_t^\zeta \quad (12)$$

To linearize equations (11) and (12), see [11], we let $\eta_t^R \approx I_3 + (\xi_t^R)_\times$. This means that $\xi_t^R \in \mathbb{R}^3$ is a small instantaneous rotation vector. We also let $\xi_t^x = \eta_t^x$, $\xi_t^\gamma = \eta_t^\gamma$, $\xi_t^\tau = \eta_t^\tau$ and $\xi_t^u = \eta_t^u$. Then we mimic the EKF methodology, and perform a first order linearization in the components of ξ , and we also neglect terms of order $\|\xi\| \|w\|$. To do this, we use a property of Lie groups: $(\xi_t^R)_\times (\hat{\omega}_t)_\times - (\hat{\omega}_t)_\times (\xi_t^R)_\times = (\xi_t^R \times \hat{\omega}_t)_\times$. This allows to identify the term $\frac{d}{dt} \xi_t^R$ using that $(a)_\times = (b)_\times \implies a = b$. We use the same denomination for ξ as for η :

$$\xi_t = \begin{pmatrix} \xi_t^R \\ \xi_t^x \\ \xi_t^\gamma \\ \xi_t^\tau \\ \xi_t^u \end{pmatrix} \in \mathbb{R}^9$$

During the propagation step, the error evolves as:

$$\frac{d}{dt} \xi_t = A_t \xi_t + w_t$$

with

$$A_t = \begin{pmatrix} 0 & -\hat{\gamma}_t & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ \hat{\gamma}_t & 0 & -\hat{\tau}_t & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \hat{\tau}_t & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\hat{\gamma}_t & 0 & 0 & 0 & -1 \\ 0 & 0 & -\hat{u}_t & \hat{\gamma}_t & 0 & -\hat{\tau}_t & 0 & 0 & 0 \\ 0 & \hat{u}_t & 0 & 0 & \hat{\tau}_t & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

This permits to write the covariance evolution during the propagation: the covariance matrix evolves with the Riccati equation (13).

$$\frac{d}{dt}P_t = A_t P_t + P_t A_t^T + Q_t \quad (13)$$

3.2.3 Gain computation and update step

The update of the state writes:

$$\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp(L_n^X(\hat{\chi}_{t_n}^{-1} Y_n)), \quad \hat{\zeta}_{t_n}^+ = \zeta_{t_n} + L_n^\zeta(\hat{\chi}_{t_n}^{-1} Y_n) \quad (14)$$

more explicitly, this can be developed as:

$$\begin{pmatrix} \hat{R}_{t_n}^+ \\ \hat{x}_{t_n}^+ \\ \hat{\gamma}_{t_n}^+ \\ \hat{\tau}_{t_n}^+ \\ \hat{u}_{t_n}^+ \end{pmatrix} = \begin{pmatrix} \hat{R}_{t_n} \exp_m[(\delta_\omega)_\times] \\ \hat{R}_{t_n} \hat{x}_{t_n} + B(\delta_\omega) \delta_x \\ \hat{\gamma}_{t_n} + \delta_\gamma \\ \hat{\tau}_{t_n} + \delta_\tau \\ \hat{u}_{t_n} + \delta_u \end{pmatrix} \quad (15)$$

where $(\delta_\omega \ \delta_x \ \delta_\gamma \ \delta_\tau \ \delta_u)^T = L_n(\hat{R}_{t_n}(Y_n - \hat{x}_{t_n}))$, \exp_m denotes the matrix exponential map in \mathcal{M}_3 , and

$$B(\delta_\omega) = I_3 + \frac{1 - \cos\|\delta_\omega\|}{\|\delta_\omega\|^2}(\delta_\omega)_\times + \frac{\delta_\omega - \sin\|\delta_\omega\|}{\|\delta_\omega\|^3}[(\delta_\omega)_\times]^2$$

The gain matrix $L_n \in \mathbb{R}^{9 \times 3}$ is computed with the Riccati equation (13), as will be explained in the following. The innovation is defined as $\hat{\chi}_{t_n}^{-1} Y_n = \hat{R}_{t_n}^T(Y_n - \hat{x}_{t_n})$, it verifies:

$$\hat{R}_{t_n}^T(Y_n - \hat{x}_{t_n}) = \hat{R}_{t_n}^T(x_{t_n} - \hat{x}_{t_n}) + R_{t_n}^T V_n = -(\eta_{t_n}^R)^{-1} \eta_{t_n}^x + \hat{R}_{t_n}^T V_n$$

thus, as $\eta_t^x = \xi_t^x$ and $(\xi_t^R)_\times \xi_t^x$ is of order two, then the linearization gives $\hat{R}_{t_n}^T(Y_n - \hat{x}_{t_n}) \approx -H\xi + \hat{R}_{t_n}^T V_n$, where $H \in \mathbb{R}^{3 \times 9}$ is defined as $H = (0_{3 \times 3} \ I_3 \ 0_{3 \times 3})$. We are now able to derive the update ξ_t^+ of ξ_t from (15) and (9), which gives:

$$\xi_{t_n}^+ = \xi_{t_n} - L_n \left[(0_{3 \times 3} \ I_3 \ 0_{3 \times 3}) \xi_{t_n} - \hat{R}_{t_n}^T V_n \right] \quad (16)$$

Finally, the Kalman gain can be computed with the full Riccati equations, with Q_t and N_n the process and measurement noise covariance respectively:

$$\begin{aligned} \frac{d}{dt}P_t &= A_t P_t + P_t A_t^T + Q_t \\ S_n &= H P_{t_n} H^T + \hat{R}_{t_n}^T N_n \hat{R}_{t_n} \\ L_n &= P_{t_n} H^T S^{-1} \\ P_{t_n}^+ &= (I_9 - L_n H) P_{t_n} \end{aligned}$$

3.2.4 Summary of the filter's equations

To sum up the results obtained before, we can write extensively the filter's equations.

1. Propagation step:

- Solve $\frac{d}{dt}\hat{\chi}_t = \hat{\chi}_t \hat{\nu}_t$ and $\frac{d}{dt}\hat{\zeta}_t = 0$
- Solve the Riccati equation $\frac{d}{dt}P_t = A_t P_t + P_t A_t^T + Q_t$

2. Update step:

- Compute the innovation $z_n = \hat{R}_{t_n}(Y_n - \hat{x}_{t_n})$
- Compute the Kalman gain $L_n = P_{t_n} H (H P_{t_n} H^T + \hat{R}_{t_n}^T N_n \hat{R}_{t_n}^T)^{-1}$ with $H = (0_{3,3} \ I_3 \ 0_{3,3})$
- Update the state $\hat{\chi}_{t_n}^+ = \hat{\chi}_{t_n} \exp((L_n z_n)_{1:6})$ and $\hat{\zeta}_{t_n}^+ = \zeta_{t_n} + (L_n z_n)_{7:9}$
- Update the covariance $P_{t_n}^+ = (I_9 - L_n H) P_{t_n}$

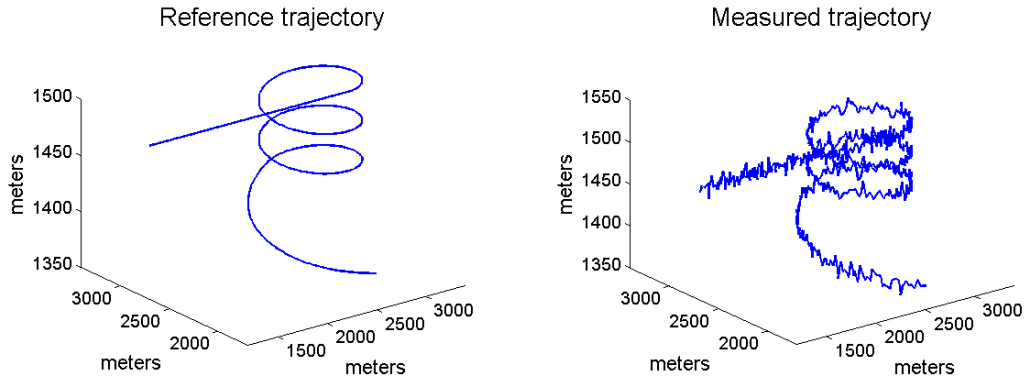


Figure 1 – Reference trajectory and trajectory with simulated noise

4 Results

The target model along with the filtering algorithm can estimate some trajectories that do not stay in a plane, and which require the use of the Interacting Multiple Model (IMM) if we use standard models such as constant velocity, constant horizontal or vertical turns or based on the linear Singer model [1]. The type of trajectory we want to track is shown on figure 1, and the estimation by the IEKF on figure 2.

The trajectory shown in figure 1 has been elaborated in three parts. The first part is a straight line motion with constant velocity, the second part is a helical motion, with different, but constant tangential velocity, torsion and curvature, and the last part is a circle in a plane, with again different velocity and curvature (the torsion is zero since we lie in a plane). Measurement noise is added by hand, independently on the three cartesian position coordinates, the amplitude of the measurement noise is visible on the figure on the right. It is quite high, so that it is more challenging for the filtering algorithm.

The estimations made by an IEKF of the position x , the curvature κ , the torsion $\tilde{\tau}$ and the norm of the velocity u are displayed in figure 2. The parameters are poorly initialized on purpose to see the behaviour of the filter when confronted to high initial errors. In practice the position is relatively well known, as well as the norm of the velocity, but the curvature or the torsion are not.

The results show that the position is very accurately estimated. The process noise tuning was done such that the estimations were precise on constant motions during the trajectory (especially for the norm of the velocity and the curvature). So for this tuning, the estimations have a small delay after the jumps, which can be reduced by increasing the process noise on these parameters but at the cost of a lesser precision on the constant parts. The process noise tuning has the same issues as for a classical Kalman filter, where one must have a balance between the precision during constant motions and the capacity to react rapidly to jumps.

The estimation of the torsion is more difficult. Indeed, the torsion comes from a third derivative of the position, so it is barely observable. In theory, it is observable, but in practice since we measure noisy position, it is hard to recover. However, we see that its estimation eventually converge, and that it becomes better after the first jump, when the trajectory is not a straight line anymore.

The lack of precision of the torsion estimation does not really affect the quality of the estimation of the other parameters, and it is not in itself a major problem. Indeed, what is most relevant to assess the performances of a filtering algorithm is either the precision of the position, and the ability of the algorithm to filter the noise, or the precision of the velocity vector, to be sure where the target is actually heading, and at which speed. It is very difficult to tune a filter to achieve both very efficiently, but we can obtain a very good balance with the IEKF. However, the torsion (and the curvature) are not as essential to the user. But we cannot suppress them of the state as they bring some necessary degrees of freedom in the trajectory (the torsion is what allows to get out of a plane).

5 Conclusions

In this paper we have presented a target model in 3D, and based on the Frenet-Serret frame. This particular frame allows us to express the model in intrinsic coordinates, which can best represent the commands of a pilot in his aircraft for instance. To perform estimation with this target model, we have seen that an Invariant Extended Kalman Filter is most appropriate, because it is suited to the evolution of a part of the state, which is not vectorial as it is usually the case.

Another advantage of the IEKF is to have more stability properties than an EKF. Indeed, we have seen that the evolution of the error does not depend on the predicted position or rotation matrix of the target, contrary to an EKF.

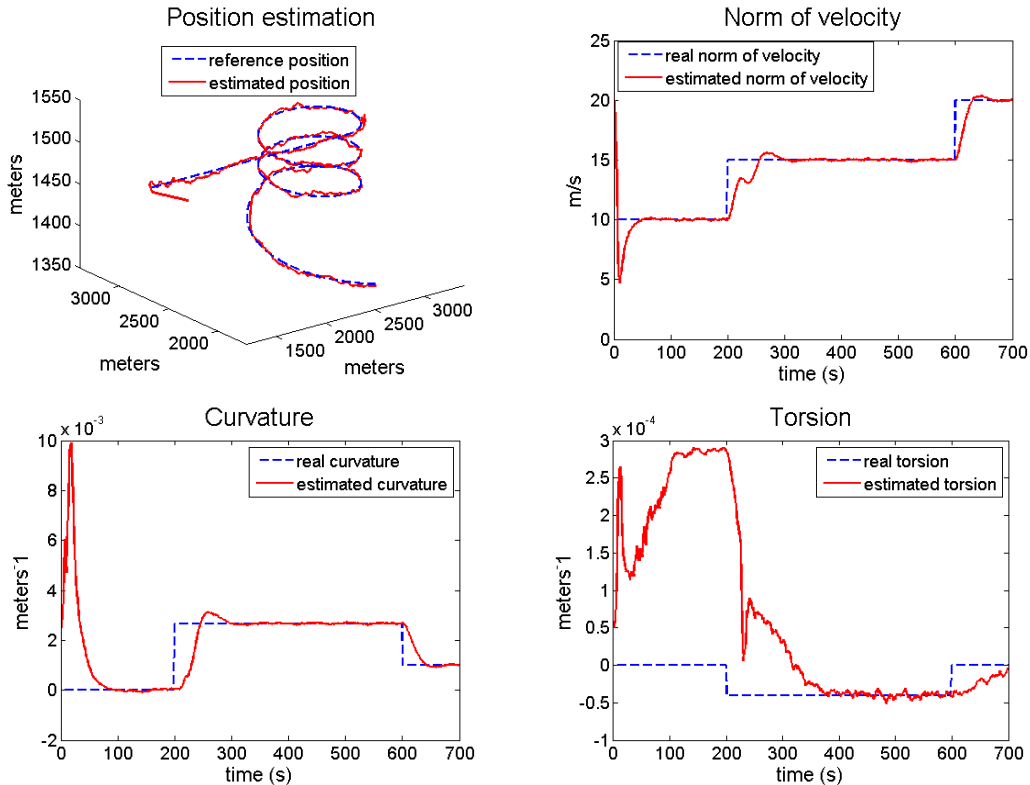


Figure 2 – Estimation of the position (top left) and of parameters u_t (top right), γ_t (bottom left), τ_t (bottom right)

It is very hard to use this model with other filtering algorithms, since they are all designed to match a vectorial state model. The orthogonality constraints implied by the use of the Frenet-Serret frame are highly nonlinear, and an EKF has difficulties with this type of formulations.

Acknowledgments

This work is supported by a DGA-MRIS scholarship and by Thales Air Systems.

6 References

- [1] R. A. Singer, “Estimating optimal tracking filter performance for manned maneuvering targets,” *IEEE Transactions on Aerospace and Electronic Systems*, no. 4, pp. 473–483, 1970.
- [2] R. E. Kalman, “A new approach to linear filtering and prediction problems,” *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.
- [3] Y. Bar-Shalom, K. Chang, and H. A. Blom, “Tracking a maneuvering target using input estimation versus the interacting multiple model algorithm,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 25, no. 2, pp. 296–300, 1989.
- [4] Y. Bar-Shalom, X. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley, 2004.
- [5] F. R. Castella, “An adaptive two-dimensional Kalman tracking filter,” *IEEE Transactions on Aerospace and Electronic Systems*, no. 6, pp. 822–829, 1980.
- [6] A. Doucet, N. J. Gordon, and V. Krishnamurthy, “Particle filters for state estimation of jump Markov linear systems,” *Signal Processing, IEEE Transactions on*, vol. 49, no. 3, pp. 613–624, 2001.
- [7] A. Doucet, N. De Freitas, K. Murphy, and S. Russell, “Rao-blackwellised particle filtering for dynamic bayesian networks,” in *Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence*, pp. 176–183, Morgan Kaufmann Publishers Inc., 2000.

- [8] T. Bullock and S. Sangsuk-Iam, “Maneuver detection and tracking with a nonlinear target model,” in *Decision and Control, 1984. The 23rd IEEE Conference on*, vol. 23, pp. 1122–1126, IEEE, 1984.
- [9] R. Bishop and A. Antoulas, “Nonlinear approach to aircraft tracking problem,” *Journal of Guidance, Control, and Dynamics*, vol. 17, no. 5, pp. 1124–1130, 1994.
- [10] P. Bunch and S. Godsill, “Dynamical models for tracking with the variable rate particle filter,” in *Information Fusion (FUSION), 2012 15th International Conference on*, pp. 1769–1775, IEEE, 2012.
- [11] A. Barrau and S. Bonnabel, “The invariant extended kalman filter as a stable observer,” *IEEE Transactions on Automatic Control*, 2017.
- [12] M. Pilté, S. Bonnabel, and F. Barbaresco, “Tracking the Frenet-Serret frame associated to a highly maneuvering target in 3D.” working paper or preprint, July 2017.
- [13] A. Barrau and S. Bonnabel, “Intrinsic filtering on lie groups with applications to attitude estimation,” *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, 2015.
- [14] S. J. Godsill, J. Vermaak, W. Ng, and J. F. Li, “Models and algorithms for tracking of maneuvering objects using variable rate particle filters,” *Proceedings of the IEEE*, vol. 95, no. 5, pp. 925–952, 2007.
- [15] M. Pilté, S. Bonnabel, and F. Barbaresco, “An innovative nonlinear filter for radar kinematic estimation of maneuvering targets in 2d,” in *Radar Symposium (IRS), 2017 18th International*, pp. 1–10, IEEE, 2017.

Time-Lapse Estimation for Optical Telescope Sequences

Estimation par Intervalle de Temps pour les Séquences de Télescopes Optiques

Mark Campbell¹ and Daniel Clark²

¹*School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, United Kingdom, mc318@hw.ac.uk*

²*Télécom SudParis, Institut Mines-Télécom, daniel.clark@telecom-sudparis.eu*

Keywords: Sensor Calibration, Multi-Target Tracking

Mots-clefs: Calibrage du Capteur, Suivi Multi-Cible

Abstract:

Typical tracking scenarios rely on the assumption that there is a constant time lapse between observations. In real life applications, this assumption is often untrue. In Space Situational Awareness (SSA) applications accurate target estimation is of importance to obtain orbital information. This paper presents recent developments in multi target detection and tracking techniques, exploiting the Single Cluster Probability Hypothesis Density (SC-PHD) filter, in order to jointly estimate the dynamic objects and the time lapse between images.

Résumé:

Les scénarios de suivi typiques reposent sur l'hypothèse qu'il y a un laps de temps constant entre les observations. Dans les applications réelles, cette hypothèse est souvent fautive. Dans les applications SSA, l'estimation précise de la cible est importante pour obtenir des informations orbitales. Cet article présente les développements récents dans les techniques de détection et de suivi de cibles multiples, en exploitant le filtre SC-PHD, afin d'estimer conjointement les objets dynamiques et le laps de temps entre les images.

1 Introduction

The usage of low cost ground-based Complementary Metal Oxide Semiconductor (CMOS) optical sensors to track objects in Low Earth Orbit (LEO) and Geostationary Earth Orbit (GEO) has dramatically increased in popularity in the past decade, leading to a chance of integrating these sensors into the formal SSA environment. These Consumer Off-The-Shelf (COTS) cameras and optics are inexpensive solutions and provide a relatively wide Field of View (FoV). A central challenge in the exploitation of these sensors is the extraction of multiple detections from a single image due to the wide FoV: some detections can stem from moving objects (e.g. satellites, debris), some others from statics objects (e.g. stars), some of them may be artefacts from the extraction process or *false alarms*. The association between the objects and the detections is unknown, as are the number of objects, both static and moving, in the FoV. In these conditions, exploiting the output of a sensor in order to estimate the number and trajectories of the detected objects becomes a multi-object estimation problem.

The Probability Hypothesis Density (PHD) filter [1] was designed as an inexpensive filtering solution in the the context of multi-object filtering. Extensions to the PHD filter such as the Second-Order Probability Hypothesis Density (SO-PHD) [2] and the Cardinalized Probability Hypothesis Density (CPHD) [3] provide additional information at the expense of computational complexity. Solving this multi-object scenario is crucial so that methods such as Initial Orbit Determination (IOD) and astrometry can take place.

The multi-target tracking methods assume a constant time-lapse between observations e.g. a constant time between image captures. However this is often not the case when operating on real datasets. This could be due to a number of reasons: operator error, corrupted or missing metadata or even sensor failure. This paper explores the exploitation of multi-object tracking algorithms developed from the Finite Set Statistics (FISST) framework [1], known as the SC-PHD filter [4, 5], in order to jointly estimate the time-lapse between the images and the objects' states in a sequence of images produced from an optical sensor. Variations of this method have been used to register microscopy [6] and astronomy [7] images.

This paper is organized as follows. Sec. 2 presents the principle of joint multi-object filtering and sensor state estimation, and introduces the two multi-object filters exploited in the paper. Sec. 3 presents the details employed for the implementation of the algorithms. Sec. 4 describes the simulation tests and Sec. 5 discusses the results. Finally, the conclusions are provided in Sec. 6.

2 Joint Multi-Object Filtering and Sensor State Estimation

2.1 Single Cluster Probability Hypothesis Density (SC-PHD) filter

The joint calibration and tracking method exploits a Simultaneous Localisation and Mapping (SLAM) based approach [4] known as the Single Cluster Probability Hypothesis Density (SC-PHD) filter. There are two sources

of uncertainty to be estimated, represented by two different *processes*:

- The sensor process Ψ estimates the sensor state: in this application, the time-lapse between subsequent images,
- The object process Φ estimates the number and states of moving objects: in this application, the orbiting objects moving through the image sequence.

Since the number and states of the objects is assumed unknown and possibly varying across the image sequence, the object process Φ is estimated through multi-object filters derived from the FISST framework: in the scope of this paper, the PHD filter [1] and the SO-PHD filter [2] [8]. These two solutions are explained in more details in Sec. 2.2.

2.2 Multi Object Filtering

This section focuses on the estimation of the object process Φ . As seen in Sec. 2.1, the estimation of the object process is conditioned upon the sensor state (i.e. time lapse between observations). In practical terms, a multi-object filter estimating the object process is maintained for each possible sensor state y . In this section, y_k denotes an arbitrary sensor state¹ at time k (i.e., in the k -th image of the sequence). Each object is described by its state x in the (*single*) *target state space* $\mathcal{X} \subseteq \mathbb{R}^d$, describing the physical characteristics of the object. An object may enter or leave the sensor FoV at any time during the image sequence, and thus the number of objects at any time is unknown and needs to be estimated. The multi-object state is represented by a Random Finite Set (RFS) Φ_k , a random object whose size and elements are unknown, and whose realization is a set of target states $X_k = \{x_1, x_2, \dots, x_{n_k}\}$ represents a specific *multi-object configuration* at time k [3]. The evolution of the objects' state between time step $k-1$ and k is described by a Markov transition function $t_{k|k-1}$. The number and states of newborn objects is described by a RFS $\Phi_{b,k}$ whose nature depends on the filter; its *first-order moment density* or *intensity* is denoted by $\mu_{b,k}(\cdot)$.

An observation collected from the sensor is described by a state z in the *observation space* $\mathcal{Z} \subseteq \mathbb{R}^d$. The set of collected observation is denoted by Z_k , which contains observations of all the targets in the image at time k . The observation process is plagued by observation noise, missed detections, and false alarms. The observation noise associated to some collected observation z is characterized by the *likelihood* function $\ell_{z,k}(\cdot|y)$, while the probability of detection for each individual object in the sensor FoV is denoted by $p_{d,k}(\cdot|y)$. The number and states of false alarms is described by a RFS $\Phi_{y,k}$ whose nature depends on the filter; its intensity is denoted by $\mu_{y,k}(\cdot|y)$, and its spatial distribution is denoted by $s_{y,k}(\cdot|y)$. For the rest of the paper, the notation $\mu(\mathcal{X})$ will be used for various intensity functions μ to denote the integral $\int_{\mathcal{X}} \mu(x) dx$.

2.2.1 Probability Hypothesis Density (PHD) Filter

The PHD filter [1] was designed as an inexpensive filtering solution in the the context of multi-object filtering, propagating only the intensity μ_k of the object process Φ_k . The key assumptions of the PHD filter is that both the predicted process $\Phi_{k|k-1}$ and false alarm process $\Phi_{y,k}$ are Poisson [1]. The prediction and update steps of the PHD filter are given by

$$\mu_{k|k-1}(x|y) = \mu_{b,k}(x) + \mu_{s,k|k-1}(x|y), \quad (1)$$

$$\mu_k(x|y) = \mu_{\phi,k}(x|y) + \sum_{z \in Z_k} \frac{\mu_{z,k}(x|y)}{\mu_{y,k}(z|y) + \mu_{z,k}(\mathcal{X}|y)}, \quad (2)$$

where the *survival* $\mu_{s,k|k-1}(x|y)$, *missed detection* $\mu_{\phi,k}(x|y)$, and *association* $\mu_{z,k}(x|y)$ terms are defined as

$$\mu_{s,k|k-1}(x|y) = \int p_{s,k}(\bar{x}) t_{k|k-1}(x|\bar{x}) \mu_{k-1}(\bar{x}|y) d\bar{x}, \quad (3)$$

$$\mu_{\phi,k}(x|y) = (1 - p_{d,k}(x|y)) \mu_{k|k-1}(x|y), \quad (4)$$

$$\mu_{z,k}(x|y) = p_{d,k}(x|y) \ell_{z,k}(x|y) \mu_{k|k-1}(x|y). \quad (5)$$

2.3 Second-Order Probability Hypothesis Density (SO-PHD) Filter

Recently, a second-order version of the PHD filter was introduced in [2] that combats some of the issues posed by the restrictive Poisson assumptions of the PHD filter. It propagating not only the intensity μ_k of the object process Φ_k , but also the variance $\text{var}_k(\mathcal{X})$ of the number of objects in the whole target state space \mathcal{X} .

The SO-PHD filter substitutes Panjer assumptions to Poisson assumptions in the PHD filter, thus providing more flexibility in the description of the number of objects and false alarms. More specifically, the Panjer

¹The estimator associated to the sensor state will be presented later in Sec. 2.4

distribution [2] describes the Poisson, binomial and negative binomial distributions in a unified formulation involving two parameters α and β which stand in one-to-one correspondence with the mean and variance of the distribution. Using this property, the SO-PHD filter is able to propagate the variances in the number of objects and number of false alarms via the Panjer parameters of the corresponding distributions.

Before stating the recursion of the SO-PHD filter, some notations inspired by those introduced in [9] for the CPHD filter will be states. The *Pochhammer symbol* or *rising factorial* $(\zeta)_n$ for any $\zeta \in \mathbb{R}$ and $n \in \mathbb{N}$ is defined with

$$(\zeta)_n := \zeta(\zeta + 1) \cdots (\zeta + n - 1), \quad (\zeta)_0 := 1. \quad (6)$$

Let $\alpha_{k|k-1}, \beta_{k|k-1}$ and $\alpha_{\mathbf{s},k}, \beta_{\mathbf{s},k}$ be the parameters of the predicted object process and clutter process at time k , respectively, and define the expression

$$Y_u^b[Z] := \sum_{j=0}^{|Z|} \frac{(\alpha_{k|k-1})_{j+u}}{(\beta_{k|k-1})^{j+u}} \frac{(\alpha_{\mathbf{s},k})_{|Z|-j}}{(\beta_{\mathbf{s},k} + 1)^{|Z|-j}} F_d^{-j-u} e_j^b(Z) \quad (7)$$

for any $Z \subseteq Z_k$ and for $u = 1, 2$, where F_d is the scalar

$$F_d := \int \left[1 + \frac{p_{d,k}(x|y)}{\beta_{k|k-1}} \right] \mu_{k|k-1}(x|y) dx, \quad (8)$$

and the so-called *elementary symmetric functions* [10] e_j^b are given by

$$e_j^b(Z) := \sum_{\substack{Z' \subseteq Z \\ |Z'|=j}} \prod_{z \in Z'} \frac{\mu_{z,k}(\mathcal{X}|y)}{s_{\mathbf{s},k}(z|y)}. \quad (9)$$

Where $s_{c,k}$ denotes the spatial clutter distribution at time k . With the help of Eq. (7), define the *corrective terms* via

$$l_u^b(\phi|y) := \frac{Y_u^b[Z_k]}{Y_0[Z_k]} \quad \text{and} \quad l_u^b(z|y) := \frac{Y_u^b[Z_k \setminus \{z\}]}{Y_0^b[Z_k]} \quad (10)$$

for $u = 1, 2$, and

$$l_2^b(z, z'|y) := \begin{cases} \frac{Y_2^b[Z_k \setminus \{z, z'\}]}{Y_0^b[Z_k]} & \text{if } z \neq z', \\ 0 & \text{otherwise.} \end{cases} \quad (11)$$

Assuming that $p_{s,k}(x) = p_{s,k}$ is constant for all $x \in \mathcal{X}$ at time k , the prediction step of the SO-PHD filter is then given by

$$\mu_{k|k-1}(x|y) = \mu_{b,k}(x) + \mu_{s,k|k-1}(x|y), \quad (12)$$

$$\text{var}_{k|k-1}(\mathcal{X}|y) = \text{var}_{b,k}(\mathcal{X}) + \text{var}_{s,k|k-1}(\mathcal{X}|y), \quad (13)$$

where the survival term for the variance $\text{var}_{s,k|k-1}$ is given by

$$\text{var}_{s,k|k-1}(\mathcal{X}|y) = (p_{s,k})^2 \text{var}_{k-1}(\mathcal{X}|y) + p_{s,k}[1 - p_{s,k}] \mu_{k-1}(\mathcal{X}|y). \quad (14)$$

The variance in the number of newborn objects $\text{var}_{b,k}$ is a parameter of the filter, and allows the operator to describe situations where the information on the number of objects entering the image frame.

The Panjer parameters of the predicted object process $\Phi_{k|k-1}$ are then given by [2]

$$\alpha_{k|k-1} = \frac{\mu_{k|k-1}(\mathcal{X}|y)^2}{\text{var}_{k|k-1}(\mathcal{X}|y) - \mu_{k|k-1}(\mathcal{X}|y)}, \quad (15)$$

$$\beta_{k|k-1} = \frac{\mu_{k|k-1}(\mathcal{X}|y)}{\text{var}_{k|k-1}(\mathcal{X}|y) - \mu_{k|k-1}(\mathcal{X}|y)}. \quad (16)$$

With these, the corrective terms l_u^b (10), (11) can be computed, and the update step of the SO-PHD follows:

$$\mu_k(x|y) = \mu_{\phi,k}(x|y) l_1^b(\phi|y) + \sum_{z \in Z_k} \frac{\mu_{z,k}(x|y)}{s_{\mathbf{s},k}(z|y)} l_1^b(z|y). \quad (17)$$

$$\begin{aligned} \text{var}_k(\mathcal{X}|y) &= \mu_k(\mathcal{X}|y) + \mu_{\phi,k}(\mathcal{X}|y)^2 \left[l_2^b(\phi|y) - l_1^b(\phi|y)^2 \right] \\ &+ 2\mu_{\phi,k}(\mathcal{X}|y) \sum_{z \in Z_k} \frac{\mu_{z,k}(\mathcal{X}|y)}{s_{\mathbf{s},k}(z|y)} \left[l_2^b(z|y) - l_1^b(\phi) l_1^b(z|y) \right] \\ &+ \sum_{z, z' \in Z_k} \frac{\mu_{z,k}(\mathcal{X}|y)}{s_{\mathbf{s},k}(z|y)} \frac{\mu_{z',k}(\mathcal{X}|y)}{s_{\mathbf{s},k}(z'|y)} \left[l_2^b(z, z'|y) - l_1^b(z|y) l_1^b(z'|y) \right]. \end{aligned} \quad (18)$$

2.4 Sensor State Estimation

This section focuses on the estimation of the sensor process Ψ . The state of the sensor is denoted by $y \in \mathcal{Y}$, the sensor space \mathcal{Y} describing the relative time difference of an image frame with respect to the previous frame in the image sequence. Following the Sequential Monte Carlo (SMC) implementation in [11], the information regarding the state of the sensor at time k is described by a probability distribution p_k approximated through a set of weighted particles $\{w_k^i, y_k^i\}_{i=1}^N$, i.e.

$$p_k(\cdot) \simeq \sum_{i=1}^N w_k^i \delta_{y_k^i}(\cdot). \quad (19)$$

Note that the number of particles, N , is maintained constant throughout the scenario. Selecting an appropriate number of particles is a key consideration of the implementation of the filter. A large number of sensor particles provides a more robust estimator, but increases significantly the computational cost of the overall algorithm: recall from the hierarchical structure (see Sec. 2.2) that a multi-object filter, either a PHD or SO-PHD filter is maintained for each possible sensor state, i.e., for each particle state y_k^i .

By construction, the initial states of y_0^i are uniformly distributed between some lower and upper parameters. At each time step, the prediction step of the sensor state follows a Markov transition model $h_{k|k-1}$ denoting the operator's knowledge in the nature of the sensor movement (i.e. linear drift). Each particle state is thus resampled according to the transition model, i.e.

$$y_k^i \sim h_{k|k-1}(\cdot | y_{k-1}^i). \quad (20)$$

The update step of the sensor state works as follows. Since the likelihood function $\ell_{z,k}(x|y)$ associating a collected observation z to an object with state x is dependent on the sensor state therefore the estimation of the object process Φ is also conditioned on the sensor state y and is exploited to update the sensor state distribution as follows:

$$w_k^i \propto \frac{\mathcal{L}(Z_k | \Phi, y_k^i) w_{k-1}^i}{\sum_{i=1}^N w_{k-1}^i}, \quad (21)$$

where the *multi-object likelihood* $\mathcal{L}(Z_k | \Phi, y_k^i)$ quantifies the match between the set of collected observations Z_k and the estimation of the object process Φ conditioned on y_k^i . These multi-object likelihoods can be seen below in Sec.2.4.1 and 2.4.2

After the updated weights w_k^i are calculated, the highest weighted particle is found. The resampling step then occurs by uniformly distributing between some minimum and maximum parameters using the highest weighted particle.

2.4.1 PHD Multi-Object Likelihood

If the object process is estimated through a PHD filter (see Sec.2.2.1), the multi-object likelihood is given by [12]:

$$\mathcal{L}(Z_k | \Phi, y_k^i) = \frac{\prod_{z \in Z_k} [\mu_{z,k}(z|y_k^i) + \int p_{d,k}(x|y_k^i) \ell_k(z|x, y_k^i) \mu_{k|k-1}(x|y_k^i) dx]}{\exp[\int \mu_{z,k}(z|y_k^i) dz + \int p_{d,k}(x|y_k^i) \mu_{k|k-1}(x|y_k^i) dx]}. \quad (22)$$

2.4.2 SO-PHD Multi-Object Likelihood

If the object process is estimated through a SO-PHD filter (see Sec.2.3), the multi-object likelihood is given by

$$\mathcal{L}(Z_k | \Phi, y_k^i) = \sum_{j=0}^{|Z_k|} \frac{(\alpha_{k|k-1})_j}{(\beta_{k|k-1})^j} \frac{(\alpha_{c,k})^{|Z|-j}}{(\beta_{c,k} + 1)^{|Z|-j}} \tilde{F}_d^{-\alpha-j} F_c^{-\alpha_{c,k}-|Z|-j} \sum_{\substack{Z' \subseteq Z_k \\ |Z'|=j}} \prod_{z \in Z'} \mu_{z,k}(\mathcal{X}|y_k^i) \prod_{z' \in Z_k \setminus Z'} \mu_{c,k}(z|y_k^i), \quad (23)$$

where

$$\tilde{F}_d = 1 - \frac{1}{\beta_{k|k-1}} \int p_{d,k}(x|y_k^i) s_{k|k-1}(x|y_k^i) dx, \quad (24)$$

$$F_c = 1 + \frac{1}{\beta_c}, \quad (25)$$

where $s_{k|k-1}(x|y_k^i) = \frac{\mu_{k|k-1}(x|y_k^i)}{\int \mu_{k|k-1}(\bar{x}|y_k^i) d\bar{x}}$ is the spatial distribution associated to the intensity $\mu_{k|k-1}$.

3 Implementation

All of the multi-object filters used in the experiments are implemented using a Gaussian Mixture (GM) approach following [13], [2] and [9]. Measurement driven birth is also used. Since the parent process is implemented using a SMC particle filter approach, a preset number of particles shall be used.

The motion of the sensor state, herein referred to as the time lapse Δ_k , (20) shall be modelled using two motion models, Brownian motion and static motion, each accounting for a different drift observed in common scenarios. The Brownian motion model, or random walk, describes a scenario where the images are perhaps taken manually by the operator at random intervals. The static model describes a fixed offset with some zero mean Gaussian noise possibly due to a control software error. The time lapse state space is characterised by a simple one dimensional space $x = \Delta_k$. When satellites and debris are observed using telescope, a “streak” is left behind in the image due to the long exposure times needed. The length of this streak can provide an estimate of the target’s velocity v , given by:

$$v = \frac{l}{T} \quad (26)$$

Where l and T are the measured streak length and camera exposure time respectively. The inclination (or heading) θ of the streak can also be measured.

This motion using the Nearly Constant Heading (NCH) motion model [14]. It provides a better representation of the target dynamics than a regular Nearly Constant Velocity (NCV) motion model since the targets move along a fairly fixed line. Dynamic targets are described via their x and y position, the speed v and the inclination θ using a four-dimensional state space $\mathcal{X} \subseteq \mathbb{R}^4$, where specific states at time k are of the form

$$x_k = [x_k, y_k, v_k, \theta_k]^T \quad (27)$$

The motion model for the dynamic objects is non-linear so an Extended Kalman Filter (EKF) [15] is used to propagate through the time steps. It is given as follows:

$$t_k(x_k|x_{k-1}) = \mathcal{N}(x_k; \hat{x}_k, Q_k), \quad (28)$$

where the intermediate state \hat{x}_k is obtained with

$$\hat{x}_k = x_{k-1} + \Delta_k v_{k-1} [\cos(\theta_{k-1}), \sin(\theta_{k-1}), 0, 0]^T, \quad (29)$$

and where Q_k is a covariance matrix of the form:

$$Q_k = \Delta_k \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{v,k}^2 & 0 \\ 0 & 0 & 0 & \sigma_{\theta,k}^2 \end{bmatrix}, \quad (30)$$

where $\sigma_{v,k}$ and $\sigma_{\theta,k}$ are the standard deviations of the velocity and inclination, respectively.

4 Simulations

In order to test the algorithm, several scenarios shall be simulated, each using a different time lapse motion model. This simulated data will be generated as follows. Firstly a number of initial targets are generated N_t , the targets’ states are then initialised following the NCH model stated above. The generated x_k, y_k positions are limited to within the sensor’s FoV, the velocities v_k are drawn from a uniform distribution between some minimum v_k^{min} and maximum v_k^{max} velocity parameters and the inclinations are drawn from a uniform distribution such that $0 \leq \theta_k \leq 360$. As stated before the target transition model follows the NCH model Eq.28. Spontaneous target birth may occur at each time step k and is modelled using a Poisson process with rate λ_b . Target death and detection rates are other factors and they are modelled using Bernoulli processes with probabilities of survival p_s and detection p_d . False alarms may also occur at each time step k , in these simulations a Poisson process with rate λ_{fa} is used to represent these clutter process. The clutter process is also assumed to be uniformly distributed over the entire sensor FoV. At each time step measurements z_k are obtained from the simulated targets’ states x_k using the observation model:

$$z_k = H_k x_k + \mathcal{N}(0, R_k), \quad (31)$$

Where H_k is given by

$$H_k = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (32)$$

And R_k is given by

$$R_k = \begin{bmatrix} \sigma_{x,k}^2 & 0 & 0 \\ 0 & \sigma_{y,k}^2 & 0 \\ 0 & 0 & \sigma_{\theta,k}^2 \end{bmatrix}, \quad (33)$$

For ease of simulation, the exposure time of the camera T is fixed to 1 second. The measured velocity is only used for initialising new targets. The simulation of the time lapse between successive images Δ_k is dependent upon the scenario. It shall be done in one of two ways:

- Static Δ_k : A value of Δ_k is generated at the start of the scenario. This remains constant throughout.
- Random Δ_k : A value of Δ_k is generated at each time step k , drawn from a uniform distribution between some minimum Δ_k^{min} and maximum Δ_k^{max} parameters

The two simulated scenarios share the same parameters with the exception of how the time lapse Δ_k evolves. These parameters are shown below in Table.1

Number of Monte Carlo (MC) Runs	40	Number of MC Particles	50
Number of Time Steps	30	Initial Number of Targets M	10
Rate of Target Birth λ_b	0.2 births / frame	Minimum Target Velocity v_k^{min}	5 pixels/image
Maximum Target Velocity v_k^{max}	15 pixels/image	Survival Probability p_s	0.95
Detection Probability p_d	0.9	False Alarm Rate λ_{fa}	5 / frame
State Space Dimensions ($X \times Y$)	1000 pixels \times 1000 pixels	Minimum Time Lapse Δ_k^{min}	0.1 seconds
Maximum Time Lapse Δ_k^{max}	4 seconds		

Table 1 – Simulation parameters

The accuracy of the time lapse estimation shall be measured using the Root Mean Square Error (RMSE) between the estimate obtained from the particle filter and the ground truth. Also the presented results are the averaged results over the MC runs. The estimates used in the results are obtained using a Maximum A Posteriori (MAP) estimate of the posterior likelihood distribution.

The output from the highest weighted particle’s PHD or SO-PHD filter, at each time step, shall be compared to the simulated ground truths using the Optimal SubPattern Assignment (OSPA) metric [16]. The OSPA metric introduces the concept of a miss distance for multi-target filters, which jointly weights the accuracy of a filter’s spatial estimates and its estimated target cardinality to produce a single value. For the OSPA results shown, a cut-off parameter of $c = 100m$ and order parameter $p = 2$ is used. The estimated number of targets alongside the ground truth shall also be shown.

5 Results

The execution times of the methods can be seen in Table.2, as expected the SO-PHD implementation is more computationally expensive by a factor of approximately 1.5. Fig.1 shows the acrmse of the Δk estimates. These show that for both scenarios the Single Cluster Probability Hypothesis Density (SC-PHD) method performs incredibly well (0.086 peak RMSE), regardless of the multi-object filter used. Note that the Δk estimate is highly dependent on the number of particles as more particles will allow a more accurate estimate to occur.

The OSPA (Fig.2) and estimated cardinality (Fig.3) results show that the estimation of the multi-object state is also accurate, with the filters often being less than 50 for both scenarios. As expected the SO-PHD filter performs better than the PHD filter. Note that the variance estimated by the SO-PHD filter is not shown here.

Method	Random	Static
PHD	0.592	0.612
SO-PHD	0.928	1.095

Table 2 – Execution time results in seconds per time step

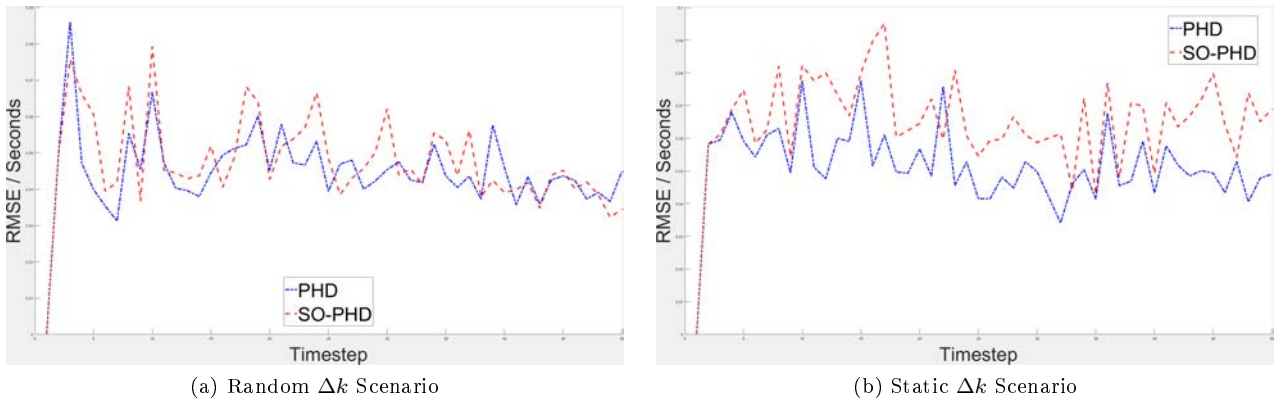


Figure 1 – RMSE results of the Δk estimates

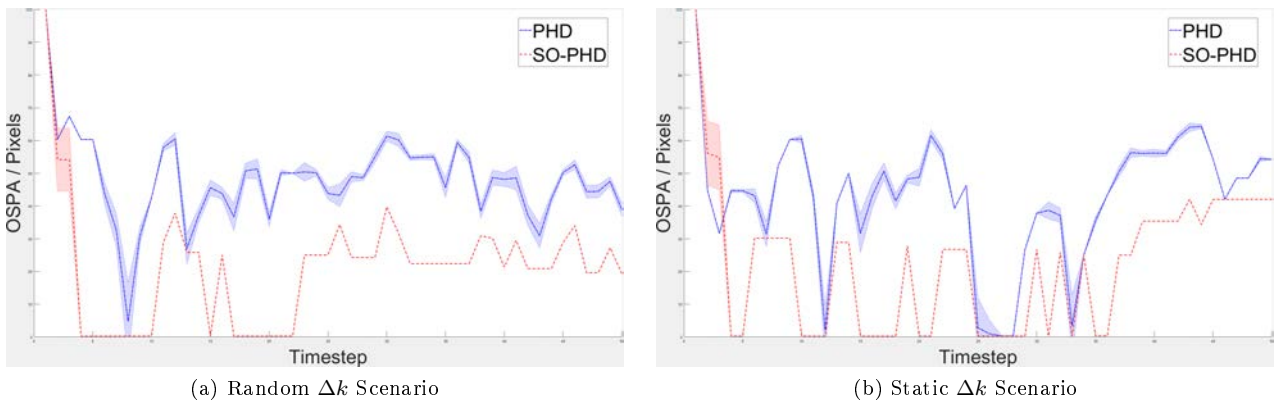


Figure 2 – OSPA results of the Δk estimates

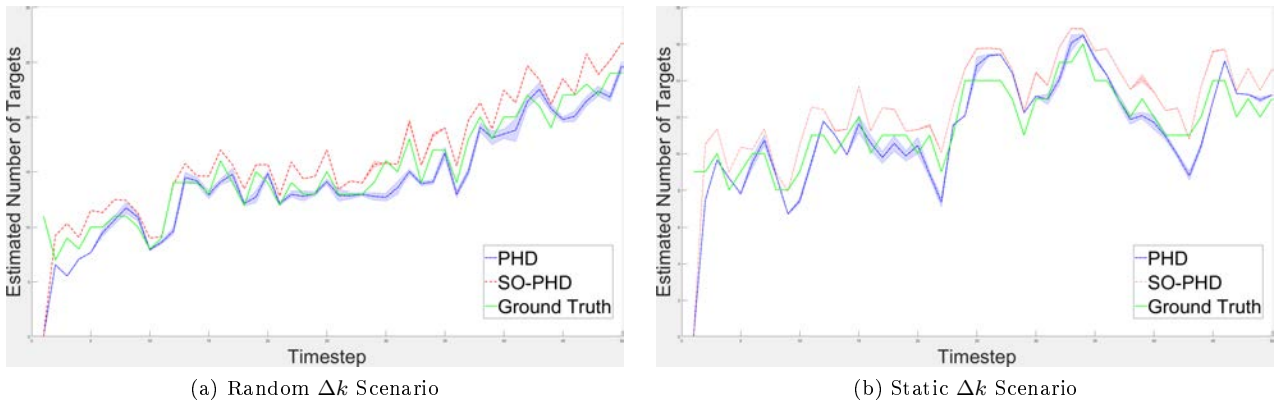


Figure 3 – Cardinality results of the Δk estimates

6 Conclusions

From the results shown above, it is clear that the Single Cluster Probability Hypothesis Density (SC-PHD) filter can accurately estimate the time lapse that occurs between images whilst also simultaneously tracking the dynamic objects in the image sequence. It has also been shown to work on real data although that is outside the scope of this paper. This method is not just limited to this application, it can be extended to almost any sensor calibration problem.

The main limitation of this implementation is that it requires some measurable information about the targets' velocity or that the targets all move at roughly the same speed. A possible extension to this method can be seen in [4], where a joint update step for the PHD filter is used to discriminate better between the different target populations. The method could be extended to the physical spherical plane (right ascension α and declination δ). This would allow a more meaningful physical interpretation of the results to occur and an easier integration with common orbital determination methods.

7 References

- [1] R. P. S. Mahler, "Multitarget Bayes Filtering via First-Order Multitarget Moments," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 39, no. 4, pp. 1152–1178, 2003.
- [2] I. Schlangen, "A Second-Order PHD Filter with Mean and Variance in Target Number," *Accepted to: Transactions on Signal Processing, IEEE Journal of*, 2016. arXiv:1704.02084.
- [3] R. P. S. Mahler, "PHD Filters of Higher Order in Target Number," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 43, no. 4, pp. 1523–1543, 2007.
- [4] C. S. Lee, D. E. Clark, and J. Salvi, "SLAM with Dynamic Targets via Single-Cluster PHD Filtering," *IEEE Journal for Selected Topics in Signal Processing (Special Issue on Multi-Target Tracking)*, 2013.
- [5] B. Ristic, D. E. Clark, and N. Gordon, "Calibration of Multi-Target Tracking Algorithms Using Non-Cooperative Targets," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 3, pp. 390–398, 2013.
- [6] I. Schlangen, J. Franco, J. Houssineau, W. T. E. Pitkeathly, D. E. Clark, I. Smal, and C. Rickman, "Marker-Less Stage Drift Correction in Super-Resolution Microscopy Using the Single-Cluster PHD Filter," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 10, pp. 193–202, Feb. 2016.
- [7] M. Campbell, I. Schlangen, E. Delande, and D. Clark, "Image Registration Using Single Cluster PHD Methods," in *Advanced Maui Optical and Space Surveillance Technologies Conference*, 2017.
- [8] I. Schlangen, *Multi-object filtering with second-order moment statistics*. PhD thesis, Heriot-Watt University, 2017. Ph.D. thesis.
- [9] B.-T. Vo, B.-N. Vo, and A. Cantoni, "Analytic Implementations of the Cardinalized Probability Hypothesis Density Filter," *Signal Processing, IEEE Transactions on*, vol. 55, no. 7, pp. 3553–3567, 2007.
- [10] I. Schlangen, D. E. Clark, and E. D. Delande, "Single-cluster PHD filter Methods For Joint Multi-Object Filtering and Parameter Estimation," May 2017. arXiv:1705.05312.
- [11] C. S. Lee, D. E. Clark, and J. Salvi, "SLAM with Single Cluster PHD Filters," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*, pp. 2096–2101, May 2012.
- [12] A. Swain, *Group and Extended Target Tracking with the Probability Hypothesis Density Filter*. PhD thesis, Heriot-Watt University, 2013.
- [13] B.-N. Vo and W.-K. Ma, "The Gaussian Mixture Probability Hypothesis Density Filter," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4091–4104, 2006.
- [14] P. A. Kountouriotis and S. Maskell, "Maneuvering Target Tracking Using an Unbiased Nearly Constant Heading Model," in *2012 15th International Conference on Information Fusion*, pp. 2249–2255, July 2012.
- [15] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*. Prentice-Hall, 1979.
- [16] D. Schuhmacher, B.-N. Vo, and B.-T. Vo, "A Consistent Metric for Performance Evaluation of Multi-object filters," *Signal Processing, IEEE Transactions on*, vol. 56, no. 8, pp. 3447–3457, 2008.

Géolocalisation et navigation en milieu difficile / Geolocation and navigation in difficult environment

Super-resolution Sparse Channel Estimation for Localization, Prospects and Issues

.....

Bernard H. Fleury¹

¹ *Department of Electronic Systems, Aalborg University,*
fleury@es.aau.dk

Mots clés (en français et en anglais) : Apprentissage Bayésien de Signaux Epars, Localisation dans les Réseaux de Senseurs Sans Fil, Sparse Bayesian Learning, Localization in Wireless Sensor Networks

Présentation en session plénière / Plenary session communication

Résumé / Abstract

In this talk we consider a class of sparse Bayesian algorithms for the estimation of wireless channels in the context of their application to localization. In a nutshell, these algorithms aim at detecting and estimating dominant “specular-like multipath components” in the channel response. Specifically, the number of said components and their parameters, such as their relative delay and complex amplitude, are estimated. This information can be exploited for localization purpose, e.g. by finger-printing or by reconstructing the corresponding physical propagation paths between transmitter and receiver, as done in SLAM.

We discuss the key properties of these algorithms, such as their ability to detect components and to resolve them in the dispersion domain (e.g. with respect to their relative delay). We also shed some light on the correct interpretation of “components” extracted by such algorithms (and actually by any parametric algorithm). We discuss the implications of these properties on localization schemes based on multipath reconstruction.

Super-resolution Sparse Channel Estimation for Localization -
Prospects and Issues

Bernard H. Fleury
Aalborg University (AAU), Denmark

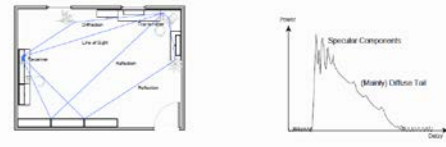
Contributors: Thomas L. Hansen (AAU), Erik Leitinger (TU Graz), Stefan
Greblen (TU Graz)

Joint Work with TU Graz

29 March 2018
JS18 URSI France

Context / Motivation I

Multipath propagation:

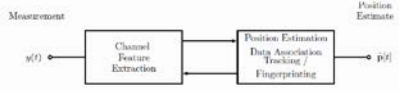


Localization exploiting multipath propagation:

- Multipath-propagation-based finger printing
- Multipath-propagation-based simultaneous localization and mapping (SLAM)

Context / Motivation II

Generic architecture of a (two-stage) localization algorithm:



Relevant issues regarding the channel estimator:

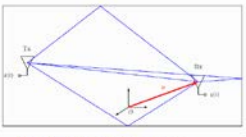
- Detection of artefacts,
- Miss of components
- Superresolution capability
- Physical reality of estimated multipath components

Table of Contents

- Context / Motivation
- Signal Model - Time Dispersion
- Signal Model - Time Dispersion
- Line Spectral Estimation using Sparse Bayesian Learning
- Experimental Results
- Conclusion
- Literature

Signal Model - Time Dispersion

The channel as a linear time-invariant system:



Frequency domain characterization:

- $X(f)$: transmitted signal with bandwidth B
- $Y(f) \triangleq Y(f, \mathbf{p})$: received signal
- $W(f)$: White Gaussian noise
- $H(f) \triangleq H(f, \mathbf{p})$: Channel (frequency) transfer function

Input-output relationship:

$$Y(f) = H(f)X(f) + W(f)$$

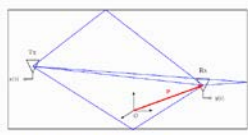
Traditional assumption:

$$H(f) = \sum_{\tilde{L}} \tilde{\alpha}_{\tilde{L}} \exp[-j2\pi f \tilde{\tau}_{\tilde{L}}]$$

\tilde{L} small: specular channel; \tilde{L} large: dense (diffuse) channel

Signal Model - Time Dispersion

The channel as a linear time-invariant system:



Channel sounding (OFDM):

- $x(t)$: transmitted signal
- $y(t) \triangleq y(t, \mathbf{p})$: received signal
- $w(t)$: white Gaussian noise
- $H(f) \triangleq H(f, \mathbf{p})$: channel (frequency) transfer function

Traditional assumption:

$$H(f) = \sum_{\tilde{L}} \tilde{\alpha}_{\tilde{L}} \exp[-j2\pi f \tilde{\tau}_{\tilde{L}}]$$

\tilde{L} small: specular channel; \tilde{L} large: dense (diffuse) channel

Discrete-time Model

$$\mathbf{y} = \mathbf{h} + \mathbf{w}$$

where

- $\mathbf{y} \in \mathbb{C}^N$: received symbol
- $\mathbf{w} \in \mathbb{C}^N$: channel noise
- We assume an all-one OFDM symbol $[1, \dots, 1]^T$.

and

$$\mathbf{h} \triangleq [h_1, \dots, h_N]^T \triangleq [H(n\Delta f) : n = 0, \dots, N-1]^T$$

with

- Δf : carrier spacing
- N : number of carrier

Under the traditional assumption:

$$\mathbf{h} = \Psi(\tilde{\theta}) \tilde{\alpha}$$

$$\Psi(\tilde{\theta}) \triangleq [\psi(\tilde{\theta}_{\tilde{L}}) : \tilde{L} = 1, \dots, \tilde{L}]$$

$$\psi(\tilde{\theta}) \triangleq [\exp[-j2\pi n \theta] : n = 0, \dots, N-1]^T$$

Frequency variable:
 $\theta \in [-1/2, +1/2]$
 $\tilde{\theta}_{\tilde{L}} \triangleq \Delta f \tilde{\tau}_{\tilde{L}}$

First- and Second-order Characterization

Expectation:

$$\mathbb{E}[\mathbf{h}] = \mathbb{E}[\mathbf{w}] = \mathbb{E}[\mathbf{y}] = 0$$

Covariance matrices:

$$\Sigma_{hh} \triangleq \mathbb{E}[\mathbf{h}\mathbf{h}^H] \quad \text{Channel covariance matrix}$$

$$\Sigma_{ww} = \beta^{-1} \mathbf{I}$$

$$\Sigma_{yy} = \Sigma_{hh} + \beta^{-1} \mathbf{I}$$

Random vectors are assumed circularly symmetric.

Rank of the channel covariance matrix:

$$L \triangleq \text{rank}(\Sigma_{hh}) \leq N$$

The rank is system-dependent: it depends on N for fixed Δf .

Carathéodory Representation

If Σ_{hh} is Toeplitz, it can be decomposed as

$$\Sigma_{hh} = \sum_{\ell=1}^L \gamma_{\ell} \psi(\theta_{\ell}) \psi(\theta_{\ell})^H$$

$$= \Psi(\theta) \Gamma \Psi(\theta)^H$$

If $L < N$, the representation is unique.

It follows that

$$h = \Psi(\theta) \alpha \quad \alpha \triangleq [\alpha_{\ell} : \ell = 1, \dots, L]^T \in \mathbb{C}^L$$

Under the reasonable assumption that h is wide-sense stationary, Σ_{hh} is Toeplitz.

We assume that $L < N$, see numerical results.

Line Spectral Estimation

Generic model:

$$y = \Psi(\theta) \alpha + w$$

$$= \sum_{\ell=1}^L \alpha_{\ell} \psi(\theta_{\ell}) + w$$

Inference problem:

Estimate L , $(\alpha_{\ell}, \theta_{\ell}) : \ell = 1, \dots, L$, and noise precision $\beta!$

The fact that L is unknown makes the problem more difficult ... and scientifically challenging.

Back to the traditional assumption:
Traditional assumption leads to the same representation of h :

$$h = \Psi(\theta) \alpha = \Psi(\tilde{\theta}) \tilde{\alpha}$$

While the components in $\Psi(\tilde{\theta}) \tilde{\alpha}$ are intended to have a physical meaning, those in $\Psi(\theta) \alpha$ do not. The latter are virtual components. They might coincide with a physical component, but not always.

A sparse channel estimator exploits the low-rank structure of Σ_{hh} . Thus it estimates the virtual components.

Signal Model for Inference

We fix the number of components:

$$y = \sum_{m=1}^M \alpha_m \psi(\theta_m) + \alpha \quad M \geq N > L$$

$$= \Psi(\theta) \alpha + \alpha$$

$$\Psi(\theta) = [\psi(\theta_1), \dots, \psi(\theta_M)] \quad \text{Dictionary matrix (frequency dependent)}$$

Rationale:
We use a sparse estimator that will set the estimates of the weight of sumnumerous components to zero.

Our Choice: Sparse Bayesian Learning
We use a Bayesian framework.

We use a sparsity inducing hierarchical prior for α , specifically a Gamma-Gaussian prior.

Sparse Bayesian Learning with Parametric Dictionary Estimation

Joint probability density function (pdf):

$$p(y, \alpha, \gamma, \theta, \beta) = \underbrace{p(y|\alpha, \theta, \beta)}_{\text{Hierarchical prior model}} \prod_{m=1}^M p(\alpha_m|\gamma_m) \prod_{m'=1}^M p(\gamma_{m'}) \underbrace{p(\beta)}_{\text{Frequency Estimation (estimation of dictionary parameters)}} \prod_{m''=1}^M p(\theta_{m''})$$

Factor graph:

Probabilistic Model in the Sparse Bayesian Framework

Joint pdf:

$$p(y, \alpha, \gamma, \theta, \beta) = p(y|\alpha, \theta, \beta) \prod_{m=1}^M p(\alpha_m|\gamma_m) \prod_{m'=1}^M p(\gamma_{m'}) \prod_{m''=1}^M p(\theta_{m''}) p(\beta)$$

where

$$p(y|\alpha, \theta, \beta) = p_{CN}(y; \Psi(\theta)\alpha, \beta^{-1}I)$$

i.e. n is a white Gaussian noise vector

$$p(\alpha_m|\gamma_m) = p_{CN}(\alpha_m; 0, \gamma_m^{-1}) \quad m = 1, \dots, M$$

$$p(\gamma_m) = p_G(\gamma; c, d) \quad m = 1, \dots, M$$

$$p(\theta_m) = p_{VM}(\theta_m; \mu, \kappa) \quad m = 1, \dots, M$$

$$p(\beta) = p_G(\beta; a, b)$$

The distribution we will consider subsequently, specifically their pdfs:

- $p_{CN}(x; m, V) = \frac{1}{\pi^N \det(V)} \exp\{-\frac{1}{2}(x-m)^H V^{-1}(x-m)\}$ complex Gaussian
- $p_{VM}(\theta; \mu, \kappa) = \frac{\kappa}{2\pi b(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}$, $\theta \in [-\pi, +\pi]$ von Mises
- $p_G(x; u, v) = \frac{v^x}{\Gamma(x)} x^{x-1} \exp\{-v/x\}$, $x > 0$ Gamma

Mean-Field Approximation

Posterior pdf:

$$p(\alpha, \gamma, \theta, \beta|y) \propto p(y|\alpha, \theta, \beta) \prod_{m=1}^M p(\alpha_m|\gamma_m) \prod_{m'=1}^M p(\gamma_{m'}) \prod_{m''=1}^M p(\theta_{m''}) p(\beta)$$

Family of proxy pdfs:
We consider a family \mathcal{Q} of pdfs with the "simpler" factorization

$$q(\alpha, \gamma, \theta, \beta) = q(\alpha) \prod_{m=1}^M q(\gamma_m) \prod_{m'=1}^M q(\theta_{m'}) q(\beta)$$

We select as an approximation of $p(\alpha, \gamma, \theta, \beta|y)$ the one element in \mathcal{Q} that is the closest:

$$q^*(z) = \arg \min_{q(z) \in \mathcal{Q}} \text{KL}(q(z)||p(z|y)) \quad z \triangleq (\alpha, \gamma, \theta, \beta)$$

Computation of $q^*(z)$:
Different iterative approaches exist to compute (approximate) the solution to the arg min problem (e.g. variational EM).

Mean-Field Approximation - The Four Estimation Tasks

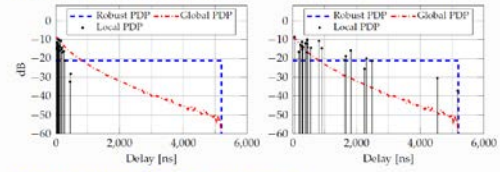
Facts about and Properties of SBL Estimators

- Various hierarchical models for α can be used that promote sparsity: Gamma-Gaussian, Bernoulli-Gaussian or Bernoulli-Gamma-Gaussian.
- SBL are Type II Bayesian estimators, while conventional methods, such as basis pursuit denoising (LASSO), atomic norm denoising, can be interpreted as Type I Bayesian estimator.
- Type II Bayesian estimators promote greater sparsity than Type I Bayesian estimators.
- SBL can also be interpreted as stochastic maximum-likelihood with model order estimation.
- SBL inherently integrates the order estimation process. No hypothesis testing problem needs to be solved, like in classical model order estimation based on information theoretic criteria.

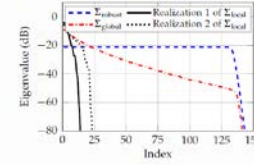
Numerical Studies - Rank of Radio Channels

A: Specular Channel - ITU-R M.2135 UMa NLOS Channel Model

Two generated impulse responses:



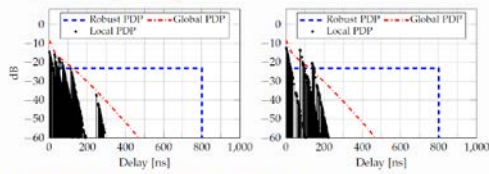
Covariance eigenvalues of the two responses:



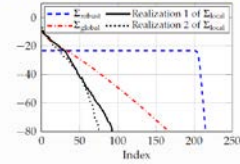
- Bandwidth: 25.6 MHz
- Sampling period: 25 KHz
- Dimension of y : 1024

B: Dense Channel - IEEE 802.15.a Outdoor NLOS Channel Model

Two generated impulse responses:

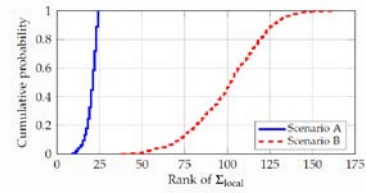


Covariance eigenvalues of the two responses:



- Bandwidth: 256 MHz
- Sampling period: 250 KHz
- Dimension of y : 1024

Effective Rank of the Synthetic Channels



Experimental Results

Description of the Measurement Experiment

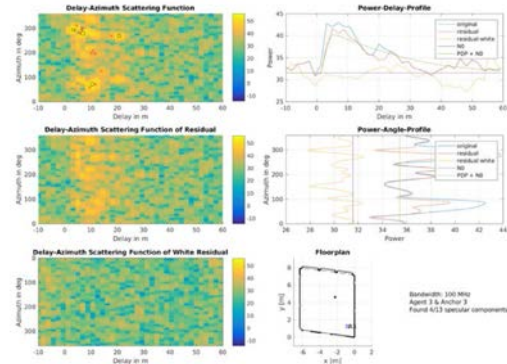
Investigated indoor environment:



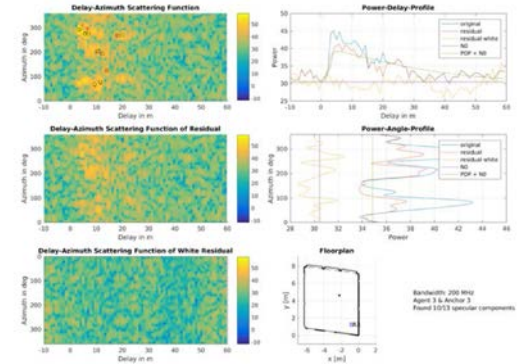
Equipment:

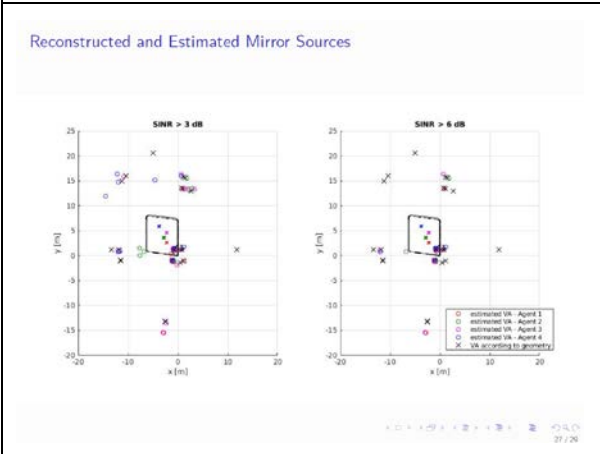
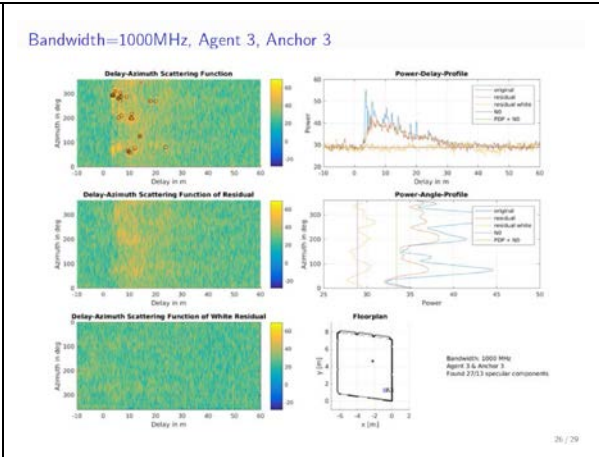
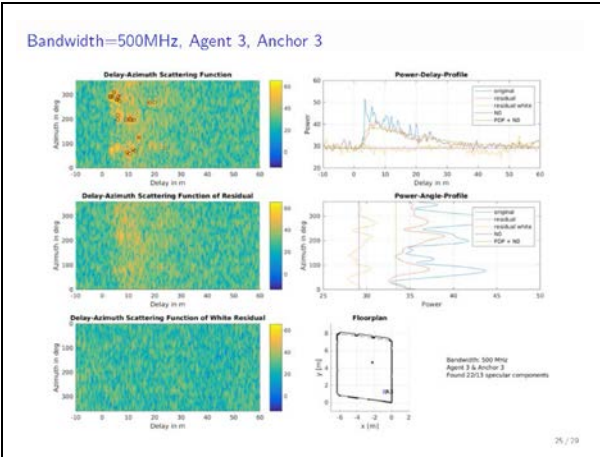
- Center frequency: 7 GHz
- Agents equipped with a 5x5 (virtual) array; array spacing: 2 cm
- Anchors equipped with a single antenna
- Bandwidth is varied: 100, 200, 500, 1000 MHz

Bandwidth=100MHz, Agent 3, Anchor 3



Bandwidth=200MHz, Agent 3, Anchor 3





Conclusion and Outlook

Conclusion:

- The inherent threshold of SBL needs to be increased in order to decrease the number of artefacts (VALSE is an alternative).
- Superresolution demonstrated in synthetic channels (half the Nyquist period)
- SBL shows a sensible behaviour in real conditions.
- Caution is needed in the interpretation of estimated components as physical ones.
- Behaviour is strongly dependent on the selected iterative implementation.
- SBL works for both specular-like and diffuse channels

Outlook:

- Implementation in a localization estimation and tracking system

28 / 29

References

[BHF17] M. A. Badiu, T. Hansen, and B. Fleury, Variational bayesian inference of line spectra, *IEEE Transactions on Signal Processing PP* (2017), no. 99, 1–1.

[HFR18] T. L. Hansen, B. H. Fleury, and B. D. Rao, Superfast line spectral estimation, *IEEE Transactions on Signal Processing PP* (2018), no. 99, 1–1.

29 / 29



Uncertainty Estimation in AoA-based Localization using PCE

Estimation de l'incertitude en localisation par angles d'arrivée avec le développement en chaos polynomial

Thomas Van der Vorst^{1,2}, Mathieu Van Eeckhaute¹, Aziz Benlarbi-Delai², Julien Sarrazin², François Quitin¹, François Horlin¹, and Philippe De Doncker¹

¹ *Université Libre de Bruxelles (ULB), Av. Roosevelt 50, 1050 Brussels, Belgium*

² *Sorbonne Université, Laboratoire d'Électronique et Électromagnétisme, UR2, F-75005 Paris, France*

Keywords: Polynomial chaos, Localization, Angle-of-Arrival

Mots-clés: chaos polynomial, localisation, angle d'arrivée

Abstract:

In this paper, polynomial chaos expansions (PCEs) are applied to angle-of-arrival-based localization. By applying a polynomial chaos expansion on the least squares estimator, a new positioning method is designed. From this expansion, three methods are discussed to obtain confidence regions of the position. Simulation results show that the proposed methods provide a substantial gain in calculation time compared to a Monte-Carlo simulation, while giving accurate approximations of the exact confidence region.

Résumé:

Dans ce travail, les développements en chaos polynomial sont appliqués à la localisation par angles d'arrivée. Une nouvelle méthode de localisation est proposée, utilisant le développement en chaos polynomial d'un estimateur par moindres carrés de la position. Trois méthodes permettant d'obtenir une zone de confiance de la position à partir de ce développement sont présentées. Les résultats de nos simulations montrent que ces méthodes apportent un gain substantiel en temps de calcul par rapport aux simulations de type Monte-Carlo, tout en donnant des approximations précises de la zone de confiance exacte.

1 Introduction

Accurate localization is one of the fundamental requirements of future Internet-of-Things (IoT) networks. In a cellular infrastructure, high localization accuracy allows network providers to offer additional services related to contextualized information delivery, targeted advertising or security applications. Traditional localization approaches based on the estimation of the signal ToA/TDoA rely on wide communication bandwidths, and can thus not be applied to the localization of IoT nodes (which typically operate with low data rates and low bandwidths). By contrast, angle-of-arrival (AoA) based estimation is less dependent on the bandwidth of the communication system, making it a suitable candidate for IoT localization.

Recently, the deployment of fixed reference nodes, referred to as anchors, which can communicate with the user equipment (UE) to be localized, has been considered [1]. In this work we propose a localization system for a network of densely deployed anchors using AoA measurements at the different anchors. Localization algorithms using AoA measurements have been previously investigated in literature. In [2] the authors use a least squares (LS) estimator, and [3] investigates the use of a linearized LS estimator. While these methods show good efficiency, they do not take into account from the outset the uncertainty of each estimated AoA. We propose to apply polynomial chaos expansion (PCE) theory to the localization of a RF transmitter in order to exploit the AoA measurements at the different anchors, associated with their known uncertainties. This allows us to obtain the location of the transmitter, as well as its statistical distribution, and subsequently, to draw confidence regions. Polynomial chaos expansions allows one to determine the statistical properties of the output of a process, based on the probability density function (PDF) of the input random variables of the process [4]. While the idea to use PCE in AoA-based localisation was initially presented in [5], this contribution presents two original ways of approximating the confidence region using the statistical moments of the position estimate obtained by the PCE. Both these methods are based on the assumption that the position is Gaussian distributed. The first one assumes the two coordinates to be uncorrelated, while the second one introduces the covariance of the two coordinates to take their correlation into account.

2 Angle-of-Arrival based localization

We consider the situation where N anchors of known position $\mathbf{x}_i = (x_i, y_i)$ collect the angle-of-arrival measurements θ_i obtained from the signal emitted by one UE of unknown position $\mathbf{x} = (x, y)$. The measurement of these AoA can be achieved with arrays of antennas by applying the MUSIC algorithm [6]. When more than $N = 2$

anchors are present, this two-dimensional problem is overdetermined. In the case where the AoA measurements are not subject to errors, this deterministic problem can be expressed in the following form, as shown in [2]:

$$\begin{bmatrix} -x_1 \sin \theta_1 + y_1 \cos \theta_1 \\ \vdots \\ -x_N \sin \theta_N + y_N \cos \theta_N \end{bmatrix} = \begin{bmatrix} -\sin \theta_1 & \cos \theta_1 \\ \vdots & \vdots \\ -\sin \theta_N & \cos \theta_N \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \quad (1)$$

that can be rewritten as

$$\mathbf{b} = \mathbf{H}\mathbf{x} \quad (2)$$

In a realistic scenario, the measurements will be prone to errors. In that case, the system of equations (1) will not have any solution. To overcome this issue, the least squares estimate of the position $\hat{\mathbf{x}}$ is calculated as:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{b} \quad (3)$$

$$\equiv \mathbf{H}^\dagger \mathbf{b} \quad (4)$$

where \dagger is the pseudo-inverse operator.

3 Angle-of-Arrival based localization using Polynomial Chaos Expansions

Consider that each measurement θ_i is associated with an uncertainty that is also estimated by the anchors. Consequently, the angles-of-arrival can be treated as random variables of known statistical distribution. The least squares estimate of the position of the UE is a function of the random variables θ_i , and is therefore a random variable. We denote this random variable by $\hat{\mathbf{x}}$.

3.1 Polynomial Chaos Expansions

The theory of Polynomial Chaos Expansions (PCE) was initially introduced by Norbert Wiener in 1938 with *The Homogeneous Chaos* [7]. The PCE is a surrogate modelling technique, through which a response surface of a computational model is constructed using multidimensional polynomials that are orthogonal with respect to the statistical distributions of the input random variables of this model. As will be explained in the next section, PCE are based on standard random variables, e.g. Gaussian or Uniform, to generate the polynomial basis of the expansion. However, it has been demonstrated that the use of isoprobabilistic transforms makes it possible to apply the theory for any arbitrary random variables [8, 9]. This is usually referred to as *Generalized Polynomial Chaos Expansions*. Recently, the PCE theory has been applied in electromagnetic engineering, in the field of antennas [10, 11, 12], and in propagation theory [13].

3.1.1 Definition of the polynomial basis

Consider a computational model \mathcal{M} , with a vector of N input parameters $\boldsymbol{\theta}$, whose output, the response of the model, is \mathbf{x} :

$$\mathbf{x} = \mathcal{M}(\boldsymbol{\theta}) \quad (5)$$

The principle of the PCE is to construct a surrogate model $\hat{\mathcal{M}}$, which is a sum of polynomials on the input variables $\boldsymbol{\theta}$, of the model \mathcal{M} , that converges to the model response as the number of polynomials in the sum grows:

$$\hat{\mathbf{x}} = \sum_{\alpha \in \mathbb{N}^N} \mathbf{c}_\alpha \Psi_\alpha(\{\theta_i\}_{i=1}^N) \quad (6)$$

$$= \hat{\mathcal{M}}(\boldsymbol{\theta}) \quad (7)$$

In this equation, the coefficients of the expansion are written \mathbf{c}_α . The polynomials $\{\Psi_\alpha(\{\theta_i\}_{i=1}^N)\}_{\alpha \in \mathbb{N}^N}$ form a polynomial chaos basis of the adequate Hilbert space containing $\hat{\mathbf{x}}$. These multivariate polynomials are products of univariate polynomials. For each input random variable θ_i , a series of univariate polynomials $\psi_k^{(i)}$, $k \in \mathbb{N}$, are constructed so that they are orthogonal with respect to the scalar product defined by the PDF of θ_i , φ_{θ_i} :

$$\langle \psi_j^{(i)}, \psi_k^{(i)} \rangle = \int \psi_j^{(i)}(u) \psi_k^{(i)}(u) \varphi_{\theta_i}(u) du = \gamma_j^{(i)} \delta_{jk} \quad (8)$$

Type of variable	Distribution	Polynomials	Hilbertian Basis $\psi_k(x)$
Uniform	$\mathbf{1}_{]-1,1[}(x)/2$	Legendre $P_k(x)$	$P_k(x)/\sqrt{\frac{1}{2k+1}}$
Gaussian	$\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$	Hermite $H_{e_k}(x)$	$H_{e_k}(x)/\sqrt{k!}$
Gamma	$x^a e^{-x} \mathbf{1}_{\mathbb{R}^+}(x)$	Laguerre $L_k^a(x)$	$L_k^a(x)/\sqrt{\frac{\Gamma(k+a+1)}{k!}}$
Beta	$\mathbf{1}_{]-1,1[}(x) \frac{(1-x)^a (1+x)^b}{B(a)B(b)}$	Jacobi $J_k^{a,b}(x)$ $\mathfrak{J}_{a,b,k}^2 = \frac{2^{a+b+1}}{2k+a+b+1} \frac{\Gamma(k+a+1)\Gamma(k+b+1)}{\Gamma(k+a+b+1)\Gamma(k+1)}$	$J_k^{a,b}(x)/\mathfrak{J}_{a,b,k}$

Table 1 – Standard univariate polynomial families used in PCE [14]

For standard statistical distributions, the associated polynomial families are analytically known. A few examples are given in Table 1. If no standard polynomials are defined for the input distribution, it is possible to define an isoprobabilistic transform to reduce the input variables to components that are distributed according to one of the standard distributions in Table 1 [14].

The multivariate polynomials $\Psi_{\alpha}(\mathbf{x})$ are obtained by the tensor product of their univariate counterparts [4]:

$$\Psi_{\alpha}(\boldsymbol{\theta}) \equiv \prod_{i=1}^N \psi_{\alpha_i}^{(i)}(\theta_i) \quad (9)$$

3.1.2 Truncation Schemes

The interest of a surrogate model being to reduce the computation time to obtain statistical information on the model response, the expression (6) is not applicable in practice and it needs to be truncated in order to limit the number of coefficients to compute. There are several ways to select the subspace of \mathbb{N}^N to which the multi-index α belongs to. However, we limit ourselves here to the simplest scheme. The standard truncation scheme corresponds to all polynomials in the N input variables of total degree less than or equal to p [14]:

$$\mathcal{A}^{N,p} = \{\alpha \in \mathbb{N}^N : |\alpha| \leq p\} \quad (10)$$

3.1.3 Calculation of the coefficients

Several methods exist to compute the coefficients of the expansion \mathbf{c}_{α} . The projection method is considered here. This method is a non-intrusive one as it is based on the post-processing of a set of model evaluations, the experimental design. The latter is chosen by an adequate sampling of the input random variables. The orthogonality of the multivariate polynomials, that follows from Eqs. (8) and (9), associated with the definition of the PCE in Eq. (6), directly leads to the following expression for the coefficients:

$$\mathbf{c}_{\alpha} = \frac{\mathbb{E}[\hat{\mathbf{x}}\Psi_{\alpha}]}{\mathbb{E}[\Psi_{\alpha}^2]} \quad (11)$$

In the latter equation, the denominator is known analytically for the standard polynomials in Table 1, whereas the numerator is calculated by an integral:

$$\mathbb{E}[\hat{\mathbf{x}}\Psi_{\alpha}] = \int_{D_{\boldsymbol{\theta}}} \mathcal{M}(\boldsymbol{\theta})\Psi_{\alpha}(\boldsymbol{\theta})\varphi_{\boldsymbol{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta} \quad (12)$$

To numerically evaluate this integral, the usual method is to use the Gauss quadrature method, by which the integral is approximated by a weighted sum:

$$\mathbb{E}[\hat{\mathbf{x}}\Psi_{\alpha}] \approx \sum_{i=1}^K w^{(i)} \mathcal{M}(\boldsymbol{\theta}^{(i)})\Psi_{\alpha}(\boldsymbol{\theta}^{(i)}) \quad (13)$$

The set of weights $w^{(i)}$, and quadrature points $\boldsymbol{\theta}^{(i)}$ (the *experimental design*) correspond to the roots of the used polynomial (see [15] for details).

3.1.4 Post-Processing

The interest of the Polynomial Chaos Expansion theory is to obtain statistical informations on the model response with less computational effort than by a Monte-Carlo simulation of the actual model. It can easily be demonstrated that the mean and the variance of the model response are respectively given by:

$$\mu_{\hat{\mathbf{x}}} = \mathbb{E}[\hat{\mathbf{x}}] = \mathbf{c}_0 \quad (14)$$

$$\sigma_{\hat{\mathbf{x}}}^2 = \text{Var} \left[\sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha} \Psi_{\alpha} \right] = \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha}^2 \|\Psi_{\alpha}\|^2 \quad (15)$$

However, in some applications, it might be necessary to have more statistical informations on the model response. The PCE of the model response may be used to derive an approximation of the probability density function (PDF) of the model response.

3.2 Uncertainty on the position using PCE

3.2.1 First method: Monte-Carlo on the PCE

To obtain an approximation of the PDF of the position of the UE to locate, using the PCE (6), the simplest way is to perform a Monte-Carlo simulation on the PCE itself. For this purpose, a large set of input vectors $\{\boldsymbol{\theta}_k, k = 1, \dots, N_{MC}\}$ is generated according to the statistical distribution of the input variables $\varphi_{\boldsymbol{\theta}}$. Then, the position corresponding to each of these AoA samples is obtained via the least squares estimator. The resulting large set of positions allows one to draw the two-dimensional statistical distribution of the position. From this PDF, the confidence region is simply obtained by selecting all points corresponding to a value of the PDF greater than a certain threshold that leads to the desired probability. This method gives very accurate results but requires a considerable computational effort, since a Monte-Carlo calculation is used. The gain in computational time of this first method compared to a simple Monte-Carlo calculation on the least squares position estimator is not significant. However, for more complex estimators, the use of PCE can significantly reduce the computational time needed to obtain similar results.

3.2.2 Second method: Ellipse

In order to exploit the potential of the PCE in terms of gain in computational time, one can make use of the mean and variances of the position estimate obtained with Eqs. (14) and (15). These values are the first and second order statistical moments of the distribution of the position coordinates. By assuming that the position coordinates $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are uncorrelated, and that both distributions are Gaussian, one can obtain an elliptic approximation of the confidence region. As the coordinates are assumed uncorrelated, the principal axes of the resulting ellipse are parallel to the x -axis and y -axis, respectively. The computational cost of this method is greatly reduced compared to the previous one, since no Monte-Carlo calculation is required. Indeed, once the PCE is calculated, the mean and variance of each coordinate is immediately obtained.

3.2.3 Third method: Oriented Ellipse

As seen previously, the second method considers that the position coordinates are uncorrelated and results in an ellipse whose principal axes are parallel to the x and y -axes. Therefore, the quality of this approximation depends on the choice of the axes in the definition of the problem. To cope with this issue, we propose to calculate the covariance between $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ and to use it for drawing the oriented ellipse. This oriented ellipse is obtained through the diagonalisation of the covariance matrix, and is therefore independent of the reference frame. By 6, the estimation of the x - and y -coordinates are expanded on the same polynomial chaos basis, but separately:

$$\hat{x} = \sum_{\alpha \in \mathcal{A}} c_{\alpha} \Psi_{\alpha}(\{\theta_i\}_{i=1}^N) \quad (16)$$

$$\hat{y} = \sum_{\alpha \in \mathcal{A}} d_{\alpha} \Psi_{\alpha}(\{\theta_i\}_{i=1}^N) \quad (17)$$

In those equations, c_{α} and d_{α} are the coefficients of the expansion of the estimates of the x - and y -coordinates, respectively. The covariance $R_{\hat{x}\hat{y}}$ is then derived using the orthogonality of the polynomials of the basis, leading to an expression similar to Eq. (15):

$$R_{\hat{x}\hat{y}} = \text{Cov}(\hat{x}, \hat{y}) \quad (18)$$

$$= \mathbb{E}[(\hat{x} - \mu_{\hat{x}})(\hat{y} - \mu_{\hat{y}})] \quad (19)$$

$$= \mathbb{E} \left[\left(\sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha} \Psi_{\alpha}(\{\theta_i\}_{i=1}^N) \right) \left(\sum_{\beta \in \mathcal{A} \setminus \mathbf{0}} d_{\beta} \Psi_{\beta}(\{\theta_i\}_{i=1}^N) \right) \right] \quad (20)$$

$$= \mathbb{E} \left[\sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} \sum_{\beta \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha} d_{\beta} \Psi_{\alpha}(\{\theta_i\}_{i=1}^N) \Psi_{\beta}(\{\theta_i\}_{i=1}^N) \right] \quad (21)$$

$$= \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} \sum_{\beta \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha} d_{\beta} \mathbb{E} [\Psi_{\alpha}(\{\theta_i\}_{i=1}^N) \Psi_{\beta}(\{\theta_i\}_{i=1}^N)] \quad (22)$$

$$= \sum_{\alpha \in \mathcal{A} \setminus \mathbf{0}} c_{\alpha} d_{\alpha} \mathbb{E} [\Psi_{\alpha}^2(\{\theta_i\}_{i=1}^N)] \quad (23)$$

Then, the covariance matrix Σ is constructed from Eqs. (15) and (23):

$$\Sigma = \begin{bmatrix} \sigma_{\hat{x}}^2 & R_{\hat{x}\hat{y}} \\ R_{\hat{x}\hat{y}} & \sigma_{\hat{y}}^2 \end{bmatrix} \quad (24)$$

The eigenvalues and eigenvectors of this matrix, obtained simply by diagonalising it, allow one to draw the oriented ellipse. The principal axes of this ellipse are defined by the eigenvectors of the covariance matrix, while the length of the minor and major axes are related to the eigenvalues of the covariance matrix.

4 Results

In the calculations, $N = 3$ anchors have been considered, each of them making an independent AoA estimation. Firstly, Gaussian distributed estimation errors were considered. To the best of our knowledge, there is no precise model for uncertainty distribution in AoA estimation in the literature. It is a common assumption to use Gaussian distribution for uncertainties. This assumption is not restrictive since by using isoprobabilistic transforms, the PCE can be applied to any arbitrary input distribution. The expansion of the position coordinates as a function of the angles has been calculated up to the order 4, with a standard truncation scheme. The choice of the order 4 was made simply by observing the confidence regions obtained for orders up to 5. As there was no significant difference between the CR obtained with the order 4 and 5, the order 4 was selected as a compromise between computational time and precision. In the Monte-Carlo calculation, 20000 runs of the least squares position estimator were used, while 125 were needed to obtain each of the two coordinates PCE. In the first method, we used 20000 realizations of the PCE to draw the confidence region. In comparison, the methods 2 and 3 are more efficient since they only require to compute the mean and the variance of both coordinates - as well as the covariance in method 3 - from Eqs (14), (15) and (23). In the first simulation, the AoA at each anchor was defined by a Gaussian random variable, centred on the actual AoA, and of standard deviation equal to 5° . The results of the three methods are given in Fig. 1, and compared to a Monte-Carlo calculation. We observe that the confidence region presents a stretched shape in the direction of the anchors. An isotropic distribution of the anchors around the UE would lead to a more circular confidence region.

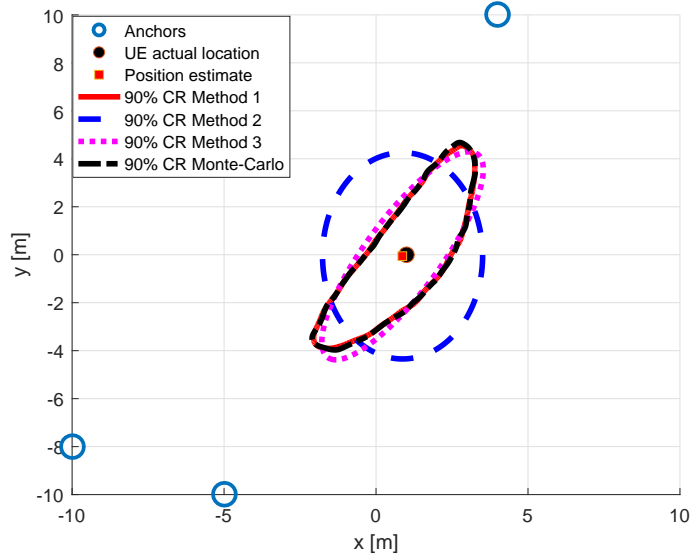


Figure 1 – 90% confidence regions, obtained by the three methods, and assessed by a Monte-Carlo calculation. The three AoA's are defined with a 5° standard deviation.

Then, the standard deviation of the AoA distribution on the anchor located at (4, 10) has been changed to observe its effect on the confidence region. In Fig. 2, this standard deviation was set to 1° , while in Fig. 3 it was set to 15° . In both cases, the other parameters, including the standard deviations of the two other anchors, were identical to the first simulation. We observe in Figs. 2 and 3 that the shape, the size and the orientation of the confidence region is affected by the standard deviation. The oriented ellipse method allows us to almost perfectly cope with these modifications, in contrary to the classical, unoriented, method.

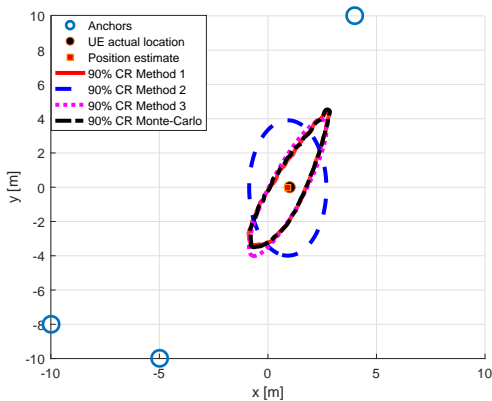


Figure 2 – 90% confidence regions, obtained by the three methods, and assessed by a Monte-Carlo calculation. The standard deviations are 5, 5 and 1° , respectively.

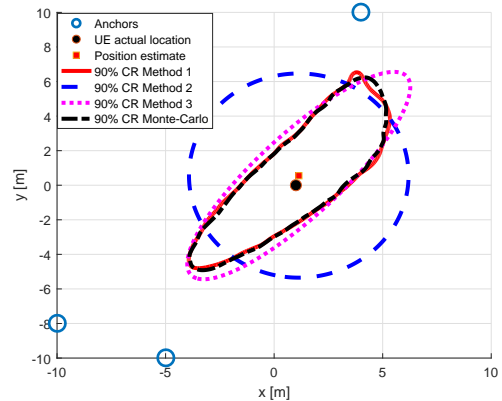


Figure 3 – 90% confidence regions, obtained by the three methods, and assessed by a Monte-Carlo calculation. The standard deviations are 5, 5 and 15° , respectively.

5 Conclusion

A new two dimension position estimation method based on polynomial chaos, least squares estimator, and AoA measurements has been proposed. Compared to traditional positioning methods, it presents the advantage of taking into account the uncertainties on the AoA estimations from the outset. This allows one to compute the covariance matrix of the position estimation, and consequently, confidence regions.

6 Acknowledgment

This work was supported by F.R.S-FNRS, and by Innoviris through the Copine-IoT project.

7 References

- [1] 3GPP TR 37.857, *Study on indoor positioning enhancements for UTRA and LTE*, 2015
- [2] A. Pages-Zamora, J. Vidal, D. Brooks, *Closed-form solution for positioning based on angle of arrival measurements*, The 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2002, pp. 1522-1526 vol.4.
- [3] D. Torrieri, *Statistical Theory of Passive Location Systems*, IEEE Transactions on Aerospace and Electronic Systems, vol. AES-20, no. 2, pp. 183-198, 1984.
- [4] B. Sudret, *Global Sensitivity Analysis Using Polynomial Chaos Expansions*, Reliab. Eng. Sys. Saf., 93, pp. 964-979, 2008.
- [5] T. Van der Vorst, M. Van Eeckhaute, A. Benlarbi-Delaï, J. Sarrazin, F. Quitin, F. Horlin, P. De Doncker, *Angle-of-Arrival based localization using polynomial chaos expansions*, Proc. of the Workshop on Dependable Wireless Communications and Localization for the IoT, 2017
- [6] R. Schmidt, *Multiple Emitter Location and Signal Parameter Estimation*, IEEE Transactions on Antennas and Propagation, 34, pp. 276-280, 1986.
- [7] N. Wiener, *The Homogeneous Chaos*, American Journal of Mathematics, vol. 60, no. 4, pp. 897-936, 1938.
- [8] C. Soize, R. Ghanem, *Physical systems with random uncertainties: chaos representations with arbitrary probability measure*, SIAM Journal on Scientific Computing, vol. 26, no. 2, pp. 395-410, 2004.
- [9] J. Li, D. Xiu, *A generalized polynomial chaos based ensemble Kalman filter with high accuracy*, Journal of computational physics, vol. 228, no. 15, pp. 5454-5469, 2009
- [10] J. Du, C. Roblin, *Statistical modeling of disturbed antennas based on the polynomial chaos expansion*, IEEE Antennas and Wireless Propagation Letters, 2016.
- [11] M. Rossi, A. Dierck, H. Rogier, D. Vande Ginste, *A stochastic framework for the variability analysis of textile antennas*, IEEE Transactions on Antennas and Propagation, vol. 62, no.12, pp. 6510-6514, 2014.
- [12] T. Van der Vorst, M. Van Eeckhaute, A. Benlarbi-Delaï, J. Sarrazin, F. Horlin, P. De Doncker, *Propagation of Uncertainty in the MUSIC Algorithm Using Polynomial Chaos Expansions*, Proc. 11th European Conference on Antennas and Propagation, 2017, pp. 820-822
- [13] A. Haarscher, Ph. De Doncker, D. Lautru, *Uncertainty propagation and sensitivity analysis in ray-tracing simulations*, PIER M, vol. 21, pp. 149-161, 2011.
- [14] S. Marelli, B. Sudret, *UQLab user manual - Polynomial Chaos Expansions*, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, 2015.
- [15] M. Abramowitz, I. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*, Courier Corporation, 1964.



A semi parametric model for RSSI-based localization *Modèle semi paramétrique pour la localisation par RSSI*

Pascal Bianchi², François Portier³, and Kevin Elgui¹

²Télécom ParisTech, pascal.bianchi@telecom-paristech.fr

³Télécom ParisTech, francois.portier@telecom-paristech.fr

¹Télécom ParisTech, kevin.elgui@telecom-paristech.fr, kevin.elgui@sigfox.com

Keywords: statistics, Internet of Things (IoT), semi-parametric, geolocation, RSSI

Mots-clefs: statistiques, Internet des objets, semi paramétrique, géolocalisation, RSSI

Abstract:

Accurate and reliable geolocation of an object in the context of an Internet of Things (IoT) network must deal with a lack of available information. The approach followed in this paper is based on Received Signal Strength Indicators (RSSI) measured at all base stations (BS) as a useful information to infer the object's position. The proposed technique is based on a maximum a posteriori (MAP) estimator derived within a flexible semi-parametric model. The non-parametric components of the model are estimated with the help of Nadaraya-Watson type estimators. Our proposal has been tested and compared with an other method found in the literature: a simple model consisting of taking a weighted barycenter of the BS as the estimator of the real position. Our approach, although computationally less efficient, has revealed to outperform other methods and to come up with some very interesting information on the behaviour of the studied variables.

Résumé:

La problématique de la géolocalisation par le réseau se heurte rapidement au manque de pertinence des informations disponibles. L'approche proposée dans ce papier est basée sur l'estimation de la position d'un objet du réseau à partir des puissances de signal reçues (RSSI) aux antennes de base (BS). Pour cela, nous proposons un estimateur maximum à posteriori (MAP) dans le cadre d'un modèle semi-paramétrique. Les composantes non paramétriques de ce modèle ont été estimés à l'aide d'estimateurs de Nadaraya-Watson. Nous testerons et nous comparerons notre approche à une méthode trouvées dans l'état de l'art, consistant à estimer la position par une barycentre des BS réceptrices.

1 Introduction

In the last few years, the Internet of Things raised a great deal of attention in very diverse fields such as agriculture or health care. Experts agree that 30 billions objects will be part of the IoT by 2020 [1] and 40% of these objects might need to be geolocated. This paper is devoted to the geolocation of connected objects in Sigfox wireless network, called LPWAN (Low Power Wide Area Network). The network has been specifically deployed in order to offer an international connectivity for objects in IoT in more than 40 countries. It provides a low energy, and economic solution for transmissions of messages. Beyond the traditional wireless access to the internet cloud, a targetted application is to use the network to geolocalize the objects. However, the singularity of the Sigfox network makes the task especially challenging.

Indeed, every message transmitted by an object in the Sigfox network occupies an Ultra Narrow Band (100 Hz - 600 Hz). This makes it difficult to address the geolocation by means of *e.g.* channel fingerprinting, because of the absence of frequency diversity of the propagation channel. In addition, the base stations (BS) of the network are not time-synchronized, and geometric approaches such as Time Difference of Arrival ([2], [3]) are not relevant. The approach followed in this paper consists in the use of the set *Received Signal Strength Indicators* (RSSI) measured at all base station (BS) as a useful information to infer the object's position. RSSI based localization is known to be a difficult problem. On the top of that, a quick inspection of the available RSSI observations reveals that the data are very noisy: RSSI measurements are subject to a significant variability. This is due to several factors (urban environment, indoor or outdoor transmission, etc.) The latter variability of the data typically rules traditional parametric approaches based on path-loss model, which aim at relating the RSSI with the source-destination distance. [4], [5]₁₀₂

Our contribution are as follows:



Figure 1 – RSSI received at a particular base station

- We propose to formulate the localization as a Maximum a Posteriori estimation problem of the position given the vector of RSSI values gathered at all BS of the network.
- We use a large amount of data in order to model the likelihood of the RSSI measurements given the position. To that end, we introduce a semiparametric model well suited to the problem at stake.
- We provide detailed numerical experiments showing that the proposed method outperform off-the-shelf solutions.

2 Semiparametric framework

We consider that an object is located at an unknown random position noted Z and belonging to a subset $\mathcal{Z} \subset \mathbb{R}^2$. The network is formed by K base stations numbered from 1 to K . The message leads to K values of RSSI, at each BS, denoted by $\mathbf{R} = (R_1, R_2, \dots, R_K)$. We allow the values of the RSSI to take the value NaN which will encode that the message is actually not received by a BS. Our first assumption deals with the parametric part of the model : the distribution we consider for \mathbf{R} given Z is Gaussian. This Gaussian assumption is illustrated in Fig. 2. We have tested this hypothesis for several z , all of them give similar histograms. Besides, this hypothesis is widely accepted in the literature (e.g. [6]).

Assumption 1 *The conditional distribution of \mathbf{R} given Z is Gaussian with mean $\mathbf{m}(Z)$ and variance Σ .*

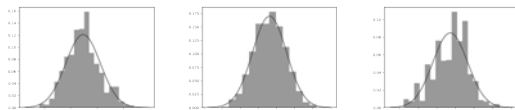


Figure 2 – Histograms of received RSSI for three emitting position.

Our second assumptions is supported by the intuition that the vector $(\mathbf{R}_1, \dots, \mathbf{R}_K)$ should be conditionally independent given the position Z (See Fig. 3).

Assumption 2 *For all $k \neq \ell$, $\Sigma_{k,\ell}(Z) = 0$ a.s.*

Hence, the likelihood function of the observations \mathbf{R} given Z can be decomposed as $p(\mathbf{R}|Z) = \prod_k p(R_k|Z)$. Under assumption 1 we have for all k :

$$p(R_k = r|Z = z) = \mathcal{N}(r; m_k(z), \sigma_k^2),$$

where $\mathcal{N}(\cdot; m, \sigma^2)$ stands for the Gaussian p.d.f. of mean m and variance σ^2 . Of course, m_k needs to be evaluated. To that end, we use the knowledge of a dataset composed of indepent copies $(Z^{(i)}, \mathbf{R}^{(i)})_{i=1 \dots n}$ of

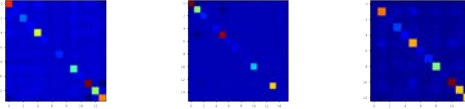


Figure 3 – HeatMap of the covariance matrix of \mathbf{R} for three different value of z .

(Z, \mathbf{R}) . To estimate the non-parametric part of the model, namely the unknown function m , we employ Nadaraya-Watson type estimates as presented for instance in [7] and defined as:

$$\hat{m}_k(z) = \frac{\sum_{i=1}^n R_k^{(i)} K_h(z - Z^{(i)})}{\sum_{i=1}^n K_h(z - Z^{(i)})}.$$

Where $K_h : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ is a kernel function such that $\int K_h = 1$, and $h > 0$ is called the bandwidth of the kernel. In the above equation and due to the lack of space, we overlooked the fact that RSSI may take NaN values, but this can be easily addressed in practice by using a slight modification of our model.

2.1 MAP Estimator of Z

The MAP estimator of Z when \mathbf{R} has been observed is as follows:

$$\begin{aligned} Z^{MAP} &= \arg \max_{z \in \mathcal{Z}} \mathbb{P}(Z \in dz, \mathbf{R} = \mathbf{r}) \\ &= \arg \max_{z \in \mathcal{Z}} \prod_k \mathbb{P}(R_k = r_k | Z \in dz) \end{aligned}$$

Consequently, we define the MAP estimator \hat{Z}^{MAP} replacing the unknown quantities in Z^{MAP} by their estimates, that is:

$$\hat{Z}^{MAP} = \arg \max_{z \in \mathcal{Z}} \prod_k \mathcal{N}(r_k; \hat{m}_k(z), \sigma_k^2)$$

An exhaustive grid search is a possible way to solve this equation.

3 Experiments

3.1 Barycenter model

We propose a very simple model to estimate the location of a device when \mathbf{R} has been observed. The location estimation is as follow:

$$\hat{Z}_{\text{bary}}(\tilde{\mathbf{r}}) = \sum_{k=1}^K \omega_k(r_k) z_{BS_k}.$$

It is simply weighted barycenter of the receiving base stations, and where the weights are increasing functions of the signal power. Despite its simplicity, this model has proven to perform well in certain contexts (e.g. high density of base stations).

3.2 Comparisons

We are giving three metrics to compare the two models. First, the expectation of distance between estimates and real position.

$$\begin{aligned} \mathbb{E}_{test}(d(Z, \hat{Z}_{\text{bary}})) &= 3.5673km \\ \mathbb{E}_{test}(d(Z, \hat{Z}_{\text{MAP}})) &= 2.6141km \end{aligned}$$

One may want to bound the error, thus we give the following information:

$$\begin{aligned} \max_{test}(d(Z, \hat{Z}_{\text{bary}})) &= 8.0573km \\ \max_{test}(d(Z, \hat{Z}_{\text{MAP}})) &= 6.1792km \end{aligned}$$

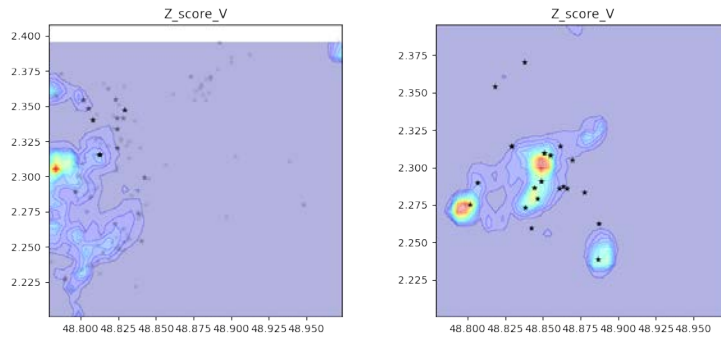


Figure 4 – Colormaps of the contour line of our learned density for two observed R

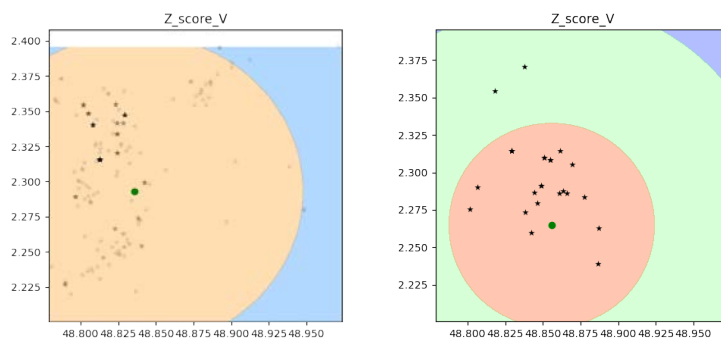


Figure 5 – Colormaps of the contour line of the barycenter approach density for two observed R .

4 Conclusion

In this paper, we propose a model for geolocalisation, with the consideration, that subjected to a disturbed environment, a message can not be received by a base station. Model performance improvements are possible in many places. At first, the relaxation of certain independence hypotheses, and thus search for "patterns" of receiving antennas, characteristic of certain positions. Finally, the reception of a message seems to carry more information about Z than the variable R , this consideration can lead us to search for \hat{Z} with a selection of candidates in several time.

5 References

- [1] C.-L. . al., “An empirical examination of consumer adoption of internet of things services: Network externalities and concern for information privacy perspectives,” in *Computers in Human Behavior*, 2016.
- [2] R. K. . al., “Accuracy analysis for tdoa localization in sensor networks,” 14th International Conference on Information Fusion Chicago, 2011.
- [3] L. Mailaender, “On the geolocation bounds for round-trip time-of-arrival and all non-line-of-sight channels,” EURASIP Journal on Advances in Signal Processing, 2008.
- [4] S. M. . al., “Robust indoor positioning provided by real-time rssi values in unmodified wlan networks,” IEEE Journal of Selected topics in signal processing, 2009.
- [5] S. C. . D. K. Dhaka, “Path loss prediction models for wireless communication channels and its comparative analysis,” in *International Journal of Engineering, Management Sciences*, 2015.
- [6] E. Elnahrawy, X. Li, and R. P. Martin, “The limits of localization using signal strength: a comparative study,” in *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004.*, pp. 406–414, Oct 2004.
- [7] A. Tsybakov, “Apprentissage statistique et estimation non-paramétrique.” Course, 2013.

Le positionnement en milieux contraints : signaux d'opportunité ou infrastructure spécifique ? *Indoor positioning: signals of opportunity or local infrastructure?*

Nel SAMAMA, Alexandre VERVISCH-PICOIS, Thierry TAILLANDIER-LOIZE

Laboratoire SAMOVAR, Télécom SudParis, CNRS, Université Paris-Saclay, Evry, France,
nel.samama@telecom-sudparis.eu

Positionnement, Milieux contraints, Intérieur, Positioning, Indoor.

Résumé/Abstract

Quelles sont les origines des difficultés rencontrées depuis des années dans la recherche d'approches viables à une échelle suffisante pour créer un véritable effet d'entraînement ? Après un état de l'art des techniques et des technologies potentielles, nous montrerons alors, par l'analyse d'une douzaine de paramètres, qu'un nombre réduit de contraintes d'emploi ou de déploiement réduit le choix réel drastiquement.

Afin d'illustrer concrètement notre propos, nous discuterons de trois approches du positionnement. La première repose sur un système de « GNSS Indoor » utilisant la transmission locale de signaux compatibles des récepteurs de navigation par satellites actuels afin de proposer un positionnement de bonne qualité et une continuité du service. La deuxième est à l'opposé en proposant un positionnement non plus continu, mais discrétisé, tant dans l'espace que dans le temps, présentant précision et fiabilité. La troisième est à la base d'un système coopératif, bien plus complexe que les précédents mais qui pourrait être une façon de contourner les limitations mentionnées ci-dessus.

La discussion conclusive synthétise les difficultés à produire une solution industrielle, mais ces dernières reflètent selon nous un intérêt réel limité de la part des parties prenantes qui sont perdues dans le foisonnement d'approches. Nous suggérons alors la création de nouveaux indicateurs de haut niveau, plus proche de l'utilisateur.

What are the origins of the difficulties encountered for years in the search for viable approaches on a sufficient scale in order to create a real industrial deployment? After a state of the art of the potential techniques and technologies, we will then show, by the analysis of twelve parameters, that a reduced number of employment or deployment constraints reduces the real choice drastically.

In order to concretely illustrate our point, we will discuss three approaches to positioning. The first one is based on a " Indoor GNSS " system using local transmission of signals compatible with current receivers in order to offer good quality positioning and continuity of service. The second one is the opposite in proposing a positioning no longer continuous, but punctual, both in space and time, presenting precision and reliability. The third one is the basis of a cooperative system, much more complex than the previous ones but which could be a way to bypass the above mentioned constraints.

The conclusive discussion synthesizes the difficulties in producing an industrial solution, but these in our view reflect a limited interest from the stakeholders, lost by the high number of available solutions. We conclude on the importance to create new high level indicators in order to take care of the real needs of the potential users.

1 Introduction

Notre interrogation porte sur l'absence de fait de solutions de positionnement en intérieur déployées à échelle industrielle dans un cadre grand public. Le constat est d'autant plus surprenant que depuis maintenant plus de vingt ans, de nombreuses communautés, scientifiques et industrielles, se sont attelées au problème. De nombreuses techniques sont explorées comme les ultrasons [1], le WiFi [2] ou encore le radar interférométrique [3]. Les avancées techniques sont indéniables mais la complexité des défis à relever reste grande. Les approches actuelles se tournent ainsi vers la fusion de diverses techniques et technologies. C'est le cas notamment de la combinaison de technologies variées, allant de l'UWB (Ultra Wide Band) et de l'inertiel [4] à la lumière et les accéléromètres [5], en passant par le WiFi et l'inertiel [6] ou encore l'inertiel et la cartographie [7]. D'autres approches s'orientent vers des systèmes complexes d'analyse de données multiples dans le but de trouver de nouvelles approches. C'est par exemple le cas de systèmes collaboratifs [8] ou de la prise en compte du « crowdsourcing » [9] qui consiste à utiliser les données nombreuses des utilisateurs afin d'améliorer les modèles et les performances. Dans ce foisonnement, certains travaux sont spécifiques comme par exemple la recherche d'une solution de positionnement en intérieur à base de GNSS (Global Navigation Satellite Systems)

et de WiFi et visant une précision centimétrique [10] ou encore l'utilisation d'antennes multiples afin de traiter le cas des trajets multiples [11]. Dans cet ensemble très riche, des synthèses sont utiles et permettent de faire le point sur les travaux en cours [12], [13] ou sur les difficultés rencontrées afin d'implémenter les derniers travaux sur les smartphones actuels [14]. Dans la présente publication, nous allons tenter de montrer que ce sont souvent les contraintes d'emploi et de déploiement qui sont à l'origine de la complexité. En particulier, l'analyse d'une quarantaine de technologies, croisées avec une douzaine de paramètres, montre qu'un travail de fond devrait être mené afin de définir un niveau de contrainte acceptable, sans que ce dernier ne mène à une impasse en termes de technologies. Au minimum, ce travail permettrait de fournir aux technologues et scientifiques des orientations claires sur lesquelles se fixer. Par ailleurs, cette analyse montre également la limite des approches de fusion qui, à quelques exceptions près, ne sont pas de nature à apporter une réponse satisfaisante de part la faiblesse des complémentarités possibles.

Dans une seconde partie, nous prendrons exemple de trois approches afin de montrer que si l'on ne « relâche » pas un peu les contraintes, les difficultés actuelles sont susceptibles de ne pas se résorber à court terme. Le premier système est un GNSS d'intérieur, le deuxième un système de proximité utilisant une technologie NFC (Near Field Communication) et le troisième un système coopératif utilisant des échanges radio afin de déterminer de proche en proche un positionnement relatif des diverses entités d'un réseau d'objets communicants. Avec les contraintes actuelles, à savoir une disponibilité sur smartphone, au moins à relativement court terme, un positionnement en continu et une relative absence d'infrastructure spécifique à déployer, les deux premières approches sont disqualifiées, la troisième nécessitant un effort de développement important.

2 Un nombre très important de technologies disponibles

Les technologies étudiées vont des accéléromètres aux systèmes de réseaux sans fil dits symboliques (pour lesquels le positionnement est donné sous la forme d'une pièce ou d'un ensemble de pièces plutôt que par des coordonnées). Sans toutes les citer, notons cependant qu'elles regroupent les systèmes inertiels (accéléromètres, gyromètres, magnétomètres), les approches à base de traitement ou d'analyse d'images (marqueurs, déplacement, SLAM – pour Simultaneous Localisation And Mapping - ou encore reconnaissance), les systèmes radio (depuis les 3 et 4G jusqu'aux signaux radio d'opportunité en passant par le WiFi ou le Bluetooth, la télévision ou la radio FM) ou encore optoélectroniques comme le Laser, le Lidar ou le LiFi (Light Fidelity). D'autres types de mesures physiques sont également présents comme ceux liés au son ou aux ultrasons, à la pression atmosphérique ou encore aux infrarouges. Les systèmes d'étiquettes électroniques RFID (Radio Frequency Identification) sont également incluses, tout comme les signaux dits à faible puissance et longue portée (LPWAN – Low Power Wide Area Networks) des systèmes SiFox ou LoRa. Les réseaux filaires sont également pris en compte car il est possible de déterminer une position par identification, dans certains cas, de l'adresse IP (Internet Protocole) d'une connexion. Nous avons aussi pris en compte les systèmes de grande proximité comme les cartes sans contact ou même les cartes bancaires. Enfin il semblait difficile d'oublier totalement les GNSS qui sont quand même à la base des interrogations actuelles sur la continuité du service de positionnement entre l'extérieur et l'intérieur.

Par ailleurs, pour chacune de ces quarante technologies, une douzaine de paramètres ont été renseignés. Dans une première série on trouve la maturité de l'infrastructure à déployer le cas échéant, la maturité technologique du terminal associé, le type de positionnement (relatif, absolu ou symbolique), le fait que le terminal puisse être un smartphone à plus ou moins court terme, la sensibilité du système de positionnement en intérieur à l'environnement ou encore le besoin en calibration de la solution. Dans la seconde série, on retrouve la précision de positionnement, la fiabilité de ce dernier, le mode de positionnement (de « continu » à « réclamant l'action de l'utilisateur »), le type de traitement réalisé (modélisation de la propagation, analyse d'image ou encore détection physique), le type de calcul mené (intersection géométrique de sphères ou de plans, fonction mathématique comme intégration par exemple, ou encore calcul ponctuel de proximité) et enfin la mesure physique réalisée.

C'est alors sur ces paramètres qu'il est possible de mettre quelques contraintes afin de sélectionner les technologies répondant au(x) critère(s). Dans un premier exemple, nous cherchons à extraire les technologies qui répondent à une infrastructure existante, une précision décimétrique, une intégration sur smartphone aisée même si elle n'existe pas forcément déjà et un mode de positionnement continu pour ressembler au GNSS. Le résultat donne neuf technologies. Le tableau 1 détaille les six premiers critères et le tableau 2 les six suivants.

Il est à noter que l'on retrouve les technologies qui sont le plus souvent employées comme le BLE ou le WiFi. Il est intéressant de remarquer également que les critères sont généraux et que dans une approche de type fusion, il faudra certainement considérer des technologies du même groupe. Par exemple BLE et Capteur de pression afin de déterminer un positionnement en 3D (3 dimensions) fiable. Certaines de ces technologies restent cependant délicates à mettre en œuvre dans un environnement instable, comme par exemple les ultrasons qui seront inexploitable au milieu de personnes se déplaçant. Les techniques de traitement d'images sont très performantes

mais réclament là encore des conditions d'emploi (bonne visibilité, images très nettes, calibrations initiales) contraignantes.

<i>Technologies</i>	<i>Maturité infrastructure</i>	<i>Maturité terminal</i>	<i>Type de positionnement</i>	<i>Smartphone</i>	<i>Sensibilité à l'environnement</i>	<i>Besoin en calibration</i>
BLE	existant	Dév logiciel	absolu	existant	élevée	plusieurs fois
Image marqueurs	-	Dév logiciel.	absolu	existant	très élevée	une fois
Image dépl. rel.	-	Dév logiciel	relatif	existant	élevée	plusieurs fois
Image SLAM	-	Dév logiciel	relatif	existant	élevée	plusieurs fois
Pressure	-	existant	relatif	facile	pas d'impact	plusieurs fois
RFID	existant	Dév logiciel	absolu	facile	faible	-
Ultra Son	existant	intégration	absolu	facile	très élevée	-
WiFi	existant	Dév logiciel.	absolu	existant	élevée	plusieurs fois
WLAN symbolique	existant	Dév logiciel	symbolique	existant	faible	-

Tableau 1 : Technologies pour lesquelles infrastructure et disponibilité sur smartphone sont presque assurées

<i>Technologies</i>	<i>Précision</i>	<i>Fiabilité</i>	<i>Mode de positionnement</i>	<i>Traitement</i>	<i>Calcul</i>	<i>Physique</i>
BLE	qq m	moyenne	~ continu	pattern matcing.	fonctions math.	ElectroMagnétique
Image marqueurs	< 1m	moyenne	~ continu	combinaison	matrices	capteur image
Image dépl. rel.	< 1m	moyenne	~ continu	combinaison	fonctions math.	capteur image
Image SLAM	< 1m	moyenne	~ continu	combinaison	fonctions math.	capteur image
Pressure	1m	élevée	continu	détection	zone	capteur ϕ
RFID	dm	élevée	~ continu	détection	ponctuel	ElectroMagnétique
Ultra Son	qq dm	faible	continu	propagation	\cap sphères	mécanique
WiFi	qq m	moyenne	continu	pattern matching	fonctions math.	ElectroMagnétique
WLAN symbolique	dam	élevée +	continu	propagation	zone	ElectroMagnétique

Tableau 2 : Technologies proposant une précision décimétrique en quasi-continu

Il serait cependant possible d'aborder d'autres critères. Le second exemple considère un système existant sur les smartphones actuels, dont la sensibilité à l'environnement est faible et dont la fiabilité est très élevée. Les tableaux 3 et 4 résument les résultats pour les cinq technologies concernées.

<i>Technologies</i>	<i>Maturité infrastructure</i>	<i>Maturité terminal</i>	<i>Type de positionnement</i>	<i>Smartphone</i>	<i>Sensibilité à l'environnement</i>	<i>Besoin en calibration</i>
Codes bar	existant	Dév logiciel	absolu	existant	faible	-
Cartes de crédit	existant	existant	absolu	existant	-	-
NFC	existant	Dév logiciel	absolu	existant	-	-
QR codes	- existant	Dév logiciel	absolu	existant	faible	-
WLAN symbolique	existant	Dév logiciel	symbolique	existant	faible	-

Tableau 3 : Technologies disponibles sur smartphone et peu impactées par leur environnement

<i>Technologies</i>	<i>Précision</i>	<i>Fiabilité</i>	<i>Mode de positionnement</i>	<i>Traitement</i>	<i>Calcul</i>	<i>Physique</i>
Codes bar	dm	élevée +	action utilisateur	pattern recognition	ponctuel	capteur image
Cartes de crédit	qq cm	élevée +	discret	détection	ponctuel	électronique
NFC	qq cm	élevée +	action utilisateur	détection	ponctuel	ElectroMagnétique
QR codes	dm	élevée +	action utilisateur	pattern recognition	ponctuel	capteur image
WLAN symbolique	dam	élevée +	continu	propagation	zone	ElectroMagnétique

Tableau 4 : Technologies proposant fiabilité très élevée

Nous obtenons une liste très différente (seule la WiFi symbolique est présente dans les deux sélections). Il semble ainsi que des critères de mise en œuvre classiques soient de nature à éliminer rapidement nombre de technologies.

3 Un « radar » GNSS inversé

Une approche moins classique que celles apparaissant dans les tableaux précédents repose sur la mise en place de transmetteurs de signaux de type GNSS en intérieur. Ces signaux sont transmis par des antennes doubles émettant chacune l'équivalent du signal d'un satellite (c'est-à-dire ayant un code déterminé). Les deux antennes sont cependant parfaitement synchronisées entre elles localement. Le récepteur est un smartphone actuel équipé d'Android à partir de la version 7 permettant d'accéder aux mesures de phases des porteuses. Le traitement consiste simplement en la différence des phases des deux porteuses. Les deux antennes étant séparées de moins d'une longueur d'onde, la mesure de cette différence de distances aux deux antennes est non ambiguë. On montre que le lieu des points obtenu pour une même différence de phases est caractérisé par une quadrique et forme géométriquement un hyperboloïde de révolution. En deux dimensions, cela se traduit par des hyperboles comme exposé à la figure 1.

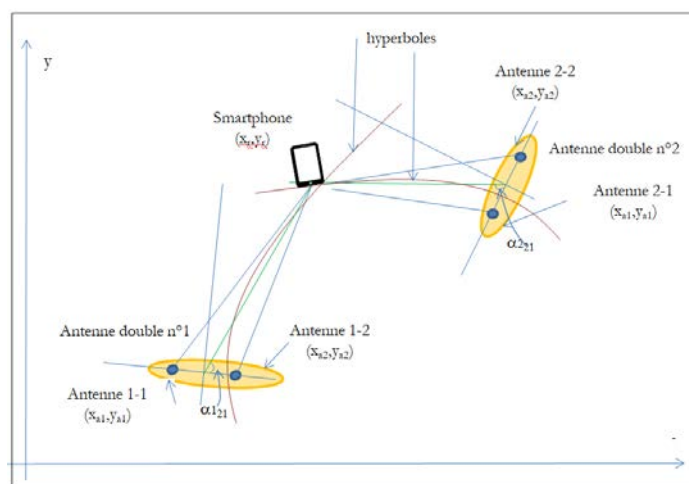


Figure 1 : Principe du radar inversé en deux dimensions

Les mesures faites à partir d'une antenne double permettent alternativement de déterminer l'angle de visée du smartphone vu de l'antenne double. Afin de positionner le smartphone, il est nécessaire d'effectuer une deuxième mesure (puis une troisième en trois dimensions) qui fournira un deuxième angle à partir de la deuxième antenne double. Il faut noter que les mesures des deux antennes doubles sont totalement indépendantes et que ces dernières n'ont pas besoin d'être synchronisées : c'est en fait le smartphone qui établit la cohérence temporelle du système.

Les avantages d'un tel système sont réels : continuité parfaite avec les GNSS utilisés en extérieur, utilisation des smartphones actuels, précision décimétrique possible, pas de synchronisation nécessaire. Cependant, il présente deux difficultés dans la perception actuelle des choses : le besoin d'une infrastructure à déployer et l'utilisation de signaux radio qui restent sensibles à l'environnement de propagation. De plus, les rares industriels qui ont proposé des approches similaires n'ont pas rencontré le succès, n'incitant pas de nouveaux acteurs à se lancer.

4 Un système NFC distribué et sa cartographie

Une approche, très différente, consiste à inclure l'utilisateur dans le système et à lui demander une action spécifique. En l'occurrence, il s'agit pour ce dernier de « flasher » un QR Codes (Quick Response Code) ou une étiquette NFC. Ces éléments sont repérés, grâce à leur identifiant unique, géographiquement dans une cartographie de l'environnement. Le fait qu'il soit nécessaire de s'approcher relativement près des étiquettes rend le système très précis, mais avec un positionnement qui est cette fois discontinu dans le temps et dans l'espace. Il semble que cette contrainte puisse être acceptée dans des sites piétons de complexité réduite car le besoin dans ce cas est principalement d'avoir une carte numérique, de savoir spécifier son emplacement et d'obtenir en retour un itinéraire particulier. Tout ceci est disponible.

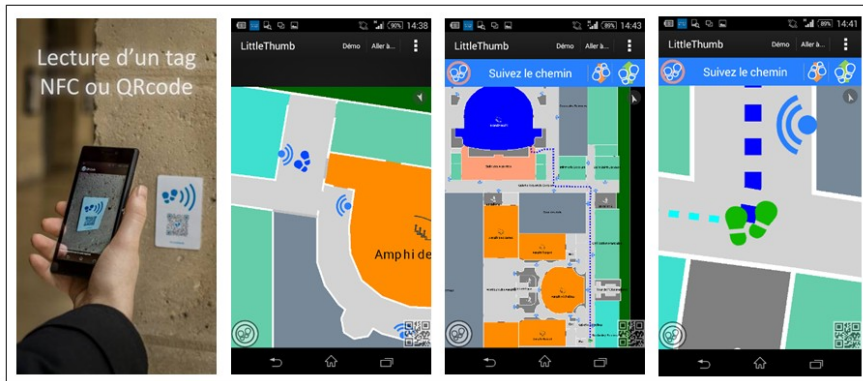


Figure 2 : Une approche à l'initiative de l'utilisateur

La cartographie est un élément majeur de tous les systèmes. Il ne s'agit pas d'une image, mais d'une structure permettant l'association d'attributs aux divers objets que sont les pièces, les couloirs ou les locaux techniques. Ces attributs vont d'une dénomination à des caractéristiques liées à la vitesse de déplacement possible dans l'objet. Il sera également important de permettre, en plus du calcul d'un itinéraire, l'estimation du temps de parcours associé. Dans l'exemple présenté à la figure 2, le magnétomètre du terminal, un smartphone ici, est utilisé afin de fournir une indication d'orientation à l'utilisateur qui ne sait pas forcément dans quel sens partir. Les « petits pieds » caractérisant l'utilisateur passent alors au vert. Ensuite, le guidage n'est plus réalisée automatiquement : si l'utilisateur en ressent le besoin, il devra alors trouver une nouvelle étiquette pour se « recalrer ». Un couplage reste possible avec soit des réseaux WiFi ou Bluetooth, soit avec un compteur de pas (fondé sur l'utilisation d'accéléromètres) pour enrichir l'expérience utilisateur.

Ce système a un coût de déploiement réduit, présente une très grande fiabilité et peut être disponible immédiatement. Cependant, il n'est pas dans la logique actuelle qui se concentre sur la continuité avec les GNSS, en particulier dans leurs aspects temporel et spatial.

5 Une approche coopérative entre terminaux communicants

Dans cette troisième version, le système cherche à calculer les positions relatives de nœuds communicants les uns par rapport aux autres. Le nœud A (en bas à gauche sur la figure 3) est la référence. Tous les nœuds estiment leur Doppler par rapport à A lorsqu'ils sont en déplacement. Par ailleurs, ils sont également constitués d'antenne de réception double permettant une estimation de l'angle d'arrivée des signaux reçus des autres nœuds. Tous les nœuds ne reçoivent pas forcément l'ensemble des signaux. La résolution géographique du système repose ici sur le calcul des divers angles repérés par des nombres sur la figure 3 (reproduite en deux dimensions pour des raisons de facilité de lecture). Ces derniers sont au nombre de 14 et il nous faut ainsi obtenir suffisamment de relations pour permettre le calcul. Cette géométrie comporte cependant de très nombreuses symétries et c'est là que les antennes doubles ont un rôle important. Même en situation statique, il est possible de croiser par couple de deux antennes doubles les mesures des signaux des autres nœuds. Il s'agit d'une approche qui ressemble au radar inversé par certains aspects, mais ici avec des entités mobiles dont les coordonnées sont inconnues.

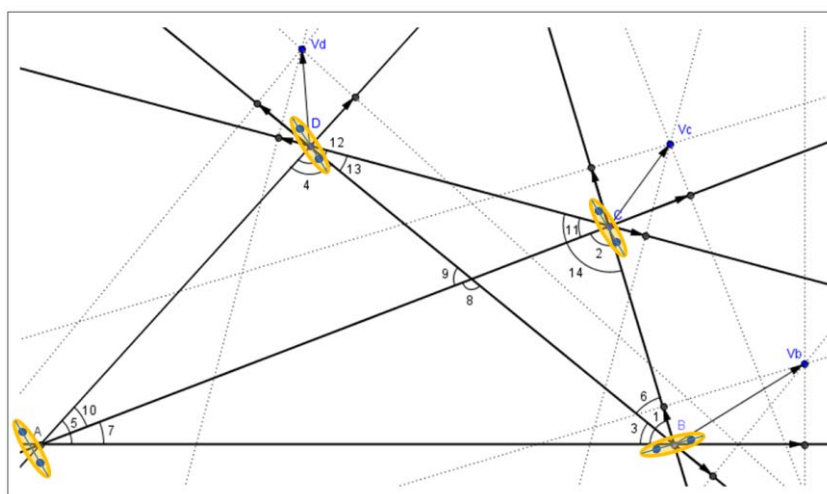


Figure 3 : Un modèle coopératif échangeant Doppler et angles d'arrivée relatifs

Trois difficultés techniques se présentent. Les mesures d'angles d'arrivée sont obtenues dans des référentiels différents, propres à chaque nœud. L'attitude, c'est-à-dire l'orientation en trois dimensions, de chaque nœud est quelconque. Enfin le calcul des intersections d'hyperboloïdes lorsque l'on est en trois dimensions se révèle particulièrement complexe dans le cas général. Une solution consiste à coupler ces mesures avec une détermination en temps réel de l'attitude de chaque nœud grâce à d'autres capteurs. Dans le cas de smartphones par exemple, ce calcul est rendu possible par la présence d'accéléromètres et de magnétomètres qui autorisent une estimation de l'attitude du récepteur. Les erreurs de positionnement résultant de l'imprécision de ces mesures restent cependant un sujet d'étude.

Dans une telle approche, il devient difficile d'imaginer que les mesures et calculs soient gérés par chaque nœud, qui devrait de plus récupérer l'ensemble des mesures réalisées. Cela générerait un trafic télécom particulièrement important. Une entité de supervision serait sans doute plus appropriée, mais réduirait de fait le côté coopératif qui peut s'entendre également par « non supervisé ». Ceci reste ouvert à discussion. Une solution consisterait alors à désigner en début de processus, le nœud qui aurait ce rôle de supervision d'un mini-réseau coopératif.

Un autre avantage de cette approche repose sur sa capacité à détecter des trajets non directs. Il est courant de nommer les trajets entre un émetteur et un récepteur, quelque soit la technologie employée, par soit LOS (Line Of Sight) dans le cas de l'absence totale d'obstacle, soit NLOS (Non Line Of Sight) dans le cas contraire. Afin de conserver à une technique ses performances nominales, il est en général nécessaire d'être en mesure d'identifier les cas NLOS, ce qui est une réelle difficulté technique qui réclame souvent la disponibilité d'une forte redondance de mesures. Dans le cas présent, les mesures d'angles d'arrivée sur deux nœuds quelconques doivent présenter une cohérence spatiale (ou géométrique) qu'il est possible d'estimer. Ainsi, sauf dans le cas de trajets non directs qui se traduiraient par des valeurs identiques à celles qu'elles auraient dues être (ce qui est très peu probable mais pas impossible), les cas NLOS devraient être identifiables aisément.

Nous voyons bien la complexité de cette approche d'un point de vue organisation des échanges, mesures, calculs et télécommunication. Certains nœuds seraient superviseurs d'un mini-réseau mais également potentiellement membres d'autres mini-réseaux. Ceci reste possible en théorie mais l'implémentation pratique est un défi. Par ailleurs, les mesures seront entachées d'erreurs très variables suivant les configurations environnementales des divers nœuds. Un mécanisme d'estimation de l'importance de ces erreurs puis d'élimination des plus dommageables devrait exister : ce dernier est une difficulté supplémentaire. Enfin, de tels terminaux ne sont pas encore disponibles et les outils de mesure associés non plus.

6 Discussions sur la « meilleure » solution

Nous ne nous sommes pas encore intéressés au paramètre « précision » qui est pourtant l'un des premiers à être mis en avant. La principale raison est qu'il existe une différence souvent importante entre les spécifications données par les équipes de développement ou de recherche et la réalité. Ceci est en général dû aux choix des environnements de tests retenus qui avantagent les uns au détriment des autres. Par ailleurs, ce paramètre devrait être associé à un second relatif à la fiabilité du positionnement. Par exemple, lorsque ce dernier dépend très fortement de l'environnement dans lequel se trouve l'utilisateur, cela réduit la fiabilité, sauf à avoir une mesure de cet environnement. C'est peut-être une voie ouverte pour les approches orientées « data science » actuelles.

Si nous revenons à la question initiale de savoir s'il vaut mieux une infrastructure locale ou des signaux d'opportunité, la réponse est immédiate : signaux d'opportunité. Cependant, tous les systèmes proposés jusqu'à présent avec cette orientation n'ont pas trouvé leur place et les solutions les plus déployées aujourd'hui restent fondées sur le déploiement de bornes spécifiques. La littérature sur les approches fusionnant diverses technologies est très importante, avec en particulier les systèmes inertiels. Ces derniers présentent en effet des caractéristiques idéales : mesures disponibles partout et à tout moment, qui de plus ne dépendent d'aucune infrastructure et ils sont d'ors et déjà présents dans les smartphones. Malheureusement, ils réclament une calibration d'autant plus fréquente que la qualité des capteurs est réduite, ce qui est le cas aujourd'hui sur ceux des smartphones.

Nous avons montré sur quelques exemples que certaines contraintes (disponibilité, maturité des technologies, fiabilité du positionnement, etc.) étaient très fortes. Les orientations qui se dégagent des travaux des nombreuses communautés impliquées [13] montrent clairement que les difficultés pratiques de mise en œuvre présente un frein aux développements industriels des approches scientifiques et techniques. Parmi les trois approches présentées dans le présent document, aucune ne semble répondre au problème, soit par besoin de déployer une infrastructure spécifique, soit par nécessité d'une action de la part de l'utilisateur et par le fait que le positionnement n'est pas continu dans le temps, ou encore par une architecture réseau non encore mature. Ces trois systèmes ne sont par ailleurs qu'une infime partie des solutions proposées.

Et si la raison de cette absence d'éclosion était ailleurs, dans en fait un intérêt réel très limité, ou une difficulté à extraire des besoins pressentis des lignes fortes qui constitueraient des orientations pour les chercheurs. Le constat semble paradoxal : l'attente pour un système de positionnement en intérieur est grand, les solutions proposées extrêmement nombreuses, mais la majorité des déploiements se concluent par un abandon progressif du dispositif mis en place et de l'application associée. Notre vision est que les besoins ne sont pas suffisamment bien spécifiés. Par exemple les performances requises sont certainement très différentes en fonction de l'usage qui est planifié. Le guidage de personnes malvoyantes dans un hôpital ne réclame pas les mêmes caractéristiques que la détection et le comptage des personnes passant devant un étalage. Cependant, la synthèse n'existe pas et ce sont souvent les chercheurs eux-mêmes qui édictent les spécifications de leur système. La rencontre entre le besoin non exprimé et la solution développée par ailleurs ne se réalise ainsi souvent pas.

De temps en temps, cette rencontre s'opère dans des cas particuliers et alors des solutions techniques sont trouvées. Cependant, sans une vision plus large, les solutions restent spécifiques. La rédaction d'un ensemble, ou de plusieurs ensembles, de spécifications qui permettraient de produire un marché suffisant et concret pour provoquer un déploiement à grande échelle est une tâche importante. Elle devrait réunir de nombreuses communautés (utilisateurs, industriels, chercheurs, etc.) et pourrait faire l'objet d'un travail institutionnel. Pour notre part, nous pouvons lancer quelques pistes de réflexion.

Tout d'abord, il nous semble que les contraintes devraient plutôt être opérationnelles que techniques. La complexité de déploiement, la sensibilité des performances aux environnements, et donc la fiabilité des performances, sont des éléments déterminants dans la perception de la solution. La constitution d'une « fiche commune » de performance permettant de comparer les technologies les unes aux autres serait très utile. Les critères d'une telle liste seraient certainement de haut niveau et orientés vers l'utilisateur. Un exemple intéressant est le projet MALIN de l'ANR [15] qui instaure des « défis » et envisage de comparer des technologies dans des cadres de mises en œuvre réels. Les critères à considérer sont sans doute également des combinaisons de critères, comme par exemple « précision et fiabilité », ou « type de positionnement et terminal ». Nous proposons de plus une classification qui permette de percevoir immédiatement la complexité d'un déploiement réel sous la forme d'un critère « couverture ». Ce dernier qualifie la portée typique de la technologie proposée sous la forme d'une donnée d'environnement : pièce, ensemble de pièces, bâtiment, ensemble de bâtiments, site, etc. Il serait également nécessaire d'associer ce critère avec un nombre de « composants » à déployer. Enfin, un paramètre sur l'efficacité énergétique de l'approche, tant pour l'éventuelle infrastructure que pour le terminal, serait utile.

7 Conclusion

Le domaine du positionnement en intérieur implique de larges communautés et propose de nombreuses solutions techniques. La différence fondamentale avec l'extérieur où les GNSS ont pris le monopôle réside dans la très grande diversité de situations. Le besoin est régulièrement rapporté mais les investissements nécessaires se heurtent souvent à l'histoire : le positionnement n'ayant pas été disponible (et ne l'étant toujours pas vraiment), les choses se sont mises en place sans lui. Quel serait ainsi l'investissement acceptable pour un besoin qui n'est en fait pas fondamental pour le moment ?

Après une analyse des technologies potentielles fondée sur l'observation de critères de haut niveau, il nous apparaît un manque de spécifications réelles, ainsi qu'une utilisation de critères bas niveau peu opérant sur les gestionnaires de bâtiments, interlocuteurs incontournables du problème. Nous proposons ainsi la création concertée d'une liste de critères mixtes qui permette une comparaison dans des conditions similaires (et représentatives) des diverses technologies. Une telle entreprise nécessite une coordination que nous appelons de nos vœux et à laquelle nous sommes prêts à prendre de notre part.

8 Références bibliographiques

[1] Alessio De Angelis; Antonio Moschitta; Paolo Carbone; Massimo Calderini; Stefano Neri; Renato Borgna; Manuele Peppucci, "Design and Characterization of a Portable Ultrasonic Indoor 3-D Positioning System," in IEEE Transactions on Instrumentation and Measurement, Volume 64, Issue 10, Pages 2616 – 2625, 2015.

[2] Suining He; S. -H. Gary Chan, "Wi-Fi Fingerprint-Based Indoor Positioning: Recent Advances and Comparisons," in IEEE Communications Surveys & Tutorials, Volume 18, Issue 1, Pages 466 – 490, 2016.

[3] Guochao Wang; Changzhan Gu; Takao Inoue; Changzhi Li, "A Hybrid FMCW-Interferometry Radar for Indoor Precise Positioning and Versatile Life Activity Monitoring," in IEEE Transactions on Microwave Theory and Techniques, Volume 62, Issue 11, Pages 2812 – 2822, 2014.

[4] Manon Kok; Jeroen D. Hol; Thomas B. Schön, "Indoor Positioning Using Ultra wide band and Inertial Measurements," in IEEE Transactions on Vehicular Technology, Volume 64, Issue 4, Pages 1293 – 1303, 2015.

- [5] Muhammad Yasir; Siu-Wai Ho; Badri N. Vellambi, "Indoor Positioning System Using Visible Light and Accelerometer," in *Journal of Lightwave Technology*, Volume 32, Issue 19, Pages 3306 – 3316, 2014.
- [6] Lyu-Han Chen; Eric Hsiao-Kuang Wu; Ming-Hui Jin; Gen-Huey Chen, "Intelligent Fusion of Wi-Fi and Inertial Sensor-Based Positioning Systems for Indoor Pedestrian Navigation," in *IEEE Sensors Journal*, Volume 14, Issue 11, Pages 4034 – 4042, 2014.
- [7] Arto Perttula; Helena Leppäkoski; Martti Kirkko-Jaakkola; Pavel Davidson; Jussi Collin; Jarmo Takala, "Distributed Indoor Positioning System With Inertial Measurements and Map Matching," in *IEEE Transactions on Instrumentation and Measurement*, Volume 63, Issue 11, Pages 2682 – 2695, 2014.
- [8] Youngtae Noh; Hirozumi Yamaguchi; Uichin Lee, "Infrastructure-Free Collaborative Indoor Positioning Scheme for Time-Critical Team Operations," in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, Volume 48, Issue 3, Pages 418 – 432, 2018.
- [9] Yuan Zhuang; Zainab Syed; You Li; Naser El-Sheimy, "Evaluation of Two WiFi Positioning Systems Based on Autonomous Crowdsourcing of Handheld Devices for Indoor Navigation," in *IEEE Transactions on Mobile Computing*, Volume 15, Issue 8, Pages 1982 – 1995, 2016.
- [10] Chen Chen; Yi Han; Yan Chen; K. J. Ray Liu, "Indoor Global Positioning System with Centimeter Accuracy Using Wi-Fi," in *IEEE Signal Processing Magazine*, Volume 33, Issue 6, Pages 128 – 134, 2016.
- [11] Abdo Gaber; Abbas Omar, "Utilization of Multiple-Antenna Multicarrier Systems and NLOS Mitigation for Accurate Wireless Indoor Positioning," in *IEEE Transactions on Wireless Communications*, Volume 15, Issue 10, Pages 6570 – 6584, 2016.
- [12] Qi Liu; Jiahui Qiu; Yi Chen, "Research and development of indoor positioning," in *China Communications*, Volume 13, Issue 2, Pages 67 – 79, 2016.
- [13] Davide Dardari; Pau Closas; Petar M. Djuric, "Indoor Tracking: Theory, Methods, and Technologies," in *IEEE Transactions on Vehicular Technology*, Volume 64, Issue 4, Pages 1263 – 1278, 2015.
- [14] Pavel Davidson; Robert Piché, "A Survey of Selected Indoor Positioning Methods for Smartphones," in *IEEE Communications Surveys & Tutorials*, Volume 19, Issue 2, Pages 1347 – 1370, 2017.
- [15] <http://www.agence-nationale-recherche.fr/suivi-bilan/historique-des-appels-a-projets/appel-detail1/challenge-maitrise-de-la-localisation-indoor-malin-2017/>

Avancées et défis de la localisation véhiculaire coopérative

.....

Gia Minh Hoang^{1,2}, Benoît Denis¹, Jérôme Härri²

¹ CEA-Leti Minatec, Systems Integration Dept.,
benoit.denis@cea.fr

² URECOM, Communication Systems Dept., Sophia Antipolis

Mots clés (en français et en anglais) : Algorithmes coopératifs, Communications V2V, Fusion de données, Localisation véhiculaire.

Présentation en session plénière / Plenary session communication

Résumé / Abstract

A terme, les applications associées aux systèmes de transports intelligents coopératifs (C-ITS) requerront une information de localisation fiable et précise. En particulier, pour les applications C-ITS de deuxième génération, telles que la conduite autonome (HAD) ou la prévention des risques d'accident pour les usagers vulnérables de la route (VRUs), le système de positionnement par satellites (GNSS) semble aujourd'hui insuffisant, notamment en matière de continuité de service. Ces applications nécessitent en effet un niveau de précision sub-métrique constant, quelles que soient les conditions d'utilisation.

En théorie, la Localisation Coopérative (CLoc) devrait permettre aux véhicules de palier la faiblesse relative de leurs propres systèmes de localisation, en s'appuyant sur 1) des échanges de données de proche en proche (ex., leurs dernières positions estimées, voire leurs données GNSS brutes), 2) l'acquisition de distances relatives les séparant (typiquement, sur la base de liens radio), 3) d'autres modalités de mesure issues de capteurs embarqués (inertiel, odométrie...) et enfin, 4) la fusion de l'ensemble de ces informations. Toutefois, dans le contexte très particulier des réseaux véhiculaires ad hoc (VANET), les approches coopératives conventionnelles ne s'avèrent pas pleinement satisfaisantes et nécessitent des ajustements importants, en raison de la spécificité des motifs de mobilité rencontrés, d'un canal de communication véhicules à véhicules (V2V) fortement contraint (en termes de capacité et de contrôle décentralisé des émissions) et de la nature des mesures à intégrer dans le processus de fusion (potentiellement non Gaussiennes et corrélées dans le temps et/ou l'espace).

Dans le cadre de cette conférence invitée, nous introduirons à titre d'exemple un schéma de fusion générique reposant sur un simple filtre particulière, ainsi que sur l'exploitation de messages coopératifs normalisés de type 'CAM' (cooperative awareness messages) prévus par le standard V2V ITS-G5. Nous décrirons en particulier une méthode permettant de synchroniser les données reçues de la part des véhicules voisins, tout en réduisant la complexité de calcul et l'impact sur les communications inter-véhiculaires (ex. via une sélection des liens/voisins les plus informatifs, l'approximation paramétrique des messages, ainsi qu'un contrôle du format et du taux des paquets émis). Nous étendrons ensuite ce schéma de manière à intégrer des mesures de distances V2V plus précises, basées sur la technologie radio impulsionnelle 'IR-UWB'. A cette occasion, nous aborderons les problèmes spécifiques de propagation d'erreur et de confiance excessive dans les résultats de la fusion, caractéristiques des filtres coopératifs. Finalement, afin de traiter certains cas particulièrement pathologiques (ex. absence de GNSS dans les tunnels, forte erreur selon l'axe orthogonal à la route), nous considérerons l'apport d'autres capteurs embarqués (centrale inertielle, compte-tour et détecteurs de voie), ainsi que d'unités de bord de route (RSUs) dotées de moyens de communication V2I.

L'approche proposée a été testée sur la base de simulations canoniques, de données de mobilité réalistes, ainsi que de données expérimentales. Ces travaux de recherche ont fait l'objet d'une thèse récente, menée au CEA-Leti (Grenoble) en collaboration avec EURECOM (Sophia Antipolis).






COOPERATIVE VEHICULAR LOCALIZATION: RECENT PROGRESSES AND CHALLENGES



G.M. Hoang^{1,2}, B. Denis¹, J. Härrig¹, D. Stock²
¹CEA-LETI, MINATEC Campus, Grenoble
²EURECOM, Communication Systems Dept., Sophia Antipolis

Observatoire de Paris, Meudon - March 29th 2018

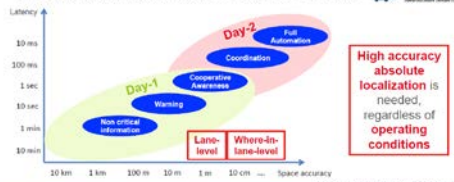
OUTLINE

- Context and Motivations
- V2V Cooperative Localization
- Hybrid V2V Cooperative Localization
- Hybrid V2X Multisensor Cooperative Localization
- Conclusions and Perspectives






COOPERATIVE - INTELLIGENT TRANSPORT SYSTEMS (C-ITS)

- Wireless communication between vehicles (V2V) and roadside infrastructure (V2I) → V2X
 - Road traffic safety
 - Road traffic efficiency
- C-ITS applications road map (C2C-CC): Day-1 & Day-2




High accuracy absolute localization is needed, regardless of operating conditions






COOPERATIVE LOCALIZATION BASICS

- Expected benefits
 - Neighbors (hopefully well positioned) → "Virtual anchors"
 - Diversity, redundancy, geometric ambiguity solving → Better accuracy/resilience





- Methods mostly validated under moderate mobility so far (e.g., WSN) → Open/unprecedented challenges in the vehicular context

CONSIDERED TECHNOLOGIES



Maturity	Technology	Frequency	Metric
Today	ITS-G5 / 802.11p	5.9 GHz	RSSI
Today	IR-UWB / 802.15.4a	~4 GHz	TOA / RT-TOF
Prospective	4G LTE V2X	2 GHz	Under specification
Prospective	5G mmWave V2X	30 – 100 GHz	AOA / AOD / TOA
Prospective	WiFi extension	2.4 GHz	Not standardized



CHALLENGES & OPEN RESEARCH QUESTIONS

- Highly dynamic mobility
- Large amount of vehicles
- Limited V2X communication channel (300 – 800 bytes, 6 Mbps)
- Unscheduled V2X communications
- Imperfect/unfavorable neighboring positions
- Unplanned geometry

Can sub-meter localization accuracy be already met through low-complexity CLoc strategies between connected vehicles with standard technologies?

GRADUAL ASSESSMENT APPROACH

NOMINAL COOPERATIVE FRAMEWORK

- At each "ego" vehicle...
 - Prediction "Synchronization"
 - Link Selection
 - Correlation Mitigation
 - Correction
 - Tx Control Strategies
 - Message Approximation

LINKS SELECTION
NON-BAYESIAN VS. BAYESIAN CRITERIA

- Link selection based on theoretical positioning performance bounds (CRLB) conditioned on a priori sub-constellations
- Non-Bayesian CRLB criterion [Hoang15a]
 - Radio link quality
 - Geometry of neighboring vehicles (GDOP)
 - All involved positions assumed deterministic (& perfect)
- Bayesian CRLB criterion [Hoang15b]
 - Radio link quality
 - Geometry of neighboring vehicles (GDOP)
 - Uncertainty of neighbors' estimated positions

Presumed probability density of local position estimates (possibly transmitted also in CAMs)

LINKS SELECTION
PERFORMANCE EVALUATION

- Large-scale GNSS error (urban canyon)
 - Saved complexity at (almost) no accuracy degradation (vs. exhaustive cooperation)
- Small-scale locally degraded GNSS capability
 - Local accuracy gains with Bayesian-CRLB criterion (vs. non-Bayesian)

IMPACT OF SPACE-TIME MEASUREMENT
NOISE CORRELATIONS

- Why is correlation a threat?
 - Inherent/specific to constrained vehicular mobility under typical refresh rates
 - Cannot properly filter out error processes (assumed white)
 - Misses hidden/fruitful location info
 - Causes filter over-confidence (in inaccurate estimates)

MITIGATION OF SPACE-TIME CORRELATIONS

- Signal level mitigation
 - Empirical cross-measurement correlations
 - Compensate for info loss
 - Differential measurements
 - Eliminate the correlated processes (i.i.d./white assumptions)
- Protocol level mitigation
 - Adaptively decreased cooperative fusion rate
 - Collect uncorrelated measurements

MITIGATION OF SPACE-TIME CORRELATIONS
PERFORMANCE EVALUATION

- ECDF of localization errors for different correlation mitigation schemes in a highway scenario (steady-state mobility)

IMPACT ON V2X COMMUNICATIONS

- Location estimation by distributed particle filter (PF)
 - Posterior by a set of random state samples
 - Any process nonlinearity and noise distribution
 - High number of particles, generating heavy communication load due to belief messages passing
- Challenges
 - Limited CAM size
 - Limited channel capacity
 - ETSI Decentralized Congestion Control (DCC)
 - Reduced CAM rate (e.g., 2 Hz) → Expected accuracy degradation

MESSAGE APPROXIMATION & TX CONTROL

- Parametric message approximation → Reduce the size of particles info
- Tx payload/rate control → Standard CAM Tx policy vs. mixed CAM traffic
- Tx power control for "tiny CAMs" (for RSSI only)

MESSAGE APPROXIMATION & TX CONTROL
PERFORMANCE EVALUATION

- ECDF of localization errors for different message approximation and transmission control strategies (1000 particles)

Challenges to Particle Filter based Fusion

- High dimensional state and high/peaky likelihood → Harmful to PF
 - Number of particles vs. state space
 - "Mismatch" between prior and likelihood
 - Particles depletion
 - Filter overconfidence
 - Bias propagation through CLoc
- PF-based GNSS+IR-UWB fusion
 - Neighbors positioned with uncertainties → High dimensional estimation space
 - Good prior not always guaranteed → Wide prior
 - Accurate ranges (e.g., IR-UWB) → Peaky likelihood
- Questionable PF efficiency in case of IR-UWB+GNSS fusion ?

JS 2018 URIS, Meudon | March 29, 2018 | 120

PF Depletion & Bias Propagation (1)

- Ex. of overconfidence in biased state estimates due to particles depletion (large prior vs. narrow likelihood) with 2500 particles

JS 2018 URIS, Meudon | March 29, 2018 | 121

PF Depletion & Bias Propagation (2)

- Ex. of unrealistically higher (10⁴) nb of particles (same scenario)
 - More particles have meaningful weights → No more overconfidence and preserved correction power from accurate observations but...
 - Unaffordable for real-time (high computational complexity)

JS 2018 URIS, Meudon | March 29, 2018 | 122

PF Depletion & Bias Propagation (3)

- Bias propagation from "Virtual Anchors"

JS 2018 URIS, Meudon | March 29, 2018 | 123

Improved PF Efficiency for High Dim. Estimation with Accurate V2V Ranges

- Scheduling (heterogeneous GNSS conditions) [Hoang16c]
- Adaptive Bayesian dithering (homogenous GNSS conditions)
 - Adaptive smoothed likelihood in perception model
 - Based on theoretical bounds e.g., BCRLB (same as for link selection)
 - Dithering noise gradually added in filter's perception so as not to outperform the BCRLB
- Main expected benefits
 - Reasonable number of particles
 - Minimized loss of accurate range info

JS 2018 URIS, Meudon | March 29, 2018 | 124

Improved PF Efficiency Settings

- Highway environment
 - 3-lane highway
 - IR-UWB network ~ 10 neighbors
 - Gauss-Markov traffic
 - $v_k = av_k - 1 + (1-a)v_k + \sqrt{1-a} \epsilon_k$
- Main simulation parameters

GNSS errors in x-y-axes (1σ)	1.5 m	Unbalanced noises
IR-UWB ranging error (1σ)	0.2 m	
Initial positional errors in x-y-axes (1σ)	1 m	large prior
Initial velocity errors in x-y-axes (1σ)	0.1 m/s	
Number of particles	1000	reasonable nb
- Performance comparisons
 - PF (GNSS, GNSS+RSSI, GNSS+IR-UWB (part. depletion vs. adapt. dithering))
 - EKF (GNSS+IR-UWB)

JS 2018 URIS, Meudon | March 29, 2018 | 125

Improved PF Efficiency Performance Evaluation

- Over-confidence depending on both
 - Actual 1-σ (68th percentile) localization errors
 - Perceived/Estimated 1-σ localization errors by fusion filters

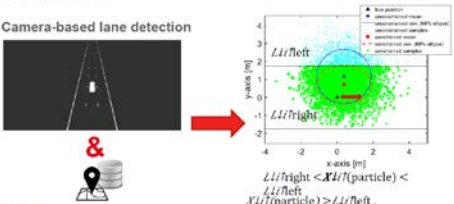
JS 2018 URIS, Meudon | March 29, 2018 | 126

Remaining Challenges & Pathological Cases

- Unbalanced vehicular geometry ~ 1-D → Singular cross-track axis
 - safety distance 50 – 100 m
 - lane width ~ 3.5 m
- Dead reckoning errors accumulation in GNSS-denied scenarios → Error propagation
 - Along-track error: Good geometry → beneficial cooperation (green smiley)
 - Cross-track error: Poor geometry counterproductive cooperation (red sad smiley)

JS 2018 URIS, Meudon | March 29, 2018 | 128

leti **MULTISENSOR FUSION FOR IMPROVED CROSS-TRACK LOCALIZATION** **EURECOM**

- IMU gyroscope ω_{ik} integration
 - $x_{ik} = x_{i,k-1} + \Delta T \omega_{ik} \cos(\theta_{ik} + 0.5 \Delta T \omega_{ik})$
 - $y_{ik} = y_{i,k-1} + \Delta T \omega_{ik} \sin(\theta_{ik} + 0.5 \Delta T \omega_{ik})$
 - $\theta_{ik} = \theta_{i,k-1} + \Delta T \omega_{ik}$
- Camera-based lane detection
 - 
 - $$L_i^{\text{right}} < X_i^{\text{left}}(\text{particle}) < L_i^{\text{left}}$$

$$X_i^{\text{left}}(\text{particle}) \geq L_i^{\text{left}}$$

$$X_i^{\text{right}}(\text{particle}) \leq L_i^{\text{right}}$$

[Hoang17b] © 2018 URSI, Monaco | March 20, 2018 | 129

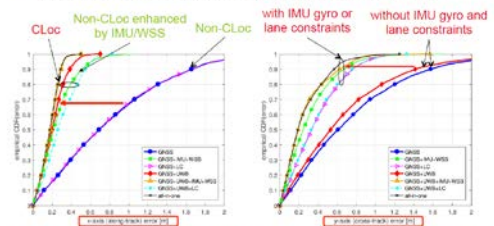
leti **IMPROVED CROSS-TRACK ERRORS SETTINGS** **EURECOM**

- Highway environment
 - 2-lane highway, 7 vehicles
 - Gauss-Markov mobility traffic
 - $v_{ik} = av_{i,k-1} + (1-a)v_i + v_i - aT_2 \cdot c_{ik}$
- Performance comparisons
 - 2 main configurations: **non-CLoc** vs **CLoc (V2V IR-UWB)**

Configuration	Result
GNSS	Non-CLoc
GNSS + IMU + WSS	Non-CLoc
GNSS + lane constraints	Non-CLoc
GNSS + V2V IR-UWB	CLoc
GNSS + V2V IR-UWB + IMU + WSS	CLoc
GNSS + V2V IR-UWB + lane constraints	CLoc
GNSS + V2V IR-UWB + IMU + WSS + lane constraints	CLoc

© 2018 URSI, Monaco | March 20, 2018 | 130

leti **IMPROVED CROSS-TRACK ERRORS PERFORMANCE EVALUATION** **EURECOM**

- ECDF of 1-D localization errors along x (left) and y (right) axes
 - 
 - Individual information source affects each component of position error differently

© 2018 URSI, Monaco | March 20, 2018 | 131

leti **CONCLUSIONS** **EURECOM**

- Sub-meter accuracy through CLoc with existing technologies?
 - (Conditionally) yes!
 - Typically, precision improved from 2 m down to 30 cm in 80% in most favorable simulated scenarios
- Various challenges inherent to the cooperative vehicular context
 - Information asynchronism
 - Space/time measurement correlations
 - Computational complexity and information selection
 - Communication constraints (imposed by underlying standards)
 - Relative geometry
- Other open questions ahead (future work)
 - Context-aware cooperative fusion (large-scale/long-term)
 - Security and privacy of involved V2X cooperative links
 - Fusion partitioning and data kind (e.g., w/rt. juridical responsibility)
 - See autonomous cars accidents)
 - New location-enabled applications and services (mapping/cartography, automotive IoT, crowd sensing...)

© 2018 URSI, Monaco | March 20, 2018 | 133



Géolocalisation et navigation à l'intérieur des bâtiments *Indoor geolocation and navigation*

Jean-Pierre Barbot¹, Isabelle Kyoko Vin¹, Pan Liu¹, Ludovic Chamoin², and Dominique Placko¹

¹Laboratoire SATIE, ENS Paris-Saclay, Jean-Pierre.Barbot@ens-paris-saclay.fr

²Laboratoire LMT, ENS Paris-Saclay, chamoin@lmt.ens-cachan.fr

Keywords: Indoor geolocalisation, fingerprinting, Kalman filtering, propagation modelisation, DPSM

Mots-clefs: Géolocalisation indoor, fingerprinting, filtrage de Kalman, modélisation de la propagation, DPSM

Abstract:

In the context of Internet of Everything (IoE), geolocation has become, in the same way as wireless connectivity to Internet, a functionality in mainly electronic communicating systems as smartphones, tablets and connected devices. While in outdoor, geolocation is mainly performed owing to satellite positioning systems (GNSS) and/or mobile networks, these systems are ineffective indoor. In this paper, we present a state-of-the-art of signals and means used for indoor geolocation, as well as related localization algorithms. Then, the parametric study of a WiFi fingerprinting method proposing different norms is shown.

Résumé:

Dans le contexte de l'Internet of Everything (IoE), la géolocalisation est devenue, au même titre que la connectivité sans fil, une fonctionnalité présente dans la plupart des systèmes communicants. Tandis qu'à l'extérieur, la géolocalisation s'opère principalement grâce aux systèmes de positionnement par satellite (GNSS) et/ou aux réseaux mobiles, ces systèmes sont inopérants à l'intérieur des bâtiments. Dans cet article, nous présenterons un état de l'art des signaux et moyens utilisés pour la géolocalisation dans les bâtiments, ainsi que les algorithmes associés. L'étude paramétrique d'une méthode de fingerprinting WiFi exploitant différentes métriques sera ensuite présentée.

1 Introduction

Les systèmes de radio-navigation ont fortement évolué ces dernières années. Ils ont en effet bénéficié de la miniaturisation des récepteurs de signaux GNSS, tels que ceux du GPS ou de GALILEO, et, de ce fait, sont désormais intégrés dans des terminaux mobiles tels que les smartphones. Les données de géolocalisation par satellite, couplées à d'autres capteurs, à des logiciels de cartographie et enrichies de données temps réel sur la congestion du trafic routier, permettent une gestion optimisée des déplacements.

Toutefois, l'estimation de la position à partir de signaux GNSS n'est techniquement possible que si quatre satellites sont en visibilité du récepteur, ce qui n'est pas toujours le cas dans des milieux urbains denses. Dans de telles situations, des capteurs sont exploités afin de maintenir une géolocalisation. Citons par exemple les accéléromètres, les magnétomètres et les gyroscopes. A ces mesures sont adjointes des informations liées aux signaux perçus : stations de base de téléphonie mobile, modems WiFi. Les localisations *a priori* connues de ces nœuds réseau sont en effet des informations d'opportunités exploitables pour une géolocalisation approximative. En revanche, dans les bâtiments, l'absence quasi-permanente des signaux GNSS et de données cartographiques numériques n'autorisent pas une géolocalisation précise. Pourtant une forte demande de géolocalisation existe dans ces environnements indoor. Les besoins s'expriment pour des applications « de confort » à usage civil, telles que la navigation interactive dans certains lieux (musées, monuments historiques ou hôpitaux). Pour y répondre, de nombreux travaux ont été initiés et des systèmes commerciaux de radio-navigation indoor commencent à être proposés [1]. D'autres besoins de géolocalisation indoor, dans le domaine de la sécurité, civile ou militaire, sont de plus en plus recensés, pour des applications de surveillance, de recherche de personnes ensevelies suite à un sinistre (avalanche, incendie, etc.), ou encore d'aide à l'intervention dans un contexte de prise d'otage [2]. La structuration de cet article sera donc la suivante. Dans un premier temps, nous recenserons les signaux et moyens permettant une géolocalisation indoor [3] : WiFi, balises Bluetooth, LiFi, RFID, balises UWB (Ultra Wide Band), capteurs inertiels, capteurs optiques, etc. Les algorithmes de traitement de données associés à ces signaux seront présentés. Parmi ces techniques, la méthode de l'identification d'empreinte (ou fingerprinting) [4] sera distinguée et fera l'objet d'une étude de cas. Cette méthode sera appliquée à des signaux WiFi et une étude paramétrique des performances expérimentales obtenues en terme de géolocalisation indoor sera présentée, avec notamment plusieurs métriques proposées.

2 État de l'art des signaux et moyens utilisés pour la géolocalisation indoor.

Afin d'opérer une localisation précise dans un bâtiment ou dans un milieu clos, il est nécessaire d'exploiter une à plusieurs sources d'informations. Ces informations peuvent être déduites de signaux radio, acoustiques, thermiques, visuels ou optiques, et/ou de mesures inertielles. Il peut s'agir de signaux d'opportunité ne nécessitant aucun déploiement particulier d'équipement émetteur, tels que les signaux visuels, les signaux infrarouges thermiques, ou encore les signaux issus des réseaux mobiles et WiFi. Il peut s'agir également de signaux transmis pour les besoins de localisation, via l'installation de balises UWB, Bluetooth, ou d'émetteurs acoustiques par exemple.

2.1 Signaux radio.

La localisation par les signaux radio est basée sur l'exploitation de caractéristiques particulières du signal. Dans le cas le plus simple, seule la puissance du signal reçu (ou RSS) est exploitée. Si le signal radio a une bande passante suffisamment large, et donc une résolution temporelle associée suffisante, le temps de vol du signal radio peut être exploité pour la localisation. Enfin, si le récepteur radio est équipé d'un réseau d'antennes, les directions (ou angles) d'arrivée estimées.

Les signaux des réseaux GSM, UMTS ou LTE sont exploitables pour une localisation directe « en un saut » [5] si la couverture du bâtiment est totale. Sinon, la coopération entre les mobiles peut permettre de remonter à la position par des techniques multi-hop [6].

Les signaux WiFi sont similairement exploités. Ce sont généralement leurs caractéristiques RSS qui servent à la localisation (fingerprinting). Alternativement, ces caractéristiques RSS peuvent être converties en une pseudo-distance via un modèle de propagation, et permettent d'effectuer soit une trilatération (dans une configuration où les signaux sont reçus par le récepteur souhaitant se localiser), soit de l'Imagerie Tomographique Radio (ITR ou RTI en anglais) - si les signaux ont été traversés et atténués par la « cible » humaine- pour localiser l'équipement ou la personne. D'autres systèmes plus complexes sont parfois déployés et exploitent la réflexion des ondes WiFi en utilisant également les caractéristiques temporelles (temps de vol) voire angulaires (angles d'arrivée) des signaux mesurés grâce à des réseaux d'antennes : citons le système RF-capture du MIT basé sur une technique proche de l'imagerie radar à travers les murs TWRI (Through-the-Wall Radar Imaging) [7].

Les signaux Bluetooth sont utilisables en géolocalisation indoor [8]. La localisation par signaux Bluetooth peut s'opérer via la coopération de terminaux mobiles. Si la densité d'équipements Bluetooth ne peut être garantie pour opérer la localisation, des balises Bluetooth peuvent être déployées dans le bâtiment.

Les signaux UWB sont très répandus en localisation indoor. Avec une bande passante comprise entre 500 MHz et quelques Gigahertz, ils offrent une grande précision de localisation. Ces signaux sont généralement obtenus par le déploiement de balises UWB fixes. On retrouve également ces signaux UWB dans les techniques de radar pulsé pour la TWRI [9] ou le radar pulsé Doppler [10].

Similairement, les signaux Terahertz permettent d'effectuer une localisation indoor via l'imagerie active radar faible bande pour la détection et la localisation à travers les murs [11], en utilisant des réflecteurs paraboliques et à des fréquences comprises entre 200 GHz et quelques THz. Ces signaux présentent l'avantage d'une très bonne pénétration à travers les murs et matériaux non-conducteurs de manière générale.

Les signaux UHF de manière générale peuvent être employés en géolocalisation passive de cibles à travers l'emploi de techniques d'Imagerie Tomographique par Radio. Des émetteurs UHF sont positionnés sur deux des quatre murs de la pièce où doit s'opérer la localisation, tandis que les récepteurs UHF se trouvent en vis à vis, sur les deux autres murs [12]. A partir des variations d'atténuations de signal sur les différents récepteurs, la position de la cible mobile peut être estimée.

Enfin, les signaux RFID sur la bande UHF sont également exploités pour les systèmes de géolocalisation indoor [13].

2.2 Signaux acoustiques.

Dans les systèmes acoustiques de localisation indoor [14], les signaux ultrasons sont exploités pour estimer des temps de vol entre l'émetteur-balise et le récepteur embarqué sur l'appareil à localiser. Leur portée est limitée à une dizaine de mètres. Ces signaux ne pénètrent pas à travers les murs, leur exploitation n'est donc efficace que s'il y a visibilité directe entre eux. Généralement, ils sont combinés avec d'autres signaux tels que les signaux RF pour gagner en précision, en exploitant par exemple la différence de temps de vol entre les signaux ultrasons et RF [15].

2.3 Signaux optiques.

Les plus anciens systèmes optiques de localisation indoor sont basés sur les signaux infrarouges pulsés, comme le système Active Badge [16]. La portée des signaux infrarouges est de l'ordre de 10 mètres et, tout comme les ultrasons, ces signaux sont inopérants en l'absence de visibilité directe émetteur-récepteur.

Les images vidéo/photo capturées par les caméras sont également des signaux exploités pour la localisation indoor. Ces capteurs optiques sont en effet présents sur de nombreux systèmes électroniques communicant. Les images permettent de détecter un élément connu dans une scène, tel qu'un tableau dans un musée, et de le localiser dans une scène en déterminant les transformations entre l'image originale et les angles de vue du capteur optique pour différentes images successives [17].

Enfin, parmi les signaux optiques exploitables pour la localisation, les plus récents sont les signaux LiFi (Light Fidelity) [18]. Ces signaux visibles, dont la longueur d'onde est comprise entre 480 et 650 nm, sont destinés à des communications visibles sans fil (Visible Light Communications – VLC) haut débit en passant par les sources d'éclairage du bâtiment. Les signaux étant monodirectionnels, la localisation s'opère par réception des signaux LiFi sur l'équipement à localiser. Comme tout signal optique, la visibilité directe est requise.

2.4 Signaux infrarouges thermiques.

Il s'agit de signaux opportunistes, naturellement émis par les objets en fonction de leur température. Ces signaux, de longueur d'onde comprise entre 2 et 15 mm, sont capturés par des capteurs thermiques infrarouges (IR). Comme pour les images vidéo « visibles », la localisation se fait à partir de l'estimation des transformations (rotation, translation) des différentes prises de vue des images thermiques IR successives [19].

2.5 Champs magnétiques et électromagnétiques MF.

Les champs magnétiques sont également utilisés pour la localisation indoor. Des boucles magnétiques sont déployées au sol à des positions fixes connues et émettent en permanence des champs magnétiques. La polarité des signaux émis sur chaque boucle magnétique du « réseau » est modulée par une séquence pseudo-aléatoire [20].

La phase du champ électromagnétique des signaux radio AM à 2 MHz a également été exploitée dans certains systèmes de localisation indoor [21].

2.6 Capteurs inertiels.

Les capteurs inertiels (accéléromètres, gyroscopes et magnétomètres) représentent une source de données très utilisée en navigation. Présents dans tous les smartphones et dans de nombreux objets connectés, ils permettent d'opérer la localisation par odométrie ou en complément avec un autre moyen de géolocalisation « absolu » [17, 22].

3 Algorithmes pour la géolocalisation indoor

Cette partie résume l'ensemble des algorithmes et techniques de géolocalisation indoor associés aux signaux et moyens précédemment cités.

3.1 Trilatération - triangulation.

Il s'agit d'algorithmes basés sur les techniques dites géométriques. La position du mobile à localiser est estimée à partir de *lignes de positions*, déduites d'une à plusieurs caractéristiques des signaux mesurés : temps de vol, RSS (converti en une pseudo-distance à l'aide d'un modèle de propagation indoor) ou angle d'arrivée. L'algorithme de trilatération consiste à calculer l'intersection d'au moins trois cercles (position 2D) ou quatre sphères (position 3D) déduits des temps de vol ou RSS mesurés. Lorsque l'instant d'émission des signaux n'est pas connu du récepteur (absence de synchronisation), une alternative est d'utiliser les temps de vol différentiels (TDOA-Time Difference of Arrival) entre les signaux issus d'émetteurs pris deux à deux. L'algorithme de triangulation utilise quant à lui des lignes de directions d'arrivée pour estimer la position du mobile.

3.2 Localisation coopérative multi-hop.

Utilisés lorsque la couverture du réseau sans fil (GPS, WiFi, réseaux 3G/4G, etc.) est limitée, les algorithmes de localisation coopérative sont basés sur la connectivité entre les mobiles (ou *devices*). Dans ce cas, des algorithmes multihop tels que Distance-Vector hop ou ses variantes [23] peuvent être utilisés. Des messages de broadcasting sont transmis par quelques « mobiles- ancres » de positions connues, et relayés par les autres mobiles environnants. Un mobile peut alors calculer sa position à partir du nombre de sauts effectués par le message de broadcasting pour l'atteindre en utilisant un indicateur de distance (RSS ou temps de vol) pour effectuer la trilatération finale. La localisation coopérative peut être également implémentée pour tirer profit d'un environnement très dense en mobiles, où le mobile à localiser est le plus souvent en non-visibilité directe des ancres principales. Parmi tous les signaux sans fil disponibles, l'algorithme déterminera les liaisons LOS (Line-Of-Sight), par des tests d'hypothèses [24] ou des méthodes de maximum de vraisemblance [25].

3.3 Radar.

En géolocalisation indoor, ce sont principalement les techniques de radar pulsé UWB multi-antennaires (MIMO), voire polarimétriques, qui sont utilisées. Dans le cas d'un radar MIMO polarimétrique, les caractéristiques spatio-temporelles (temps, angles d'arrivée et amplitudes polarimétriques associées) des signaux sont estimées par des algorithmes haute résolution de type MUSIC ou SAGE [26]. Dans les systèmes TWRI, un algorithme de correction des effets des murs, basé sur un modèle de propagation, est appliqué, suivi d'un algorithme de localisation des *cibles* (*i.e. éléments de la scène*) tel que la trilatération ou la rétroprojection [27]. Certains systèmes radar pulsés exploitent le spectre Doppler, au lieu du temps de vol aller-retour, pour estimer la distance cible/radar.

Les techniques radar à onde continue CW (*Continuous Wave*) sont potentiellement utilisées en localisation à travers les murs, notamment avec les ondes THz radar [11]. Des signaux en bande étroite sont transmis en continu et le signal de retour est traité en permanence, ce qui ne permet d'extraire que la vitesse de la cible par effet Doppler.

3.4 Imagerie Tomographique par Radio (ITR).

Tout comme les techniques radar précédemment décrites, l'ITR est une technique de localisation ne nécessitant pas le placement d'un capteur sur l'élément cible à localiser. À partir de l'atténuation des signaux sur les différents récepteurs, un tableau 2D représentant l'environnement maillé est défini, et l'algorithme ITR détermine les valeurs associées à chaque cellule [12] et en déduit la position de la personne à localiser.

3.5 Simultaneous Localization And Mapping (SLAM).

L'algorithme SLAM, très répandu pour les applications de localisation de robot et de véhicule autonome, consiste à *i*) construire une carte de l'environnement et *ii*) se localiser dans cette carte [17], simultanément. L'algorithme opère alors de manière itérative en quatre temps. À l'initialisation (*étape 1*), l'algorithme se réfère à une estimation (à l'instant précédent $k - 1$) de la position du mobile dans la carte reconstruite à $(k - 1)$. Lors d'un déplacement (*étape 2*), les nouvelles positions du mobile et des amers de l'environnement sont prédites à partir des nouvelles données proprioceptives (issus des capteurs inertiels, encodeurs, etc.). À l'*étape 3*, une carte locale de l'environnement est construite à partir des données extéroceptives (images de caméra par exemple), et à partir de la mise en correspondance entre cette carte et la carte prédite par les données proprioceptives, l'algorithme calcule les erreurs et donc la correction nécessaire à apporter aux estimations. Enfin, l'estimation des positions du mobile et des amers à l'étape 2 est améliorée en appliquant la correction déterminée dans l'étape 3. Les éventuels nouveaux amers détectés à l'étape 3 sont également corrigés et intégrés dans la carte estimée.

3.6 Fingerprinting.

La méthode du fingerprinting consiste à discrétiser, dans une étape préliminaire (*phase hors ligne*), l'environnement en M cellules où doit s'opérer la localisation. Pour chaque cellule, une empreinte du signal reçu (RSS, temps de vol, angle d'arrivée, etc.) qui est issu de chacun des N émetteurs, est stockée dans une base de données (BD) de dimension $M \times N$. Un inconvénient du fingerprinting est la construction de la BD, d'où l'apparition récente de méthodes d'interpolation [28] et de méthodes de construction participative de BD [29].

Dans l'étape de localisation (*phase en ligne*), l'empreinte mesurée par le mobile est comparée à celles de la BD via une métrique de similarité. Parmi les algorithmes de fingerprinting, on recense les algorithmes déterministes tels que *k-Nearest Neighbors*, les algorithmes probabilistes basés sur l'inférence bayésienne, ou encore les réseaux neuronaux [30].

3.7 Algorithmes de poursuite : filtres Kalman et particuliers.

Les algorithmes de poursuite [31] permettent d'estimer et de suivre la position d'un mobile de manière dynamique en tenant compte des mesures – bruitées – et des trajectoires et mouvements aux instants précédents. Le filtre de Kalman simple est adapté pour des problèmes pouvant se modéliser sous forme linéaire et soumis à du bruit gaussien. Il consiste à estimer récursivement l'état courant (la position actuelle du mobile) à partir de l'estimation de l'état précédent (position précédente estimée) et des mesures actuelles. Dans une première étape, la prédiction de l'état actuel et de la covariance de l'erreur de son estimation est effectuée à partir de l'équation d'évolution entre les instants $(k - 1)$ et (k) exprimée sous forme matricielle. Dans la seconde étape, la valeur de l'état estimée et de la covariance de l'erreur d'état est mise à jour à partir de l'équation de mesure reliant l'état à la mesure (données inertielle et/ou données GPS par exemple) par de simples opérations matricielles. Des extensions du filtre Kalman (EKF, UKF) s'appliquent dans le cas où le problème n'est pas linéaire. Alternativement aux filtres de Kalman, les filtres particuliers, basés sur les techniques de Monte-Carlo séquentielles, permettent de couvrir

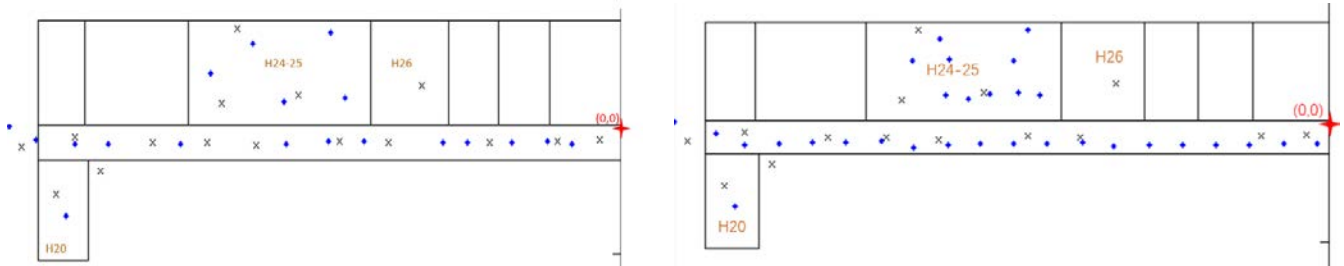


FIGURE 1 – Vue aérienne des deux campagnes de mesures dans le bâtiment IDA. Les croix noires matérialisent les points de références (PR), les losanges bleus sont les points de mesures (PM) à localiser.

Campagne de mesure	$n^{\circ}1$	$n^{\circ}2$	$n^{\circ}3$
Nombre de Points de Référence (PR)	19	31	15
Nombre de points de mesure (PM)	16	15	15
Echantillonnage spatial $\Delta\ell$	$1, 19 \leq \Delta\ell \leq 5,02$	$\Delta\ell \leq 1,5$	$1, 19 \leq \Delta\ell \leq 5,02$
Étage	Rez-de-chaussée	Rez-de-chaussée	Sous-sol

TABLE 1 – Scénarios de mesures.

les cas non pris en compte par les filtres de Kalman (linéarisation non optimale des filtres EKF et UKF, bruits non gaussiens). La démarche des filtres particuliers consiste à rechercher la meilleure estimation de la densité de probabilité des processus aléatoires en jeu dans le système. Considérant un modèle de Markov caché, le filtre estime séquentiellement les valeurs des états cachés à l’instant k à partir des valeurs d’observation précédentes de $n = 0$ à k . La distribution *a posteriori* de l’état à l’instant k sachant les mesures précédentes est estimée par simulations de Monte Carlo.

4 Étude paramétrique d’une méthode Fingerprinting basée sur le WiFi.

4.1 Méthode appliquée.

La méthode de fingerprinting WiFi proposée repose sur une BD (ou carte de référence) construite à partir de statistiques en termes de puissance. L’environnement est échantillonné spatialement sur N_{PR} positions de références. Pendant la phase en ligne, l’équipement souhaitant se localiser mesure les RSS des différents signaux WiFi reçus et les compare avec celles de la BD via une métrique de similarité (ou *norme*) probabiliste. La méthode de calcul des *k-Nearest Neighbors* (kNN ou NNSS-AVG [30]) est ensuite appliquée pour estimer la position de l’équipement, en calculant le barycentre des k plus proches voisins, chacun étant pondéré par l’inverse de sa norme de similarité avec la mesure RSS de l’équipement.

4.2 Description des scénarios et du protocole expérimental.

Les cartes de référence ont été mesurées à l’aide d’un ordinateur portable grâce à un logiciel développé pour cette expérimentation. Ce dernier permet de réaliser l’acquisition et la sauvegarde de mesures (100 mesures successives au même positionnement) et des métadonnées (nom du réseau WiFi, adresse MAC, canal, fréquence). Plusieurs campagnes de mesures ont été réalisées, sur trois étages, au sein d’un bâtiment du laboratoire SATIE à l’ENS de Paris-Saclay (site de Cachan), présentées sur la figure 1. Le positionnement de chaque point de référence, dans le référentiel du bâtiment, est réalisé à l’aide d’un télémètre laser ayant une précision de mesure inférieure au centimètre. Le tableau 1 résume les caractéristiques des campagnes de mesures.

4.3 Étude paramétrique des performances expérimentales du fingerprinting WiFi.

Afin de déterminer les k plus proches voisins dans l’espace signal (*i.e.* espace des empreintes RSS), différentes normes ont été proposées et testées, numérotées de $L1$ à $L4$ (voir système d’équation (1)).

	Norme	L_1	L_2	L_3	L_4
Campagne 1 (RDC)	Erreur moyenne (m)	4.9	3.5	2.5	2.9
	Ecart-type (m)	4.3	2.8	1.7	1.8
Campagne 2 (sous-sol)	Erreur moyenne (m)	3.0	4.4	2.2	3.2
	Ecart-type (m)	4.4	4.5	1.8	2.4

TABLE 2 – Comparaison des performances d’estimation en fonction de la norme utilisée (dans le cas où $1.19 \text{ m} \leq \Delta\ell \leq 5.02 \text{ m}$).

Echantillonnage spatial $\Delta\ell$ (en m)	$\Delta\ell \leq 5$	$\Delta\ell \leq 1,5$
Erreur moyenne (m)	2.5	2.0
Ecart-type (m)	1.7	0.9

TABLE 3 – Comparaison des performances de localisation pour différents pas d’échantillonnage $\Delta\ell$ (norme L_3).

$$\left\{ \begin{array}{l}
L_1 = \frac{1}{N_{AP}} \sum_{n=1}^{N_{AP}} (P_{mes}(n) - P_{ref}(n))^2 \\
L_2 = \frac{1}{N_{AP}} \sum_{n=1}^{N_{AP}} \left(\frac{P_{mes}(n) - P_{ref}(n)}{P_{ref}(n)} \right)^2 \\
L_3 = \frac{1}{N_{AP}} \sum_{n=1}^{N_{AP}} \frac{(P_{mes}(n) - P_{ref}(n))^2}{P_{mes}(n) P_{ref}(n)} \\
L_4 = \frac{1}{M} \sum_{i \neq j} \max \left(\frac{P_{mes}(i)/P_{mes}(j)}{P_{ref}(i)/P_{ref}(j)}, \frac{P_{ref}(i)/P_{ref}(j)}{P_{mes}(i)/P_{mes}(j)} \right)
\end{array} \right. \quad (1)$$

La norme L_1 est la norme euclidienne, fréquemment exploitée dans le cas d’un positionnement par fingerprinting. moyennée sur les N_{AP} réseaux en commun entre un PR ($P_{ref}(k)$) de la carte de référence et le PM à localiser ($P_{mes}(k)$). Les puissances $P_{ref}(k)$ et $P_{mes}(k)$ sont exprimées en Watt.

La métrique L_1 étant non normalisée, elle attribue à tort de fortes valeurs d’erreur à des signaux de forte puissance, relativement à des signaux de moindre puissance. La métrique L_2 introduit une normalisation en divisant par $P_{ref}^2(k)$, permettant la prise en compte de ces fortes différence de puissance. La norme L_3 normalise la différence de puissance par le produit $P_{ref}(k) \times P_{mes}(n)$. La norme L_4 a été introduite pour tenir compte de récepteurs WiFi d’un type différent lors de l’établissement de la carte de référence et des mesures : tablette, smartphone, objets connectés, *etc.* Dans ce cas, ce sont les rapports de puissance entre les réseaux captés qui sont exploités. Ceux-ci sont pris par paire (i, j) avec M le nombre total de paires possibles de réseaux en commun entre PR et PM.

Les tableaux 2 et 3 synthétisent les résultats de localisation obtenus expérimentalement, respectivement en fonction de la norme utilisée et du pas d’échantillonnage spatial pour la cartographie de référence.

Concernant l’influence de la norme utilisée, comme il était attendu, les mesures de similarité obtenues avec les métriques L_3 et L_4 permettent une meilleure précision de localisation grâce à la normalisation introduite.

Le pas d’échantillonnage spatial $\Delta\ell$ avec lequel sont construites les cartes de références influence la qualité des résultats. La précision de localisation est ainsi meilleure avec un pas d’échantillonnage spatial plus resserré. Toutefois, l’amélioration des performances s’infléchit à partir d’un pas de l’ordre du mètre.

Concernant k , le nombre de plus proches voisins, l’analyse des résultats expérimentaux conduit à retenir un nombre fixe $k_{opt} = 3$. Une méthode de détermination dynamique de k_{opt} a été testée, à partir la distribution statistique des valeurs de la norme sur les N_{PR} points de références candidats, en adaptant la technique de détermination automatique de seuil en traitement d’image - pour des distributions unimodales - proposée par Rosin dans [32]. Les résultats de localisation ont été très légèrement améliorés pour notre étude de cas, mais la détermination dynamique de k_{opt} reste recommandée dans le cas général, au lieu de fixer un paramètre de manière empirique.

Enfin, toutes ces expérimentations ont été effectuées plusieurs fois, afin de déterminer la variabilité dans le temps des cartes de référence. Les résultats obtenus ne montrent pas de variations significatives, certainement parce que l’intervalle de temps (1 mois) entre ces mesures n’était pas suffisant.

5 Conclusions

La navigation terrestre par géolocalisation est maintenant d'un usage quasi-quotidien. Basés essentiellement sur l'exploitation des signaux des réseaux GNSS, ainsi que des signaux des réseaux de télécommunication, ces systèmes ne sont plus opérants dans les bâtiments du fait de la disparition des données GNSS.

Dans cette contribution, nous avons recensé divers signaux auxquels il serait possible de recourir afin de se géolocaliser dans un bâtiment. Nous avons également établi une liste des stratégies et des traitements du signal applicables afin d'estimer une position et une trajectoire à partir de ces données.

Nous avons ensuite appliqué la méthode du *fingerprinting* aux données de puissance des réseaux WiFi (RSS). Pour cela nous avons conduit plusieurs campagnes de mesures dans un bâtiment. D'une part nous avons montré que la méthode du fingerprinting était sensible à la métrique utilisée. Deux métriques parmi les quatre testées ont été distinguées. D'autre part, nous avons évalué expérimentalement la sensibilité de cette méthode vis à vis du pas d'échantillonnage spatial avec lequel était établie la carte de référence. Des performances en terme d'erreur moyenne et d'écart-type de positionnement ont été données.

L'amélioration des performances relevées passe par l'exploitation de signaux issus d'autres capteurs : capteurs inertiels, magnétomètre, capteurs optique et caméra, *etc.* Pour cela nous consolidons actuellement nos premiers résultats grâce à un filtrage de Kalman.

Une modélisation de la puissance électromagnétique rayonnée à partir d'une borne WiFi est également étudiée. Il s'agit de la méthode DPSM (Distributed Point Source Method) développée par Dominique Placko [33], méthode semi-analytique, déjà mise en oeuvre dans des problèmes de contrôle non destructif, d'acoustique, de mécanique des fluides et d'électromagnétisme. Cette méthode pourrait être une alternative au modèle de Motley et Keenan [34]. Le caractère distribué de la méthode DPSM permet une modélisation plus fine des affaiblissements rencontrés. La formulation de cette méthode s'écrivant sous forme d'un problème d'algèbre linéaire, la cartographie des puissances des bornes WiFi peut être aisément obtenue et remise à jour dans le cas d'une maintenance de type Simultaneous Localization And Mapping (SLAM) [35].

6 Références

- [1] D. Lymberopoulos and J. Liu, "The Microsoft Indoor Localization Competition : Experiences and Lessons Learned," *IEEE Signal Processing Magazine*, vol. 34, pp. 125–140, Sept 2017. 1
- [2] P. Millot et al., "An UWB Through-The-Wall radar with 3D imaging, detection and tracking capabilities," in *2015 European Radar Conference (EuRAD)*, pp. 237–240, Sept 2015. 1
- [3] G. Deak, K. Curran, and J. Condell, "A survey of active and passive indoor localisation systems," *Computer Communications*, vol. 35, no. 16, pp. 1939–1954, 2012. 1
- [4] I. Vin, D. P. Gaillot, P. Laly, M. Liénard, and P. Degauque, "Radio science for connecting humans to information systems / L'homme connecté," *Comptes rendus - Physique*, vol. 16, no. 9, pp. 862–873, 2015. 1
- [5] X. Ye, X. Yin, X. Cai, A. P. Yuste, and H. Xu, "Neural-Network-Assisted UE Localization Using Radio-Channel Fingerprints in LTE Networks," *IEEE Access*, vol. 5, pp. 12071–12087, 2017. 2.1
- [6] K. McDermott, R. M. Vaghefi, and R. M. Buehrer, "Cooperative UTDOA Positioning in LTE Cellular Systems," in *2015 IEEE Globecom Workshops (GC Wkshps)*, pp. 1–6, Dec 2015. 2.1
- [7] F. Adib, Z. Kabelac, and D. Katabi, "Multi-Person Localization via RF Body Reflections," in *12th USENIX Symposium on Networked Systems Design and Implementation (NSDI 15)*, (Oakland, CA), pp. 279–292, USENIX Association, 2015. 2.1
- [8] H. J. P. Iglesias, V. Barral, and C. J. Escudero, "Indoor person localization system through RSSI Bluetooth fingerprinting," in *2012 19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 40–43, April 2012. 2.1
- [9] P. K. Kumar and T. K. kumar, "UWB Impulse Radar for Through-The-Wall Imaging," *International Journal of Electromagnetics and Applications*, vol. 1, no. 1, pp. 19–23, 2011. 2.1
- [10] Y. Wang, Q. Liu, and A. E. Fathy, "Simultaneous localization and respiration detection of multiple people using low cost UWB biometric pulse Doppler radar sensor," in *2012 IEEE/MTT-S International Microwave Symposium Digest*, pp. 1–3, June 2012. 2.1
- [11] K. B. Cooper et al., "THz Imaging Radar for Stair¹²⁶ Personnel Screening," *IEEE Transactions on Terahertz Science and Technology*, vol. 1, pp. 169–182, Sept 2011. 2.1, 3.3

- [12] J. Wilson and N. Patwari, "Radio Tomographic Imaging with Wireless Networks," *IEEE Transactions on Mobile Computing*, vol. 9, pp. 621–632, May 2010. 2.1, 3.4
- [13] A. Abdelgawad, "Localization system for indoor robot using RFID," in *2014 IEEE Symposium on Industrial Electronics Applications (ISIEA)*, pp. 157–160, Sept 2014. 2.1
- [14] J. Zhao and Y. Wang, "Autonomous Ultrasonic Indoor Tracking System," in *2008 IEEE International Symposium on Parallel and Distributed Processing with Applications*, pp. 532–539, Dec 2008. 2.2
- [15] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan, "The Cricket Location-support System," in *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking, MobiCom '00*, (New York, NY, USA), pp. 32–43, ACM, 2000. 2.2
- [16] A. L. Cambridge, *The Active Badge System*. <http://www.cl.cam.ac.uk/research/dtg/attachive/ab.html>, 2002. 2.3
- [17] N. Mostofi, M. Elhabiby, and N. El-Sheimy, "Indoor localization and mapping using camera and inertial measurement unit (IMU)," in *2014 IEEE/ION Position, Location and Navigation Symposium - PLANS 2014*, pp. 1329–1335, May 2014. 2.3, 2.6, 3.5
- [18] S. Ma, Q. Liu, and P. C. Y. Sheu, "Foglight : Visible Light-Enabled Indoor Localization System for Low-Power IoT Devices," *IEEE Internet of Things Journal*, vol. 5, pp. 175–185, Feb 2018. 2.3
- [19] S. Vidas, P. Moghadam, and S. Sridharan, "Real-Time Mobile 3D Temperature Mapping," *IEEE Sensors Journal*, vol. 15, pp. 1145–1152, Feb 2015. 2.4
- [20] Polhemus, *Polhemus*. www.polhemus.com, 2005. 2.5
- [21] M. S. Reynolds, *Low Frequency Indoor Radiolocation*. PhD thesis, Massachusetts Institute of Technology, 2002. 2.5
- [22] J. Li, M. Guo, and S. Li, "An Indoor Localization System by Fusing Smartphone Inertial Sensors and Bluetooth Low Energy Beacons," in *2017 2nd ICFST*, pp. 317–321, 2017. 2.6
- [23] Y. Wang, X. Wang, D. Wang, and D. P. Agrawal, "Range-Free Localization Using Expected Hop Progress in Wireless Sensor Networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 20, pp. 1540–1552, Oct 2009. 3.2
- [24] S. Frattasi and M. Monti, "On the Use of Cooperation to Enhance the Location Estimation Accuracy," in *2006 3rd International Symposium on Wireless Communication Systems*, pp. 545–549, Sept 2006. 3.2
- [25] H. Wymeersch, J. Lien, and M. Z. Win, "Cooperative Localization in Wireless Networks," *Proceedings of the IEEE*, vol. 97, pp. 427–450, Feb 2009. 3.2
- [26] T. Chen, H. Wu, and L. Liu, "A Joint Doppler Frequency Shift and DOA Estimation Algorithm Based on Sparse Representations for Colocated TDM-MIMO Radar," vol. 2014, 11 2014. 3.3
- [27] O. B. Daho, *Radar ULB pour la vision à travers les murs : mise au point d'une chaîne de traitement de l'information d'un radar imageur*. PhD thesis, Université de La Rochelle, 2014. 3.3
- [28] K. Cujia et al., "DactyLoc : A minimally geo-referenced WiFi GSM-fingerprint-based localization method for positioning in urban spaces," in *2012 International Conference on Indoor Positioning and Indoor Navigation, IPIN 2012, Sydney, Australia, November 13-15, 2012*, pp. 1–9, 2012. 3.6
- [29] M. Lee et al., "Elekspot : A Platform for Urban Place Recognition via Crowdsourcing," in *2012 IEEE/IPSJ 12th International Symposium on Applications and the Internet*, pp. 190–195, July 2012. 3.6
- [30] A. Roxin, J. Gaber, M. Wack, and A. Nait-Sidi-Moh, "Survey of Wireless Geolocation Techniques," in *2007 IEEE Globecom Workshops*, pp. 1–9, Nov 2007. 3.6, 4.1
- [31] S. Sand, A. Dammann, and C. Mensing, *Positioning in wireless communications systems*. John Wiley & Sons, 2014. 3.7
- [32] P. L. Rosin, "Unimodal thresholding," *Pattern recognition*, vol. 34, no. 11, pp. 2083–2096, 2001. 4.3
- [33] D. Placko and T. Kundu, *DPSM for modeling engineering problems*. Wiley-Interscience, 2007. 5
- [34] A. J. Motley and J. M. P. Keenan, "Personal communication radio coverage in buildings at 900 MHz and 1700 MHz," *Electronics Letters*, vol. 24, pp. 763–764, June 1988. 5
- [35] J. Tang et al., "Fast Fingerprint Database Maintenance for Indoor Positioning Based on UGV SLAM," vol. 15, pp. 5311–5330, 03 2015. 5

GNSS Integrity Enhancement for Urban Transport Applications by Error Characterization and Fault Detection and Exclusion (FDE)

Renforcement de l'Intégrité GNSS pour les Applications de Transport Urbain par Caractérisation des Erreurs et la Détection et l'Exclusion des Défauts (FDE)

Ni Zhu¹, David Bétaille², Juliette Marais¹, and Marion Berbineau³

¹Univ Lille Nord de France, IFSTTAR, COSYS, LEOST, F-59650 Villeneuve d'Ascq, France, {ni.zhu, juliette.marais}@ifsttar.fr

²IFSTTAR, COSYS, F-44344 Bouguenais, France, david.betaille@ifsttar.fr

³IFSTTAR, COSYS, F-59650 Villeneuve d'Ascq, France, marion.berbineau@ifsttar.fr

Keywords: GNSS, integrity monitoring, FDE, urban environments

Mots-clés: GNSS, la surveillance de l'intégrité, la détection et l'exclusion des défauts (FDE), l'environnement urbain

Abstract:

In the past decades, more and more Global Navigation Satellite Systems (GNSS)-based urban transport applications emerged. Among these applications, the liability critical ones, such as Electronic Toll Collection (ETC) and Pay as you Drive insurance, have high requirements for positioning accuracy as well as integrity since large errors can lead to serious consequences. Yet urban environments present great challenges for GNSS positioning due to the existence of multipath effects and None-Line-of-Sight (NLOS) receptions. This article presents a complete integrity monitoring scheme for urban transport applications. This scheme is realized in several levels. Firstly, measurement errors are better characterized by using weighting models with the help of an Urban Multipath Modeling (UMM). Secondly, several Fault Detection and Exclusion (FDE) methods are applied in order to detect and exclude erroneous measurements. Finally, Horizontal Protection Levels (HPLs) are computed and the probability of Misleading Information (MI) is analyzed.

Résumé:

Au cours des dernières décennies, de plus en plus d'applications de transport urbain basées sur les systèmes de positionnement par satellites (GNSS) ont vu le jour. Des applications exigent une fiabilité critique comme le télépéage basé sur l'utilisation du GNSS, pour lesquelles des erreurs de positionnement peuvent entraîner de graves conséquences. Pourtant, les environnements urbains présentent de grands défis pour le positionnement GNSS en raison de l'existence des trajets multiples et de signaux NLOS (None-Line-of-Sight). Cet article présente un système complet de surveillance de l'intégrité pour les applications de transport urbain. Tout d'abord, les erreurs de mesure sont mieux caractérisées en utilisant des modèles de pondération tirant parti d'un modèle de multi-trajet urbain (UMM). Deuxièmement, différentes méthodes de détection et d'exclusion de défauts (FDE) sont appliquées afin de détecter et d'exclure des mesures erronées. Enfin, les niveaux de protection horizontale (HPL) sont calculés et la probabilité d'événements redoutés (MI) est analysée.

1 Introduction

GNSS integrity is one criteria to evaluate GNSS performance, which was first introduced in the aviation field. It is defined as a measure of trust which can be placed in the correctness of the information supplied by the total system [1]. Since the complexity of urban environments, GNSS performance could be severely degraded if no special measures are taken. At the same time, the algorithms developed for the aerospace domain cannot be introduced directly to the GNSS land applications. This is because a high data redundancy exists in the aviation domain and the hypothesis that only one failure occurs at a time is made, which is not the case for urban users[2]. The main objective of this article is to present a complete integrity monitoring scheme for GNSS-based urban transport applications. With real data collected in urban environments, on the one hand, the measurement errors will be better characterized and reduced by different approaches. On the other hand, different FDE methods will be used in order to detect and exclude the faults. Finally, the HPL is calculated so as to provide a statistical position error bound.

The paper is organized as follows: After the introduction, two weighting models will be introduced: the C/N_0 (carrier-power-to-noise-density ratio) weighting model [3] and a new Hybrid model[4], which combines the C/N_0 , the satellite elevation and an Urban Multipath Model (UMM), which is an improved version of the Urban Trench Model (UTM) described in [5]. In the next section, five different FDE method will be described briefly, including the Subset Test (ST) (also named exhaustive search), the Local Test (LT), the Forward-Backward (FB) Test, the Danish method as well as the Classic Test, [6][7]. Then, several HPLs will be discussed in order to choose the one which can better fit the urban canyons. The next section will present the real GPS data set as well as the evaluation of the accuracy performance with the Weighted Least Square (WLS) estimator compared to the Ordinary Least Square (OLS) estimator. Finally, five FDE algorithms will be applied with each WLS estimator and HPL will be calculated. The integrity performance will be analyzed from different aspects.

2 Error Characterizations: Signal Weighting

GNSS signals are often influenced by different error sources during the transmission in the propagation channels. Distributing different weights to each pseudo-range measurement according to the severity of its contamination is often an economic way to enhance the accuracy of positioning for the commercial GNSS receivers. This approach can be highly efficient if the propagation errors could be properly characterized. Conventional weighting models are often determined by signal transmission delay, such as ionospheric, tropospheric delay, etc, far from considering the real-time local effects, such as the multipath and NLOS reception. In fact, these local effects are predominant in the total GNSS measurement error budget for urban transport applications, since they can lead up to several kilometers' measurement error. In this paper, we will design the weighting matrix according to the real-time signal reception state.

Generally, the linearized GNSS pseudo-range observation equation can be written as:

$$\delta\rho = \mathbf{H} \times \delta x + \epsilon \quad (1)$$

where, $\delta\rho$ is the deviation between the actually measured pseudo-ranges and the predicted noiseless pseudo-ranges with a specified initial user state;

\mathbf{H} denotes the geometry matrix in which describes the satellite-receiver relative position information;

δx represents the offset vector of the user states ($x, y, z, \delta t$) with the initial states;

ϵ denotes the measurement error vector.

The Weighted Least Square (WLS) solution of δx is expressed as:

$$\delta x = [\mathbf{H}^T \mathbf{W} \mathbf{H}]^{-1} \mathbf{H}^T \mathbf{W} \times \delta\rho \quad (2)$$

here, \mathbf{W} represents the measurement weighting matrix.

In fact, the weighting matrix \mathbf{W} is a diagonal matrix, whose diagonal elements are the reciprocal of each variance of measurement as shown in Eq.3. That is to say, measurements with smaller error variances will contribute to larger weights in the estimation. On the contrary, measurements with larger error variances, such as NLOS signals, will get smaller weights or even approximately canceled from the measurement set by a nearly zero weight. Thus, in the following parts, error variance estimation according to different error models will be implemented in this matrix and the navigation solutions will be estimated with the WLS according to Eq. 2.

$$\mathbf{W} = \begin{pmatrix} \frac{1}{\sigma_1^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sigma_m^2} \end{pmatrix} \quad (3)$$

2.1 C/N_0 -based Weighting model

The carrier-power-to-noise-density ration, *i.e.*, C/N_0 represents the ratio of signal power and noise power per unit of bandwidth. It can be seen as an indicator of the signal quality. Several weighting models exist in literature using C/N_0 as criteria[3][8][9]. In this paper we will choose the model in [3], which is firstly build up for geodetic receivers. And the following formula of variance estimation was proposed for all types of receivers:

$$\sigma^2 = a + b \cdot 10^{-0.1 \times C/N_0} \quad (4)$$

where, a (in m^2) and b (in $m^2 Hz$) are model parameters.

The two model parameters a and b strongly relate to the receiver characteristics and signal reception environments. Calibrations are needed according to specific datasets and receivers.

2.2 A new Hybrid model: Contribution of the Urban Multipath Modeling (UMM)

Despite its non-complicated implementation, the C/N_0 weighting model presents some drawbacks especially in dense urban canyons. On the one hand, constructive multipath interference leads to an increase in C/N_0 , while destructive multipath interference leads to a decrease. Thus all signals with high C/N_0 do not possess high quality but they all contribute to a larger weight in the C/N_0 weighting model when calculating a position. Thus, we propose here a new hybrid model which is inspired by [4]. This model takes into consideration not only the C/N_0 , but also the satellite elevation θ and a LOS/NLOS indicator k from the UMM. This model of measurement error variance can be written as:

$$\sigma^2 = k \times \frac{m \cdot 10^{-0.1 \times C/N_0}}{\sin^2(\theta)} \quad (5)$$

where,

$$k = \begin{cases} 1 & \text{if } LOS \\ 10 & \text{if } NLOS \end{cases}$$

m is a multiplier for model calibration, which depends on receivers and environments.

The UMM is an improved version of the Urban Trench Model (UTM) described in [5]. Based on 3D map data, the UTM calculates the mask of satellite visibility and the additional pseudo-range distance due to signal reflections. An important hypothesis of UTM is that, the length of the streets is infinite, where comes from the model name "Urban Trench". In fact, this hypothesis is an approximation of the map geometry which fits well the reality when the vehicle is inside the homogeneous streets which have approximately the same geometric features as an example of a typical urban trench shown in Fig.1a. However, this assumption is no longer true especially in the no-homogeneous streets or the intersections as shown in Fig.1b.

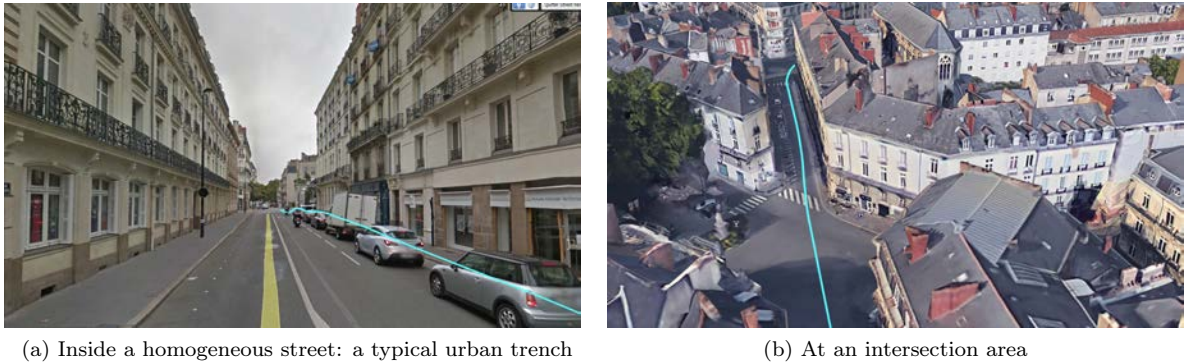


Figure 1 – Example of two kinds of streets

In order to avoid the drawback of the UTM, the UMM is proposed while getting rid of the infinite length street assumption. The process is first to detect if any facade exists in every satellite azimuth, with a height enough to occult the corresponding satellite elevation. This initial step will identify LOS/NLOS satellites. Fig.2 shows an example of this step at a crossroad in the *Chaussée de la Madeleine* of Nantes: Fig.2a is an illustration with the information of digital map, where the polygons represent upside view of buildings. The red lines represent the directions of each received satellite azimuth. Fig.2b is the street view of the same location provided by Google Earth.

The next step is to examine every facade locally, to detect whether it could make a specular reflection with the occulted satellite previously identified. In this case, the final step is to check that no other facade may occult the reflected ray, whether in between the antenna and the impact point, or whether in between the impact point and the satellite. In case several facades exist with reflected rays, the one with the largest incidence angle is preferred. Facades are regarded as vertical planes, with the height of the polygon they belong to. Fig.3 shows an illustration of this step provided respectively by digital map (Fig.3a) and the same location provided by Google Earth (Fig.3b). The green line represents a NLOS satellite signal which is reflected by a facade (in blue) and then received by the user.

Finally, the additional distances will be calculated for all the NLOS receptions according to the geometrical rules of specular reflection. These additional distances will be corrected when applying the hybrid error model.

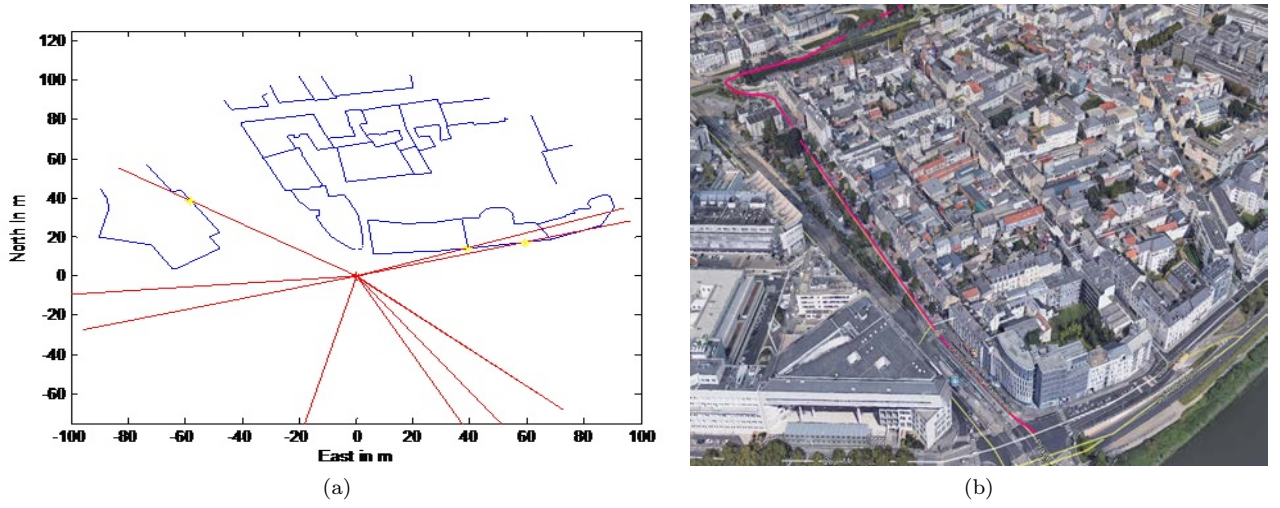


Figure 2 – An illustration of the first step of the UMM

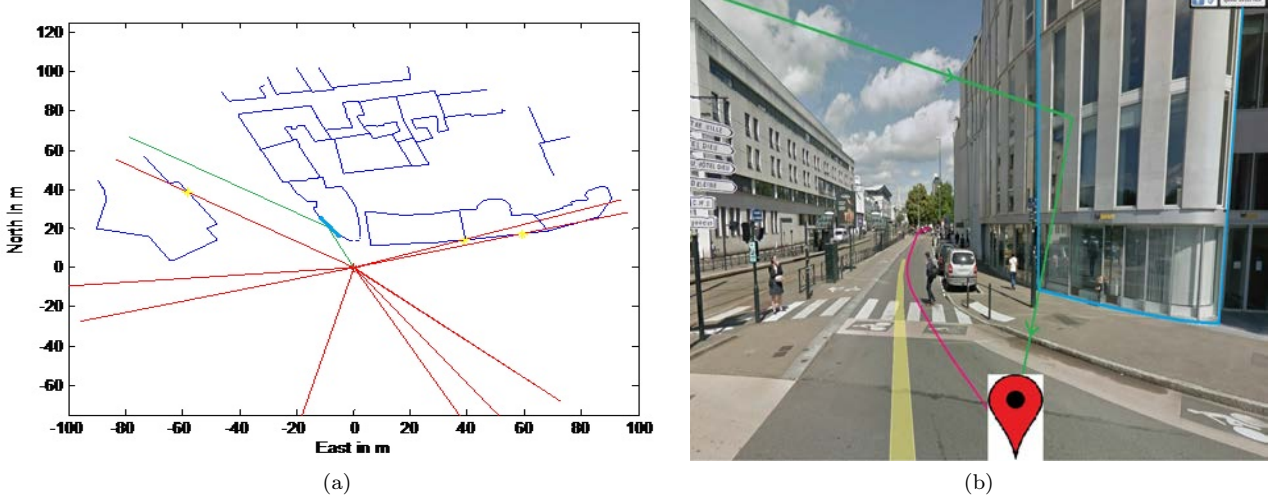


Figure 3 – An illustration of the second step of the UMM

3 Fault Detection and Exclusion (FDE) Algorithms

Fault Detection and Exclusion (FDE) algorithms are generally based on the consistency check of the difference between the measured pseudorange ρ and the predicted pseudorange based on the estimated solution \hat{x} , which is called residuals. The WLS residual vector r can be written as:

$$\hat{r} = \delta\rho - H\delta\hat{x} \quad (6)$$

Two principle statistical hypothesis tests exist in FDE algorithms: the Global Test (GT) and the Local Test (LT). The GT is generally implemented at the primary stage in order to detect whether faulty measurements exist. And if a fault exists, the LT is able to further identify the faulty measurement.

3.1 Global Test

In the framework of the residual-based FDE, the test statistic t used in the GT is called Normalized Sum of Squared Error (NSSE), which can be expressed as:

$$t = \hat{r}^T \Sigma^{-1} \hat{r} \quad (7)$$

where, the matrix Σ denotes Variance Covariance Matrix (VCM) of measurements, which is also the inverse of the weighting matrix \mathbf{W} mentioned previously in Eq.3.

Under fault-free conditions, the test statistic t follows a central χ^2 distribution; under faulty conditions, t follows a noncentral χ^2 distribution [10][11], which is illustrated in Fig.4a. With the pre-defined probability of false alarm P_{fa} and the probability of missed detection P_{md} , a threshold Th can be deduced and the statistical hypothesis test can be conducted as follows:

$$\begin{aligned} H_0 : (\text{fault-free conditions}) & \quad t \leq Th \\ H_a : (\text{faulty conditions}) & \quad t > Th \end{aligned} \quad (8)$$

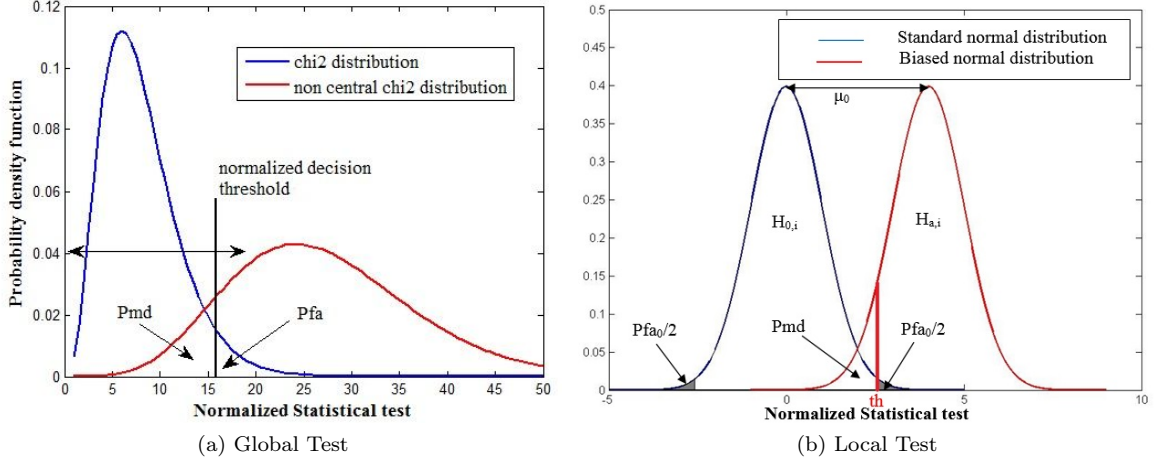


Figure 4 – Global Test and Local Test

If the null hypothesis H_0 is rejected and the alternative hypothesis H_a is accepted, an inconsistency of the total measurement set is detected. Then, the fault identification procedure is needed to isolate and exclude the faulty measurement. The LT is one of methods to realize the fault identification.

3.2 Local Test

If a fault is detected with GT, that means an outlier exist in one or several measurements. LT can be carried out so as to identify the outlier. The LT uses the normalized residuals as test statistic, which can be written as follows:

$$w_i = \left| \frac{\hat{r}_i}{\sqrt{(C_{\hat{r}})_{ii}}} \right| \quad (9)$$

where, $(C_{\hat{r}})_{ii}$ represents the i th diagonal element of the covariance matrix of the residuals $C_{\hat{r}}$. The covariance matrix of residual is calculated as follows:

$$C_{\hat{r}} = \Sigma - H(H^T \Sigma^{-1} H)^{-1} H^T \quad (10)$$

If the i th measurement is not an outlier, w_i is supposed to follow a standard normal distribution, which is the local null hypothesis $H_{0,i}$. Otherwise, w_i follows a biased normal distribution, which corresponds to the local alternative hypothesis $H_{a,i}$ as shown in Fig.4b. In Fig.4b, th denotes the local threshold, which can be calculated with the predetermined local significance level P_{fa0} . As a result, the local test is realize as follows:

$$\begin{aligned} H_{0,i} : (\textit{i}th \text{ measurement not an outlier}) & \quad w_i \leq th \\ H_{a,i} : (\textit{i}th \text{ measurement an outlier}) & \quad w_i > th \end{aligned} \quad (11)$$

This procedure can be repeated several times in a loop until no outlier exists in current measurement set.

3.3 Different FDE Schemes

Different FDE methods exist by combining the GT and LT in different ways. In this paper, we will implement five FDE techniques which allow eliminating multiple faults. Most of them are explained in detail in our previous work [6]. So they will be briefly described here. In case of GT failure, the following FDE techniques can be carried out:

- Subset Testing (ST): each subset (number of satellite > 4) of the initial measurement set will be used to calculate an user position solution, among which the ST with the most satellites and the smallest test statistic which passes the GT will be chosen;
- Sequential Local Test (LT): each measurement will be examined by LT. In each iteration, the measurement with the biggest local test statistic exceeding the threshold will be eliminated as outlier. The procedure stops until no outlier exists or lack of redundancy;
- Forward-Backward (FB) Test: the forward loop will be conducted as the sequential LT described previously. Then in the backward loop, the eliminated satellites in the forward loop will be reintroduced with all the possible combination until the optimal measurement set is found. The main advantage of the FB technique is, on the one hand, to avoid the erroneous rejection of a good measurement since a huge measurement error can sometimes distribute and hide in other measurements' residuals due to special satellite geometry; on the other hand, the effect of re-introduction of previous excluded satellite can enhance the satellite geometry which is usually poor in urban canyon.
- Danish method (DAN): an iteratively reweighting procedure on the pre-estimated measurement variance will be carried out according to the ratio between local test statistic w_i and the local threshold th . What should be highlighted is that there is no exclusion step in this method, which can well keep the initial satellite geometry. If the algorithm cannot converge after several iteration (here, we fix it as 10), the solution will be declared as unreliable.
- The classic method: the measurement with largest normalized residual will be excluded in one iteration. The procedure stops until the GT passes or lack of redundancy.

During all the FDE procedures, if there is not enough redundancy to realize the FDE or the FDE cannot successfully identify the outliers, the position solution will be declared as unreliable.

4 Horizontal Protection Level (HPL) Estimation

Protection Level (PL) is a statistical error bound computed so as to guarantee that the probability of the absolute position error exceeding the said number is smaller than or equal to the target integrity risk[2][12]. Several methods of HPL computation exist in the literature especially for aviation utility. Generally speaking, a 'complete' HPL is composed of two terms: the impact of measurement noise on user position and the impact of measurement bias on user position. The noise term HPL_n is calculated according to the error propagation, which is defined in [12] as the HPL for Satellite-Based Augmentation System (SBAS):

$$HPL_n = K(P_{md}) \times d_{\text{major}} \quad (12)$$

where, d_{major} corresponds to the error uncertainty along the semi-major axis of the error ellipse. K is an inflation factor in order to meet specified integrity risk, which is usually obtained in the corresponding χ^2 table with 4 degree of freedom for conservative purpose. More details can be found in [12][13].

The noise term of HPL in Eq. 12 is often not big enough to bound the position errors especially in the presence of bias in one or several measurements. Thus, the bias term of HPL should be added. The bias term HPL_b generally consists of information of satellite geometry and the detectable bias in the domain of test statistic, which can be expressed as follows:

$$HPL_b(p_{\text{bias}}) = \max_i(SLOPE_i \cdot \sigma_i) \times p_{\text{bias}} \quad (13)$$

where,

$$SLOPE_i = \sqrt{\frac{(\mathbf{H}_{N,i}^+)^2 + (\mathbf{H}_{E,i}^+)^2}{\mathbf{S}_{ii}}} \quad (14)$$

with $\mathbf{H}^+ = (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W}$ and $\mathbf{S} = \mathbf{I} - \mathbf{H} \mathbf{H}^+$;

σ_i represents the standard deviation of the i th measurement;

p_{bias} denotes the bias in the domain of test statistic which can take different values according to the degree of conservative. This will be detailed in the following text.

In fact, $SLOPE$ represents the sensibility of HPE to the bias of i th satellite[13]. That is to say, the satellite with the largest $SLOPE$ is the most difficult to detect because given the same HPE, it yields the smallest test statistic. Also, given a test statistic, the satellite with the highest $SLOPE$ produces the largest position error.

Several variations of HPL_b exist by taking different values of p_{bias} , such as following cases:

- $p_{bias} = \sqrt{Th}$ [13], where Th represents the threshold of the GT which can be determined according to the predefined P_{fa} ;
- $p_{bias} = \sqrt{\delta}$ [14], where δ denotes the non centrality parameter in the non central χ^2 distribution described in Fig.4a;
- $p_{bias} = \sqrt{NSS\bar{E}}$.

Finally, the complete HPL can be obtained as the sum of the noise term and the bias term:

$$HPL = K(P_{md}) \times d_{major} + \max_i(SLOPE_i \cdot \sigma_i) \times p_{bias} \quad (15)$$

The size of HPL is also an important issue for urban transport applications since it will be less effective if the size of HPL is too big compared to the size of current road where situated the user vehicle. Thus, a compromise should be made between the efficiency of the HPL (*i.e.*, HPL well bound the HPE) and its size.

In this paper, we will take $p_{bias} = \sqrt{NSS\bar{E}}$. This is because: firstly, with the same integrity probability specifications and the same degree of freedom, we always have $Th < \delta$ which can be verified from χ^2 tables. Secondly, we can always make the GT pass after FDE so that $NSS\bar{E} \leq Th$. As a result, by choosing $p_{bias} = \sqrt{NSS\bar{E}}$, the HPL size can be well reduced. This is coherent with our results obtained from real data, in which the other two choices of p_{bias} often produce too big HPLs. Such conservative HPLs don't fit well with urban transport applications.

5 Application of the Complete Integrity Monitoring Scheme

5.1 Description of System and Dataset

The complete framework consists of two modules: an accuracy enhancement module and an integrity enhancement module as shown in Fig. 5. In the accuracy enhancement module, measurement errors will be characterized with error models mentioned in section 2. They will be compared with the results obtained from the Ordinary Least Square (OLS) estimator, in which, all the measurements have the same weight. Then, in the integrity enhancement module, five FDE methods mentioned in section 3 will be implemented and evaluated. In this paper, we fix $P_{fa} = P_{md} = 10^{-2}$. Finally, the HPL will be calculated according section 4.

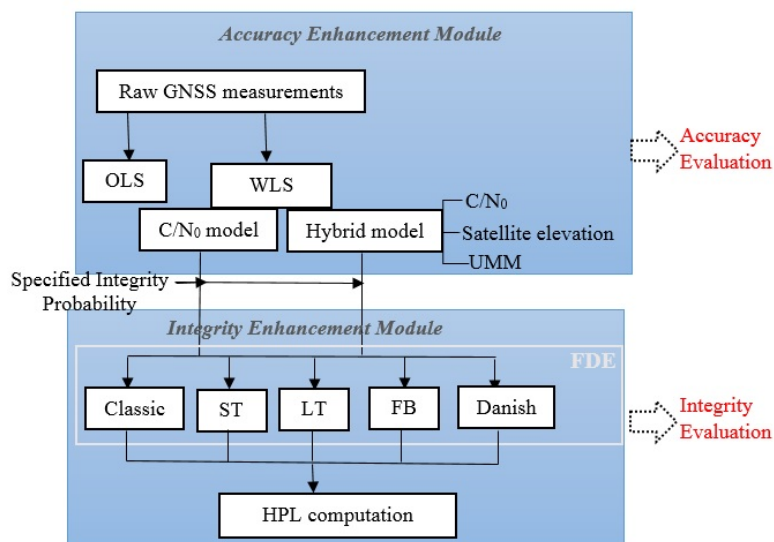


Figure 5 – An overview of the complete integrity monitoring scheme

The GPS data used to test the complete system was collected in the city center of Nantes, France. Two kinds of receivers are used: an Ublox-LEA-6T, which is a typical equipment used in car navigation systems and a dual-frequency Novatel-DLV3 receiver, which is a high precision receiver used to provide the reference trajectory. The total trajectory have 17903 epochs which takes about one hour at $5Hz$. Fig. 6 shows us an overview of this trajectory with Google Earth. We can see that the traveled streets are in deep urban canyons with different heights of buildings.



Figure 6 – An overview of the trajectory in the city center of Nantes, France

	Horizontal Position Error (m)			
	Mean	Median	95%	std
OLS	10.14	5.66	30.74	77.39
WLS C/N_0	4.81	2.64	13.78	26.79
WLS Hybrid	3.61	2.33	10.83	4.43

Table 1 – WLS Accuracy Comparison (without FDE applied)

5.2 Experimental Validation and Performance Analysis

5.2.1 Performance of accuracy

In this sub-section, only the error models are applied (without FDEs). Fig. 7 shows the reference trajectory (in green) together with the trajectories estimated respectively by OLS (in blue), by WLS with C/N_0 model (in red of Fig. 7a) and by WLS with hybrid model (in red of Fig. 7b). Then Fig. 8 presents the Cumulative Distribution Function (CDF) of the HPE. We can see that with the two error models, the accuracy is well improved compared to the OLS, especially with the hybrid model. Table 1 makes a summary about the HPE in terms of mean, median (50th percentile), 95th percentile and standard deviation (std), which is used to evaluate the dispersion of HPE. We can see obviously that the WLS with hybrid model has the best performance concerning the accuracy. The main advantage of the hybrid model compared to the C/N_0 model is the capability of reducing huge errors since the 95th percentile of HPE passes from 13.78m to 10.83m and the std passes from 26.79m to 4.43m.

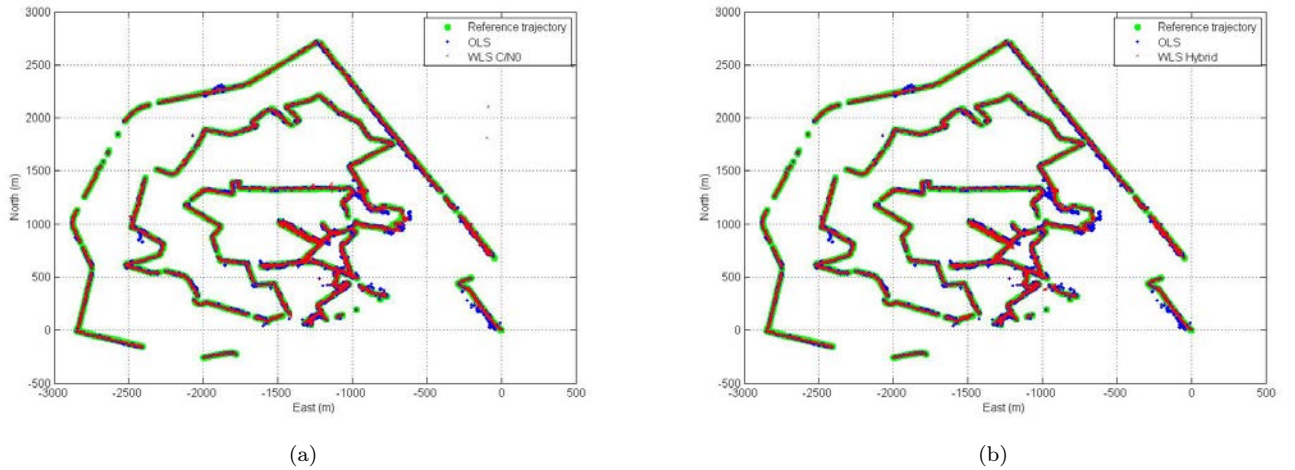


Figure 7 – Comparison of Estimated Trajectories

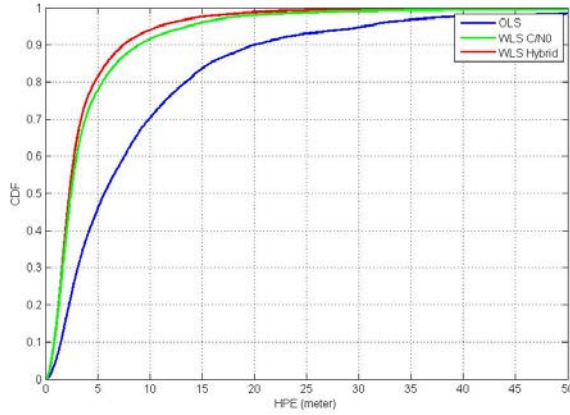


Figure 8 – Cumulative Distribution Function (CDF) of the HPE

	mean (HPE)	median (HPE)	95% HPE	std (HPE)	median(HPL) Pmi
ST	4.61m	2.45m	14.44m	8.02m	18.26m 0.76%
LT	3.56m	2.38m	9.56m	4.41m	17.76m 0.11%
FB	3.43m	2.32m	9.90m	4.36m	17.46m 0.14%
DAN	4.13m	2.50m	11.18m	6.17m	19.21m 0.28%
Classic	4.55m	2.46m	13.61m	8.07m	18.29m 0.75%

Table 2 – Performance summary of C/N_0 model with FDE

5.2.2 Performance of integrity

After the accuracy enhancement, the integrity module is implemented and the HPL is calculated in this section. The main performance evaluated after application of FDE are mainly accuracy, the size of HPL as well as the efficiency of HPL. The last one is assessed by the probability of misleading information P_{mi} , which represents the probability of HPE exceeding HPL being the system declared the position solution is reliable.

Table 2 gives us a summary about the WLS with C/N_0 model with different FDE methods. We can see that the accuracy performance is again improved especially huge errors are removed since the standard deviations of HPE are much reduced. Among all the FDE methods, the FB has the best global performance and Fig. 9 presents some details. We can see in Fig.9a the reference trajectory (in green), the trajectory estimated with C/N_0 WLS (in sky blue) as well as the trajectory estimated with WLS C/N_0 and FB (the reliable positions are in magenta and unreliable positions are in black). Especially in the red circles, we can see that the majority of huge HPEs are either reduced or flagged as unreliable. All the P_{mi} are smaller than 1%.

Table 3 reports the performance of each FDE method with hybrid model. We can see that, with LT, FB and DAN, accuracy can be slightly improved compared to the one with only WLS hybrid model. Unfortunately, with ST and classic method, the accuracy is slightly degraded. Additional measures should be taken in the following research, such as special control of satellite geometry, in order to avoid this degradation. Compared to the performances of total integrity scheme with C/N_0 model, the sizes of HPL calculated with hybrid model are smaller. And all the P_{mi} can be guaranteed smaller than 1%.

As mentioned in Section 3.3, the FDE algorithms are not all the time available. When there is not enough

	mean (HPE)	median (HPE)	95% HPE	std (HPE)	median(HPL) Pmi
ST	4.07m	2.39m	12.59m	6.34m	12.50m 0.95%
LT	3.42m	2.29m	10.59m	3.94m	12.36m 0.42%
FB	3.48m	2.28m	10.76m	4.03m	12.50m 0.48%
DAN	3.55m	2.30m	10.81m	4.67m	13.28m 0.21%
Classic	4.01m	2.39m	11.96m	7.16m	12.45m 0.80%

Table 3 – Performance summary of Hybrid model with FDE

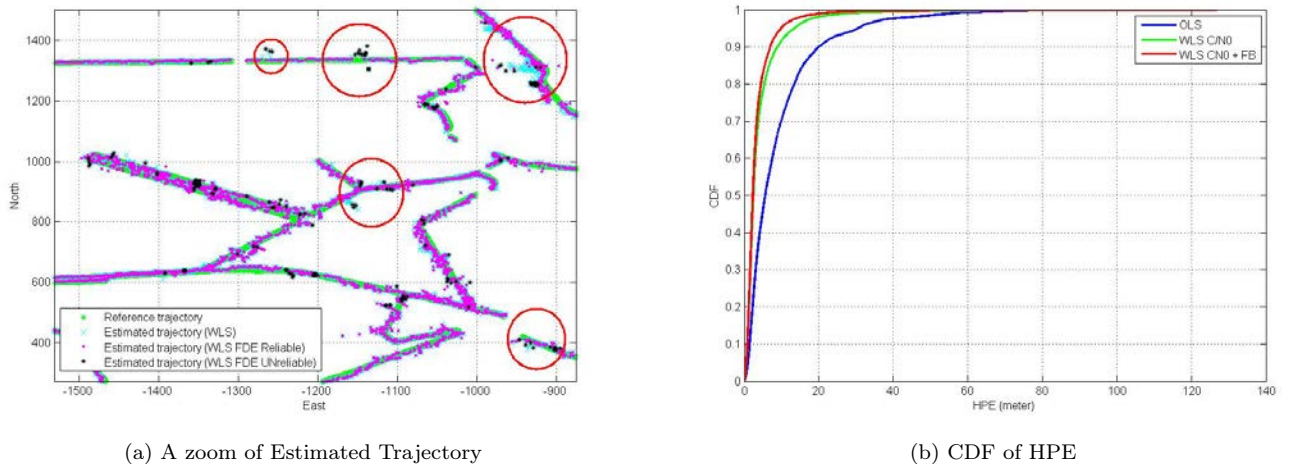


Figure 9 – Performance of WLS C/N_0 with FB

(%)	ST	LT	FB	DAN	Classic
C/N_0	98.58	93.55	95.01	94.81	98.57
Hybrid	98.08	89.53	91.22	96.39	97.98

Table 4 – FDE Availability

redundancy to check the consistency or when the FDE algorithms are not able to identify the outlier, the system will be declared as non available. These positions will be flagged as unreliable. Table 4 reports the availabilities of each scheme. Thus, how to enhance the availability is also an issue for future research.

6 Conclusion

In this paper, we propose a complete integrity monitoring scheme for urban transport applications. This framework consists of two modules: one for accuracy enhancement, inside which, a new error model with the contribution of map information is proposed; the other for integrity enhancement, inside which, five FDE methods are implemented.

The validation with real GPS data collected in urban canyons shows that accuracy can be significantly improved by two error models compared to the classic Ordinary Least Square method especially the new hybrid model. All the FDE algorithms are able to help improve the accuracy of C/N_0 model but for hybrid model, only the Local Test, Forward-Backward Test and the Danish method can improve accuracy while the Subset Testing and the Classic method cannot due to satellite geometry degradation. Thus, special measures may be taken in order to avoid this degradation for future work. The HPL is able to well bound the HPE with P_{mi} smaller than 1%. HPLs still have a median value above 10 m in size (for HPE of 2 m in median), which is an issue to target for future research.

Acknowledgement

This work has been performed with the financial support of CNES and the Hauts de France Region Council in the framework of the SMARTIES project of the CPER ELSAT 2020 program which is co-financed by the European Regional Development Fund, the French state and the Hauts de France Region Council.

References

- [1] P. B. Ober, *Integrity prediction and monitoring of navigation systems*. Integricom Publishers Leiden, 2003, vol. 1.
- [2] N. Zhu, J. Marais, D. Bétaille, and M. Berbineau, “GNSS position integrity in urban environments: A review of literature,” *IEEE Transactions on Intelligent Transportation Systems*, 2018.

- [3] A. Wieser and F. K. Brunner, "An extended weight model for GPS phase observations," *Earth, planets and space*, vol. 52, no. 10, pp. 777–782, 2000.
- [4] S. Tay and J. Marais, "Weighting models for GPS pseudorange observations for land transportation in urban canyons," in *6th European Workshop on GNSS Signals and Signal Processing*, 2013, p. 4p.
- [5] D. Betaille, F. Peyret, and M. Ortiz, "How to enhance accuracy and integrity of satellite positioning for mobility pricing in cities: the Urban Trench method," in *Transport Research Arena TRA 2014*, PARIS, France, Apr. 2014, p. 8p, transport Research Arena TRA 2014, PARIS, FRANCE, 14-/04/2014 - 17/04/2014. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-01206082>
- [6] N. Zhu, J. Marais, D. Betaille, and M. Berbineau, "Evaluation and comparison of GNSS navigation algorithms including FDE for urban transport applications," *Proceedings of ION ITM, Monterey, CA, USA*, 2017.
- [7] H. Kuusniemi, A. Wieser, G. Lachapelle, and J. Takala, "User-level reliability monitoring in urban personal satellite-navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, 2007.
- [8] F. Brunner, H. Hartinger, and L. Troyer, "GPS signal diffraction modelling: the stochastic SIGMA- δ model," *Journal of Geodesy*, vol. 73, no. 5, pp. 259–267, 1999.
- [9] H. Hartinger and F. Brunner, "Variances of GPS phase observations: the SIGMA- δ model," *GPS solutions*, vol. 2, no. 4, pp. 35–43, 1999.
- [10] S. Kuang, *Geodetic network analysis and optimal design: concepts and applications*. Ann Arbor PressInc, 1996.
- [11] A. Leick, L. Rapoport, and D. Tatarnikov, *GPS satellite surveying*. John Wiley & Sons, 2015.
- [12] RTCA/DO-229D, "Minimum Operational Performance Standards for Global Positioning System/Wide Area Augmentation System airborne equipment," *RTCA SC-159*, 2006.
- [13] T. Walter and P. Enge, "Weighted RAIM for precision approach," in *PROCEEDINGS OF ION GPS*, vol. 8. Institute of Navigation, 1995, pp. 1995–2004.
- [14] R. G. Brown, *GPS RAIM: Calculation of Thresholds and Protection Radius Using Chi-square Methods; a Geometric Approach*. Radio Technical Commission for Aeronautics, 1994.

Solution d'intégration de COTS offrant un service de navigation piéton collaboratif fiable, résilient et sécurisé

Integrating COTS to deliver secured and reliable navigation in a pedestrian collaborative radionavigation environment

Authors :

Dominique HEURGUIER (Thales TCS), Grégory GAILLIARD (Thales TCS), Frederique YWANNE (Thales TCS), Alain LEMER (Thales TCS), Goulven EYNARD (DGA MI).

Keywords :

GNSS ; RANGING ; RESILIENCE ; INTEGRITE ; COLLABORATIF

Résumé :

Cet article résume les leçons tirées de la conception d'un démonstrateur visant à prouver la faisabilité d'intégration de composants sur étagère bon marché, pour délivrer une navigation fiable et sécurisée. Destiné à un groupe de piétons interconnectés par un réseau radio MANET¹, ce démonstrateur repose notamment sur une approche de radiolocalisation collaborative. Celle-ci exploite des mesures de télémétries radio entre voisins en compléments des mesures GNSS² disponibles. Parallèlement à l'amélioration de disponibilité du service de navigation, ces travaux pointent la vulnérabilité d'une telle approche face à des malveillances par intrusion et démontrent l'intérêt de compléter les mesures de cyber protection par un contrôle d'intégrité efficace.

Abstract :

This paper sums up lessons learned from a demonstrator design, aiming to prove the integration feasibility of low cost COTS³ components, to provide to the customer enhanced PNT⁴ reliability and security. Intended for a group of pedestrians, interconnected by a MANET radio network, this demonstrator bases in particular on a collaborative radio localization approach. It exploits measures of radio telemetry between neighbors in complements to the available GNSS measures. Complementary to the improvement of availability of the navigation service, these works point out the vulnerability of such approach face of hostilities by intrusion and demonstrate the interest to complete the cyber protection by an efficient integrity monitoring.

1. Introduction

The concept of use of the demonstrator is a small set of collaborative pedestrian nodes, operating completely independently from any telecommunication infrastructures. This last point means no radio cellular network available, no fixed anchors deployed and finally no external GNSS alternative navigation system or dedicated augmentation system available.

It is assumed that certain nodes can be somehow corrupted and used - intentionally or not - to degrade the overall collaborative navigation of each nodes. It is also initially assumed that usual cyber protections, like authentication and cryptography, will not be part of the defense that the system will be able to deploy to protect himself against integrity attacks (in particular against stolen equipment). The objective is to examine the opportunities offered by civilian COTS for collaborative navigation, to identify the vulnerability points of such collaborative navigation, and to study and develop security solutions to protect it. The main degradation considered will be therefore intentional: false navigation data injection coming from one node or local denial of access (could be experienced for example when enduring GNSS jamming).

Precision of the collaborative PNT delivered by the system will not be the prime key performance indicator optimized. Indoor navigation is also deliberately out of the scope of the demonstrator's concept of use.

¹ Mobile Ad hoc NETWORKs

² Global Navigation Satellite System

³ Commercial Off-The-Shelf

⁴ Positioning, Navigation and Timing

Lastly but not least, GPS receivers based on military standards can be integrated in the mock up in order to take potentially into account high-integrity PNT, coming from an external secured and trusted source. Proof of concept of this integration has been done using Polaris receivers.

The paper structure is as follow: on the first part the overall architecture of the demonstrator, and then the overall architecture of each node is presented. The choices of the low cost sensors retained to provide information to the navigation filter are justified. On the second part of the paper, an integrity-based data fusion algorithm is presented. On the third part, the Human Machine Interface of the Android application, and the integration of all the sensors on one node, is presented. Finally, the abstract present some trials results obtained, exhibiting the added-value of integrity-oriented navigation algorithms.

2. Demonstrator architecture

The demonstrator consists of a network of 7 nodes: 5 collaborative “blue” nodes, one “red” node playing the role of an intruder, and one supervisor node, used to define scenario, supervise the situation in real time and record the results for potentially further analysis (Figure 1).

A blue node is made of Android smartphone, one GW5204 embedded board from Gateworks including WiFi card, one ranging module from Nanotron, battery, external antennas and optionally GPS watch and/or a GPS Polaris receiver from Rockwell & Collins (Figure 2).

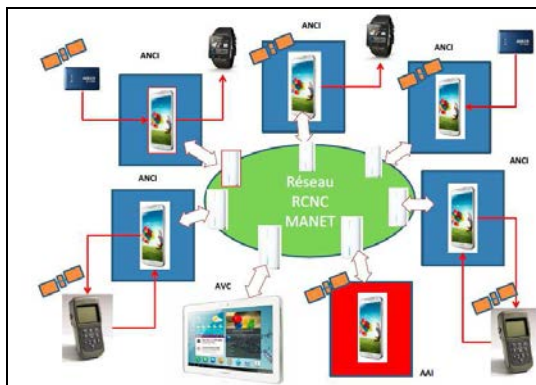


Figure 1: Functional diagram of the collaborative radio navigation mock-up

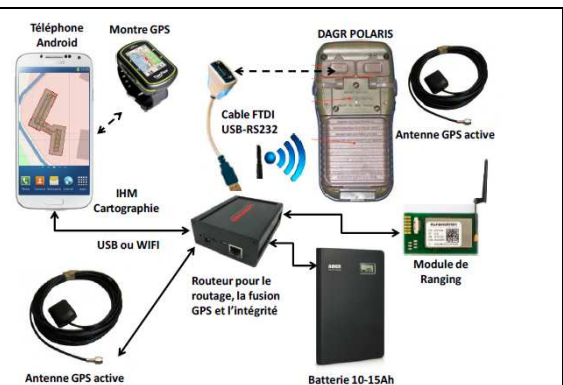


Figure 2: Sensor integration of one node in the collaborative radio navigation

1.1 Embedded navigation sensors

1.1.1 Inertial sensors

The basic use of smartphone embedded MEMS⁵ accelerometer with double integration is not realistic due to the drift of these current sensors. The zero update velocity, usually used with rigid mobile structure or with Inertial Motion Unit (IMU) mounted foot for pedestrian, to reset the drift on zero velocity (i.e. zero when the foot is on the ground), cannot be performed with a smartphone used in the usual way. Nevertheless, trials have confirmed that it is possible to offer a continuity of navigation service to a walking pedestrian for some time, using accelerometer and gyroscope from an accurate initial position and orientation, requiring an initial smartphone calibration phase and an estimate of the stride length of the user. This approach has not been implemented due to the IMU using constraints which are not compatible with usual operational missions (any kind of pedestrian motion, like running, crawling, etc.).

1.1.2 Barometer

It has been experimentally shown that smartphone embedded barometer could be very useful to contribute to differential measurement of altitude. But this has not been implemented due to the focus on 2D only navigation. Nevertheless, this sensor, when present in the smartphone, is used for the integrity monitoring. An abnormal difference of pressure between two nodes, meaning in outdoor an abnormal difference of altitude, is used as a coherence criteria. Abnormal differential pressure leads to alert and data possible rejection.

⁵ MicroElectroMechanical Systems

1.1.3 Inter-node ranging module

Low cost COTS performing inter-node radio ranging now exists and can be integrated into a collaborative navigation system. A promising COTS is the module from Nanotron performing ranging in ISM⁶ band on Chirp Spread Spectrum (CSS) waveform with Symmetric Double-Sided TWR⁷ (SDS_TWR) protocol. First trials have proven that accuracy of about 1 or 2 meters is achievable in outdoor environments. This module may be used in an autonomous way thanks to the swarm mode. It is embedded in the router board and connected to the router in order to deliver ranging measurements to the collaborative navigation data fusion algorithm.

1.2 Mobile Ad-hoc NETWORK (MANET)

Existing open source projects developing Android applications for wireless networking in MANETs like SPAN [1] and SERVAL [2] use rooted devices to address this issue. This solution was not retained for the demonstrator design, in order to keep a requirements guarantee regarding the application portability and modularity, as well as a protection against malware provided by an always up-to-date operating system.

Ad hoc and mesh networking modes are however not natively supported for 802.11 WLAN on non-rooted Android OS. User requests to support Wi-Fi ad hoc networking and mesh networking protocols) on the Android open source project issue tracker were classified as obsolete and respectively closed in April 2015 and December 2014.

Departing from these statements, the chosen solution relies on an external Linux COTS router that supports ad hoc/mesh modes and MANET routing protocols. The module retained is the GW5204 embedded board from Gateworks company. Smartphone and its accompanying router forming a wireless ad-hoc routing node may be connected either by an USB wire interface or by Wi-Fi. No porting or integration with Android is required compared to SPAN and Serval applications on rooted smartphones.



Figure 3: Gateworks router



Figure 4: inter-Ranging Nanotron Module

1.3 Integration of the navigation sensors / communications device on each node

The demonstrator has been integrated in a tactical gilet. Smartphone, Router board and optional Polaris GPS receiver are integrated in dedicated front pockets. Antennas (WiFi, Nanotron, GNSS L1 band for router receiver and GPS for optional Polaris) are positioned on the shoulders. Batteries are placed in the back of the gilet and wires are fully integrated in the vest. A general interrupter is available in the router board pocket.



Figure 5: Integration of a node on a tactical gilet

⁶ Bande « Industriel, Scientifique et Médical”

⁷ Two Way Ranging

1.4 Hybridization and sensor fusion for a non-centralized collaborative navigation

The fusion algorithm developed for the demonstrator is completely decentralized and is implemented on each node. It uses positioning measurements delivered by the local GNSS receiver when available, ranging measurement of local node on nearby nodes, ranging measurement of friends and positions of friends when available (resulting of the fusion algorithm). This fusion algorithm delivers periodic navigation solution (each 2 s) thanks a two steps process:

- Geolocalisation process using local GNSS measurement (when available), ranging measurement and position of friends (when available) to elaborate an estimation of local position by minimizing a Less Mean Square criteria thanks to a Gauss-newton algorithm. Due to strong non linearity of the criteria, an efficient initialization is implemented.
- Tracking process based on Extend Kalman Filter with a Nearly Constant Position motion model, suitable for pedestrian nodes.

This algorithm provides a localization solution of the current node even in lack of GNSS signal, under availability of a sufficient number of ranging measurements (node positioning observability condition).

1.5 Integrity monitoring for secured collaborative navigation

If an intruder is able to disseminate corrupted data in the collaborative navigation network without been detected by the passive security protections of the navigation system (case for example of stolen equipment), the only way to detect it consists to analyze navigation data to find abnormal data or behavior in the current navigation context. That is performed first, by detecting abnormal data compare to a priori knowledge and secondly, by searching potential incoherencies between data themselves (if there is enough redundancy). Some algorithms have been developed to check data likelihood by comparing navigation data to some reference models (thanks to a priori knowledge database) and to detect incoherence in different subsets of data using statistic tests (chi2 typically). If there is enough redundancy between data, these algorithms allow detect integrity problem, identifying corrupted data, corrupted sources or nodes. This process lead finally to detect attack and raise integrity alerts, identify attacking node and reject data coming from corrupted node, offering so reliable collaborative navigation with a good guaranty of the integrity of the localization. An operator may also decide to exclude temporary or permanently a node (if he is for example systematically alerted about corruption of a particular node by the integrity monitoring process). In this case, the navigation data and measurement coming from the excluded node will not be used by the navigation process, neither by the integrity monitoring process.

1.6 Human Machine Interface (Android application)

The Android smartphone is used for its embedded sensors and, mainly, as a Human Machine Interface. Firstly, a local Geographical Information System supports map used by the android application to display Blue Force Tracking data and main characteristics of the different member node of the collaborative network.

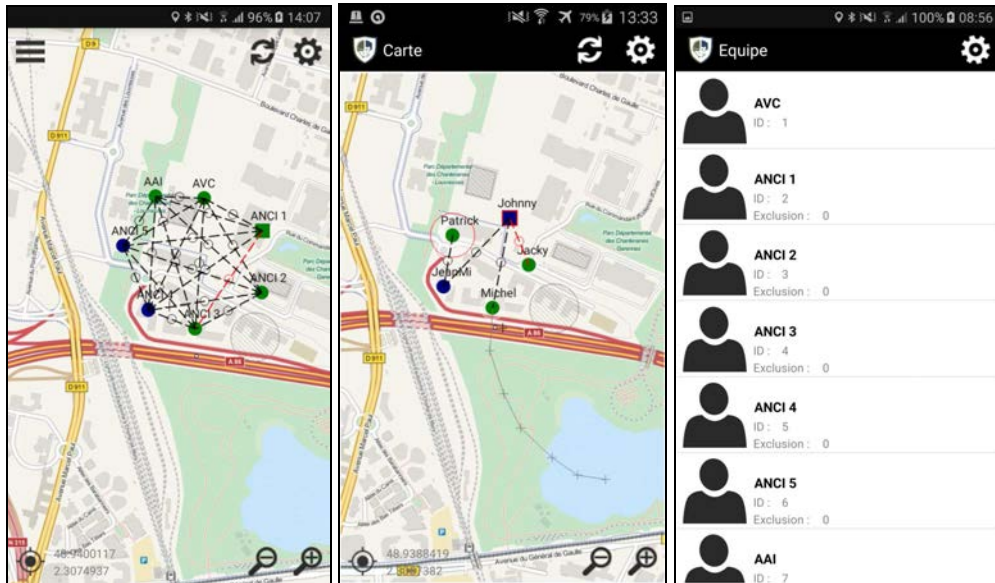


Figure 6: Smartphone HMI

The Human Machine Interface (HMI) of the Android application allows also displaying the measurements: local GNSS data (PVT, NMEA frame and satellite card), ranging and barometer measurements.

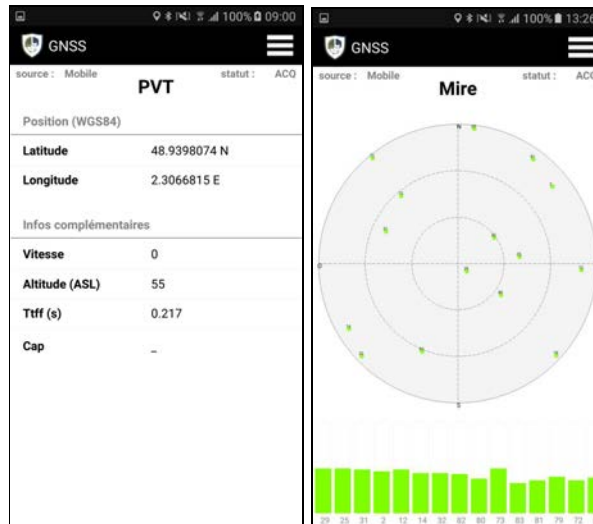


Figure 7: Smartphone HMI for sensors

The Android Human Machine Interface allows also to display status of radio links and the alerts related to the state of the radio network (in blue), the state of GNSS receiver (in green), the integrity monitoring (in orange) and the node exclusion by operators (in red). The operator may also give orders, typically to exclude a node or change the display parameters.

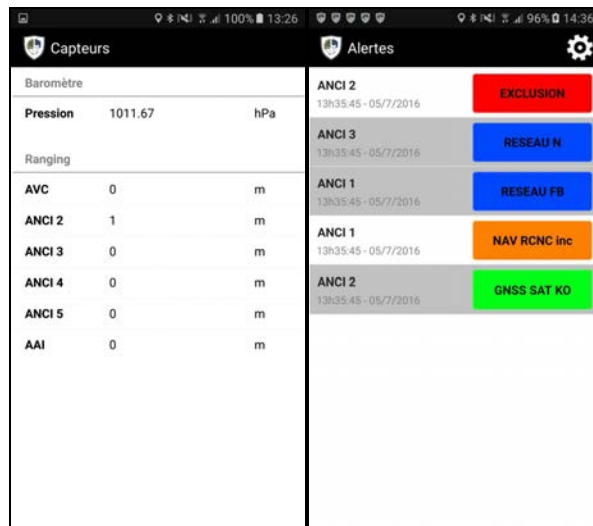


Figure 8: Smartphone HMI for alerts

3. Experimentations and results

Different trials has been performed in an open outdoor area in order to :

- Evaluate the benefit of the collaborative navigation in jamming GNSS context in comparison to the classical autonomous GNSS navigation,
- Evaluate the benefit of the integrity monitoring against integrity attacks in comparison to collaborative navigation without integrity monitoring.

Several fixed point trials have shown the negative effect of a hostile intruder in the collaborative navigation but they have also shown the capability to detect, identify and reject the intruder when the collaborative navigation is coupled to an integrity monitoring.

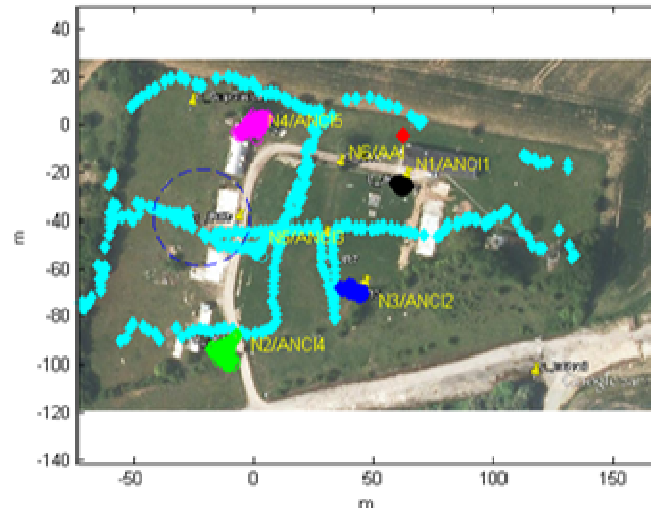


Figure 9: Free run trial

This free-run trial, with a moving GNSS jammed node, has clearly shown the capability to maintain a positioning and navigation with the hybridization and sensor fusion of the collaborative approach.

4. Conclusions

Experience gained during the design of the collaborative navigation demonstrator shows that generic non-rooted Android smartphones have today some limitations for an implementation relying on a non-centralized collaborative outdoor pedestrian navigation.

Some of these limitation rose during the design of the demonstrator was the absence of modular MANET chipset, the limited precision of the possible Wi-Fi ranging, or of the incorporated GNSS sensor. These limitations can however be overcome by the use of additional external COTS. At the end, it appears that today the smartphone mainly constitutes a very good Man-Machine Interface, and that secured navigation has to be delivered by auxiliary components.

Trials made with the demonstrator shows that potential vulnerability of collaborative navigation against integrity attacks could be efficiently treated by integrity monitoring algorithms. The demonstrator provides a first example of a possible implementation of a secured non centralized collaborative navigation.

References

- [1] Thomas J. and Robble J, "Off-Grid Communication with Android – Meshing the mobile world", The MITRE Corporation, 2012
- [2] Gardner-Stephen, P.; Challans, R.; Lakeman, J.; Bettison, A.; Gardner-Stephen, D.; Lloyd, M., "The serval mesh: A platform for resilient communications in disaster & crisis," IEEE Global Humanitarian Technology Conference (GHTC), 2013, pp.162,166, 20-23 Oct. 2013

Système autonome orienté données :
Du signal à l'information
Autonomous system for data collection: from signal to information

Tullio Josph Tanzi¹, Jean Isnard²

¹ Institut Mines-Télécom-Télécom ParisTech, Univ. Paris-Saclay, LCTI. France,
tullio.tanzi@telecom-paristech.fr

² URSI France, jisnard-isti@club-internet.fr

*Mots clés (en français et en anglais) : Capteurs embarqués, système autonome, temps réel.
Embedded Sensors, Autonomous system, Real-Time*

Proposition de communications

Le besoin en systèmes de mesure fonctionnant en totale autonomie existe depuis longtemps, mais ne pouvait être couvert jusqu'à présent que par des enregistreurs interrogeables sur place. Les données acquises n'étaient de ce fait jamais actualisées et ne donnaient qu'un instantané de l'environnement exploré plutôt qu'une surveillance permanente de celui-ci. L'utilisation de réseaux de téléphonie mobile et d'Internet permet aujourd'hui d'associer des systèmes de mesure et d'acquisition variés et distribués et de procéder à l'exploitation en temps réel des données relevées [1]. Mais ces systèmes demandent le déploiement d'infrastructures fixes et en état de fonctionner.

L'arrivée de véhicules autonomes intégrant des capteurs de technologies modernes (LIDAR, Radar, vidéo, etc.) [2,3] ouvre de nouvelles perspectives dans ce domaine. Ces équipements constitueront des moyens de plus en plus sophistiqués pour acquérir des informations de toute nature afin d'explorer un environnement donné. L'utilisation de capteurs couplée à un système autonome permet de réaliser une mission sans intervention externe. Cette approche permet de s'affranchir des contraintes d'infrastructure existantes ou de communication difficile [4] comme par exemple des environnements présentant des infrastructures en panne ou détruites telles que les situations faisant suite à des accidents ou des catastrophes. Les informations captées auront une finalité duale : d'une part, comprendre et modéliser l'environnement pour une bonne réalisation de la mission et d'autre part, réutiliser cette modélisation dans un cadre plus large d'aide à la décision par les équipes de secours dans le cadre d'une catastrophe [5,6].

Les données traitées seront produites par un ensemble de capteurs variés et déployés en réseaux pour la collecte en temps réel [7-9]. Les données produites sont de divers types : données de distance obtenues par des capteurs ultrasonores (temps de vol), données de distance obtenues par des capteurs optiques laser (LIDAR), données de position et d'attitude produites par les systèmes inertiels (accéléromètres, magnétomètres, gyromètres, etc.), données d'odométrie, données d'environnement telles que la température, la pression, etc. Ces données seront combinées pour détecter différents éléments de l'environnement.

Les difficultés résident d'abord dans l'utilisation de processeurs de basse puissance (capacité de calcul faible) pour des raisons d'optimisation énergétique [12,13] dans un contexte de système embarqué critique et de prise en compte de l'incertitude des données [11]. Cette difficulté intervient aussi bien en ce qui concerne la réduction des données acquises, leur traitement, que leur protection par des mécanismes logiques et cryptographiques [10].

Une autre question intéressante porte sur l'utilisation des données acquises : dans le cas où ces données doivent être protégées par des techniques cryptographiques, comment les rendre utilisables par le système autonome pour sa propre navigation tout en les protégeant contre des fuites intempestives ?

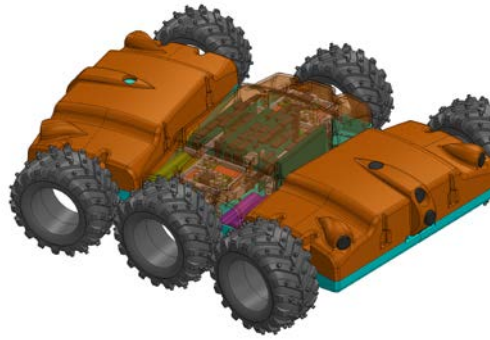


Figure 1 : Module Arcturius, destiné à la recherche de personnes enfouies après un séisme

La contribution présente la genèse et l'état actuel du prototype. Des scénarios de terrain ont été développés pour une évaluation dans un contexte opérationnel. Des résultats pourront être présentés selon le niveau d'avancement de ces essais.

Bibliographie

1. Tanzi, Tullio Joseph; Isnard, Jean. (2015). "Introduction to Public safety networks" Chapter of book, In Public safety Book volume 1: Overview and Challenges. Published by Wiley-ISTE. 2015.
2. Madhu Chandra and Tullio Joseph Tanzi, (2017). "Drone-Borne GPR Design: Propagation Issues," in Journées scientifiques de l'URSI-France (JS'17)
3. Sylvie Servigne, Yann Gripay, Ozgun Pinarer, John Samuel, Atay Ozgovde & Jacques Jay, (2016). "Heterogeneous Sensor Data Exploration and Sustainable Declarative Monitoring Architecture: Application to Smart Building". First International Conference on Smart Data and Smart Cities, 30th UDMS, 9 septembre 2016, Split (Croatie), pp. 97-104, doi : 10.5194/isprs-annals-IV-4-W1-97-2016.
4. Tanzi, Tullio Joseph; Roudier, Yves; Apvrille, Ludovic. (2015). "Towards a new architecture for autonomous data collection". ISPRS Geospatial Week 2015: Workshop on civil Unmanned Aerial Vehicles for geospatial data acquisition. 1-2 October, 2015. La Grande Motte (Montpellier), France.
5. Ludovic Apvrille, Tullio Joseph Tanzi, Yves Roudier, and Jean-Luc Dugelay. (2017) "Drone "humanitaire : état de l'art et réflexions," Revue Française de Photogrammétrie et de Télédétection, pp. 63-71, N 213-04-26. 2017. ISSN 1768-9791
6. Tullio Joseph Tanzi, Madhu Chandra, Jean Isnard, Daniel Camara, Olivier Sebastien, Fanilo Harivelo. (2016) "Towards rone-Borne Disaster Management: Future Application Scenarios ". 06/2016; III-8:181-189. DOI:10.5194/isprs-annals-III-8-181-2016
7. Jane W. S. Liu, (2000). "Real-time Systems", Prentice Hall
8. Arnold S. Berger. (2002). "Embedded Systems Design: An Introduction to Processes, Tools, and Techniques" CMP Books. 237 pages, ISBN: 1578200733
9. Sandeep K. Shukla, IIT Kanpur, (2016). "Introduction to Embedded Systems "TEQIP Course at IITK, Aug 29, 2016
10. Yves Roudier and Tullio Joseph Tanzi, (2017). "A State of the Art of Drone (In)Security," in Journées scientifiques de l'URSI-France (JS'17)
11. Claudia Gutierrez Rodriguez & Sylvie Servigne (2013). "Managing Sensor Data Uncertainty: a data quality approach". International Journal of Agricultural and Environmental Information Systems (IJAEIS), vol. 4, pp. 35-54.
12. Ozgun Pinarer, Yann Gripay, Sylvie Servigne, Ozgovde Atay & Atilla Baskurt (2017). "Dynamic energy-aware sensor configuration in multi-application monitoring systems". Pervasive and Mobile Computing, vol. 41, pp. 192-204
13. Ozgun Pinarer, Yann Gripay, Sylvie Servigne & Atay Ozgovde. (2016). "Energy Enhancement of Multi-application Monitoring Systems for Smart Buildings". CAiSE'16: Conference on Advanced Information Systems Engineering - EnBIS: Energy-awareness and Big Data Management in Information Systems, 17 juin 2016, Ljubljana (Slovénie), pp 131-142. doi : 10.1007/978-3-319-39564-7_14

Synchronisation et calibrage entre un Lidar 3D et une centrale inertielle pour la localisation précise d'un véhicule autonome

Synchronization and calibration between a 3D Lidar and an inertial measurement unit for the accurate localization of an autonomous vehicle

Pierre MERRIAUX¹, Remi BOUTTEAU¹, Romain ROSSI¹, Guillaume CORU¹, Vincent VAUCHEY¹, and Xavier SAVATIER¹

¹Institut de Recherche en Systèmes Electroniques Embarqués, Normandie Univ, UNIROUEN, ESIGELEC, IR-SEEM, 76000 Rouen, France, {merriaux,boutteau,rossi,coru,vauchey,savatier}@esigelec.fr

Mots-clés: Lidar, centrale inertielle, localisation, synchronisation, calibrage, véhicule autonome

Keywords: Lidar, IMU, localization, synchronization, calibration, autonomous vehicle

Résumé:

La télédétection par laser (Lidar) est une technologie de plus en plus utilisée en particulier dans les fonctions de perception et localisation nécessaires à la conduite autonome. L'acquisition des données Lidar doit être couplée à la mesure du mouvement du véhicule par une centrale inertielle. Ces capteurs n'étant pas conçus pour fonctionner ensemble nativement, il est nécessaire de maîtriser leur synchronisation et leur calibrage géométrique. Cet article présente une méthode pour caractériser les décalages temporels entre un Lidar 3D et une centrale inertielle. Il explique aussi comment mettre en œuvre les méthodes de la littérature pour le calcul de la pose entre centrale inertielle et Lidar sur un véhicule utilisé en conditions réelles.

Abstract: Laser remote sensing (Lidar) is a technology increasingly used especially in the perception layers of autonomous vehicles. As the vehicle moves during measurement, Lidar data must be referenced in a fixed frame which is usually done thanks to an inertial measurement unit (IMU). However, these sensors are not designed to work together natively thus it is necessary to synchronize and calibrate them carefully. This article presents a method for characterizing timing offsets between a 3D Lidar and an inertial measurement unit. It also explains how to implement the usual methods for pose estimation between an IMU and a Lidar when using such sensors in real conditions.

1 Introduction

La télédétection par laser (Lidar) est une technologie qui s'est massivement diffusée ces dernières années, en particulier dans le domaine du véhicule autonome [1, 2]. De nombreux travaux et expérimentations exploitant des Lidars, initialement 2D et maintenant 3D, ont montré l'apport de ce type de capteur pour deux fonctions essentielles du véhicule autonome : la génération de cartes haute définition afin de localiser précisément le véhicule [3, 4] et la détection des obstacles autour du véhicule. La mesure par un Lidar embarqué sur un véhicule impose de coupler l'acquisition à une observation du mouvement propre à partir d'une centrale inertielle. La précision de la mesure 3D – et par extension la localisation du véhicule et des obstacles – est liée non seulement aux erreurs de mesure du Lidar et de la centrale inertielle mais aussi :

- à la qualité de la synchronisation entre le Lidar et la centrale inertielle, le véhicule se déplaçant en même temps que l'acquisition des mesures ;
- à l'estimation précise de la transformation géométrique entre le Lidar et la centrale inertielle de manière à rejeter les points Lidars dans un repère géoréférencé.

Ces deux points sont peu traités dans la littérature bien qu'ils soient des éléments critiques de la performance globale du système de conduite autonome. Dans cet article nous présentons dans un premier temps une méthode expérimentale permettant de qualifier la synchronisation entre un Lidar 3D et une centrale inertielle. Cette démarche nous a permis de valider la datation de chaque impact Lidar par rapport à la date de la centrale inertielle. Dans un second temps, nous abordons la mise en œuvre de deux méthodes de calibrage géométrique entre centrale inertielle et Lidar 3D souvent citées dans la littérature mais rarement détaillées. Nous présentons la méthodologie, les bonnes pratiques pour leur implémentation et des résultats de calibrage issus de mesures sur véhicule en situation réelle.



(a) Dispositif placé sur un véhicule routier



(b) Dispositif placé sur petit robot-
véhicule Segway RMP440

FIGURE 1 – Coffre de toit de référence installé sur deux types de véhicules compatibles du laboratoire.

Mode	GPS RTK	60s sans GPS
True Heading (°)	0.01	0.01
Roulis/Tangage (°)	0.005	0.005
Position X,Y (m)	0.02	0.1
Position Z (m)	0.05	0.07

TABLE 1 – Précision de la localisation du coffre de toit

2 Dispositif de mesure par Lidar 3D embarqué sur véhicule

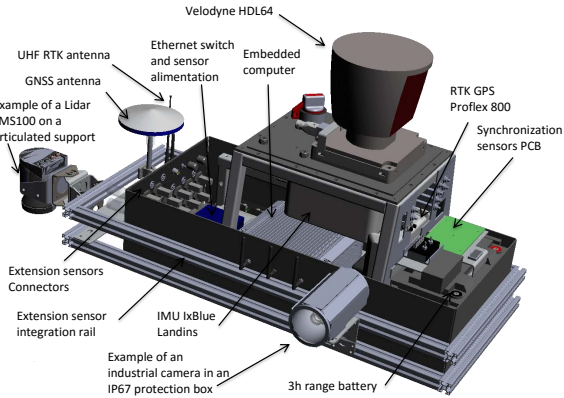
Dans le cadre de projets de recherche sur le véhicule autonome, nous avons développé un équipement embarqué se présentant comme un coffre de toit qui a été instrumenté pour permettre le recueil des données de perception 3D autour du véhicule. Ce dispositif est présenté sur la figure 1 et peut être installé sur le toit d'un véhicule routier ou sur petit robot-véhicule. Une présentation détaillée est montrée sur la figure 2a. La perception 3D de la scène est obtenue en équipant le coffre avec un Lidar 64 nappes (Velodyne HDL64) comme on le voit sur la figure 1 ou un Lidar 16 nappes (Velodyne VLP16); ces deux Lidars sont actuellement les plus communément utilisés. Le positionnement centimétrique (cf. table 1) peut être atteint grâce à un GPS-RTK (Proflex 800) couplé à une centrale inertielle de haute précision (Landins). Une baie d'acquisition avec synchronisation matérielle permet l'enregistrement de données de capteurs additionnels (caméras comme montré sur la figure 1 ou tout autre type de capteur de perception). Les données sont enregistrées par un PC, l'horodatage des flux étant assuré grâce au logiciel RT-MAPS d'Intempora. Le système a une autonomie de 3H. Ce dispositif a été conçu pour recueillir les données de référence nécessaire à la mise au point d'algorithmes pour le véhicule autonome, principalement la position précise du véhicule et la représentation 3D précise de la scène autour du véhicule à chaque instant. Si la précision du Lidar et de la centrale inertielle autorisent en principe une reconstruction 3D centimétrique de la scène, cette résolution est fortement dépendante de mesure de la transformation géométrique entre les repères de mesure du Lidar et de la centrale. Ces paramètres sont estimés dans une phase dite de calibrage extrinsèque. D'autre part, le véhicule étant en mouvement et le Lidar en rotation pendant toute son acquisition, la synchronisation des deux capteurs doit être la plus parfaite possible pour limiter les erreurs de mesure. Ces deux points - calibrage et synchronisation - sont l'objet de cet article et sont abordés dans les paragraphes qui suivent.

3 Synchronisation d'un Lidar avec une centrale inertielle

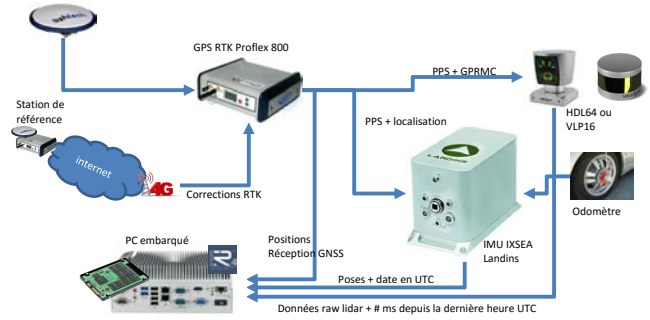
3.1 Influence sur la perception Lidar

Une reconstruction 3D consiste à projeter chaque point Lidar mesurés avec la position fournie par la centrale inertielle. Les fréquences de rotations des Lidars (Velodyne) sont comprises entre 5 et 20Hz. Prenons l'hypothèse d'un véhicule se déplaçant à 55km (15m/s) et un Lidar à 10Hz, si les acquisitions sont totalement asynchrones, la mesure est entachée d'une erreur allant jusqu'à un scan complet, soit 1.5m d'erreur. Les effets d'une mauvaise synchronisation sont encore plus marqués dans le cas d'une rotation du véhicule. Les points mesurés à une distance de 50m du Lidar lors d'une rotation de 25°/s comportent une erreur de 2.19m.

La synchronisation des données est donc essentielle pour obtenir des reconstructions 3D de qualité. Du fait



(a) Instruments contenus dans le coffre de toit de référence. La localisation est assurée par la fusion d'IMU à fog, d'un GPS RTK et d'un odomètre. Un ensemble de capteurs peut être ajouté sur les rails d'extension en fonction des expérimentations.



(b) Synoptique des la partie localisation du coffre de toit : Une station de référence installée sur le toit de notre laboratoire fournit les corrections RTK par le réseau 4G au GPS embarqué. La centrale inertielle Landins assure la fusion avec les GPS et l'odomètre. Le PC embarqué est chargé d'enregistrer l'ensemble des capteurs grâce au framework RTMaps.

FIGURE 2 – Coffre de toit instrumenté développé par l'IRSEEM. Ce moyen d'essai constitue notre référence pour les expérimentations externes sur véhicule, aussi bien pour la validation de la précision d'algorithmes que pour l'acquisition de données géo-référencées.

que les Lidars tournent mécaniquement pour réaliser l'acquisition d'un scan complet pendant que le véhicule se déplace, nous ne pouvons pas nous contenter d'une synchronisation pour chaque scan, il faut chercher à obtenir la date de chaque impact Lidar.

Dans un premier temps nous allons présenter le matériel et le logiciel utilisés dans notre coffre de toit permettant d'assurer cette synchronisation. Puis nous décrirons l'expérimentation mise au point pour valider cette synchronisation. Enfin, nous présenterons les résultats obtenus à la fois en simulation et sur des données réelles.

3.2 Description de l'acquisition des données

Comme décrit au paragraphe 2, la partie localisation du référence du coffre de toit est constituée d'une centrale inertielle qui fusionne un GPS RTK et un odomètre. Le GPS RTK reçoit ses corrections d'une station de base installée sur le toit du laboratoire par le réseau 4G. La centrale obtient la position du GPS par une liaison série et un signal de synchronisation de type Pulse Par Seconde (PPS).

Les Lidars Velodyne HDL64e et VLP16, renvoient les données sous forme de paquets Ethernet ; chaque paquet est constitué de 12 blocs de données appelés *fire blocks* contenant eux-mêmes 32 tirs laser. Chaque paquet Ethernet comprend un champ *timeStamp*, qui représente le nombre de microsecondes depuis l'heure courante en temps universel coordonné (UTC). Pour cela le Lidar a besoin de connaître l'heure UTC. Cette information peut être reconstituée en lisant l'heure UTC fourni dans les trames du GPS (champ GPRMC de la trame NMEA) via la liaison série et en se recalant sur le signal de synchronisation PPS. Il faut se référer à la documentation du fabricant afin de respecter scrupuleusement le calage temporel entre le signal de synchronisation et le champ GPRMC fournissant l'heure. Ces différents échanges de données sont résumés sur la figure 2b.

D'un point de vue logiciel, le décodage des trames Ethernet est assuré par plusieurs composants comme décrit dans la Figure 4. Le Lidar nous fournit des données exprimées dans un repère sphérique (distance, azimut et hauteur du point mesuré).

- Décodage des trames UDP : conversion de chaque paquets Ethernet en 384 distances, 12 angles de rotation et une datation faite matériellement par le capteur (*hardware timeStamp* TS), plus précisément le nombre de μs depuis l'heure UTC courante.
- Conversion sphérique en cartésien : on utilise les données de calibrage intrinsèque du Lidar pour convertir les impacts dans un repère cartésien XYZ, chaque point valide étant associé à un *TimeStamp*. Comme nous n'avons qu'un *TimeStamp* pour pour chaque paquet Ethernet, Velodyne fournit une table de correction de délais des tirs Lidar pour chaque position du *fire block*.
- Correction 3D de scan : le véhicule se déplaçant en même temps que le lidar fait ses acquisitions, les nuages de points résultants sont déformés par les mouvements du véhicule. Pour corriger ces déformations, nous avons étendu une méthode développée au laboratoire [5] aux mouvements 3D. Ce composant recalcule également les *TimeStamps* UTC en μs .

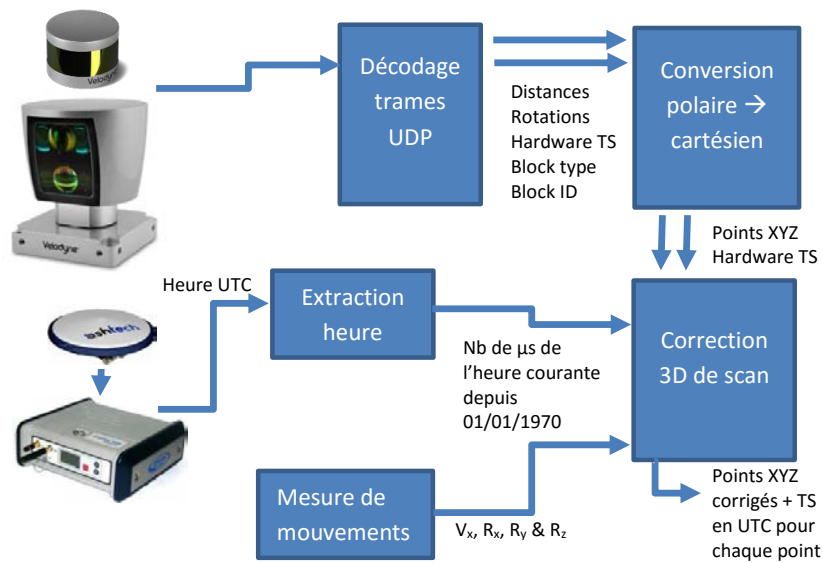


FIGURE 3 – Architecture des composants logiciels RTMaps développés pour l’acquisition et la datation des points Lidar

- Extraction heure : composant permettant d’extraire l’heure et la date courante depuis les trames RMC du GPS.
- Mesure de mouvements : le composant de Correction 3D des scans nécessite une mesure du mouvement du véhicule, dans notre cas la vitesse longitudinale V_x ainsi que les trois vitesses de rotation R_x , R_y et R_z extraites à partir de la centrale inertielle.

3.3 Méthode mise en place pour la validation de la datation des impacts Lidar

Que ce soit d’un point de vue matériel ou logiciel, le système de synchronisation au complet est complexe. Dans cette section nous présentons la méthodologie que nous avons utilisée pour valider toutes les parties de la synchronisation.

L’expérimentation mise au point se décompose en deux étapes :

- la vérification de la précision et de la gigue (*jitter*) du signal PPS, puisque la qualité de notre mesure dépend de celui-ci,
- la validation effective de la datation des impacts Lidar.

Pour valider la précision et la gigue du signal PPS, nous avons choisi d’observer celui-ci à l’aide d’un oscilloscope. Cependant, la précision de l’horloge de référence de l’oscilloscope ne permet pas de mesurer une durée longue (ici une seconde) avec certitude. Pour pallier cette limitation, nous avons choisi de synchroniser l’appareil (SIGILENT SDS1102X) sur l’horloge de référence interne d’un synthétiseur de fréquence (R&S SMB100A). L’erreur de fréquence de l’horloge de référence du synthétiseur étant inférieure à 10^{-7} , nous obtenons une erreur de mesure maximale de 100ns sur une période de 1s.

La figure 4 décrit l’architecture de l’expérimentation mise au point afin de valider la date des impacts Lidar. Nous utilisons une cible plane de petite dimension de manière à n’avoir qu’une nappe qui vient l’échantillonner. La cible est équipée d’une photodiode en son centre permettant de détecter l’impact laser du Lidar. Le signal de la photodiode est acquis par l’oscilloscope (SIGILENT SDS1102X) dont la sortie de *trigger* est envoyée dans une entrée *event Marker* de la centrale inertielle prévue pour dater des événements matériels (fronts montants ou descendants d’un signal) en heure UTC.

Nous obtenons d’un côté par notre chaîne décrite en section 3.2 la date des impacts Lidar dans la cible et d’un autre côté la date de l’impulsion de la photodiode. Le Lidar et la centrale inertielle étant synchronisés par le même signal PPS, nous devrions avoir une différence la plus faible possible. Afin de rendre le système automatisable, nous avons ajouté un traitement logiciel pour n’extraire que les impacts sur la cible des centaines de milliers de points par seconde acquis par le Lidar et fournir une moyenne des *TimeStamps* des points ayant impactés la cible.

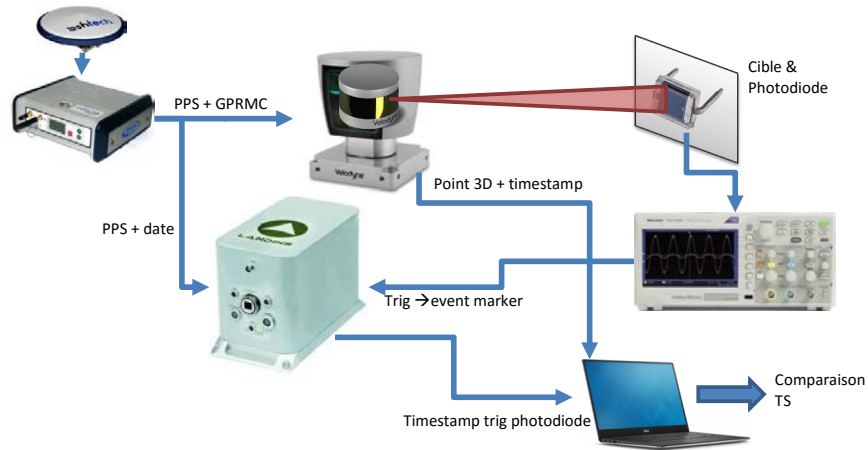


FIGURE 4 – Synoptique de l'expérimentation de validation du datage des points Lidar. Les dates des points Lidars impactant la cible sont comparées à la date de l'impulsion de reçu sur la photodiode.

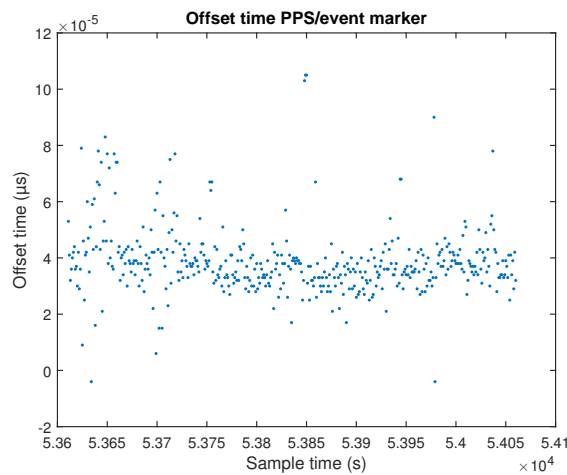


FIGURE 5 – Erreur de datation de l'entrée event marker de la centrale inertielle IXSEA Landins

3.4 Résultats et discussion

La précision de la méthodologie de validation de la synchronisation (section 3.3) peut être affecté par deux points : le *jitter* du signal *PPS* issu du GPS et la précision de datation de l'entrée *event Marker* de la centrale. Pour ce qui est du *jitter* du signal *PPS*, nous l'avons mesuré avec l'oscilloscope et le synthétiseur de fréquence comme référence de fréquence externe. Le *jitter* est inférieur à la précision de mesure de l'appareil, soit 100ns. Pour la précision de la datation de l'entrée *event marker*, nous avons directement entré le signal *PPS* dans l'entrée *event marker*. Comme la centrale fait une datation en UTC, nous obtenons directement l'erreur de datation de l'entrée *event marker*. Ces résultats sont présentés sur la figure 4. L'erreur moyenne est de $38.9\mu s$ et l'écart-type de $12.4\mu s$. Ces valeurs permettent de déterminer le bruit systématique que nous retrouverons dans l'étude de la synchronisation Lidar/Centrale Inertielle.

La figure 6 et le tableau 2 présente les résultats finaux de cette étude. Pour le VLP16, sachant que l'erreur de datation moyenne de l'entrée *event marker* est de $38.9\mu s$, une synchronisation à $8.14\mu s$ est très correcte. Par contre, pour le HDL64, malgré nos efforts pour identifier un éventuel problème nous n'avons pas trouvé d'où venait la source d'erreur qui faisait que le Lidar était en avance d'environ un *fire block* sur la centrale inertielle. Nous avons donc fini par nous résoudre à ajouter logiciellement un offset de $288\mu s$, d'où la présentation d'un résultat brut et corrigé sur la figure 6b.

	VLP16	HDL64 brut	HDL64 corrigé
Moyenne	8.14	-248.7	39.3
Ecart type	23.3	32.2	32.2

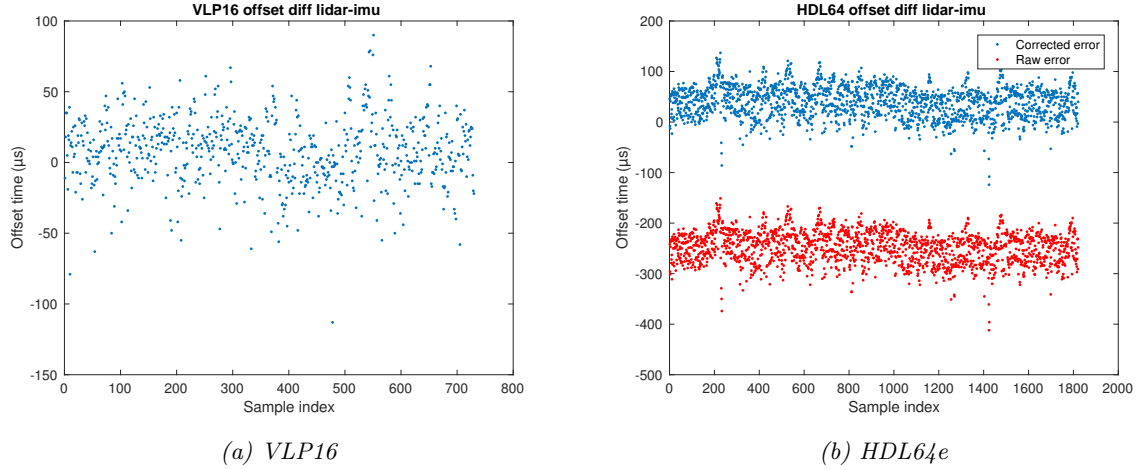
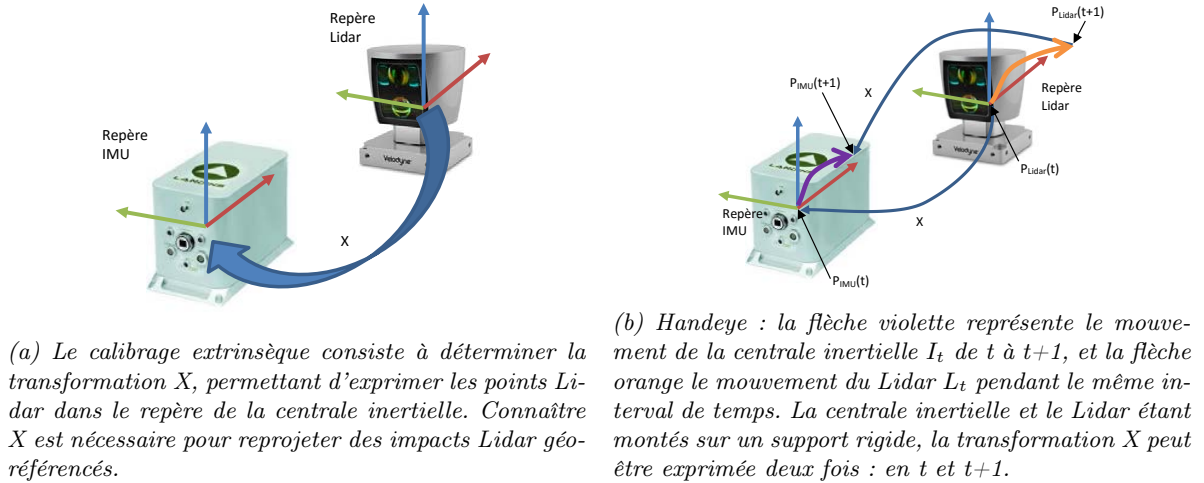


FIGURE 6 – Différence de timestamp entre la moyenne des points 3D sur la cible et le front de la photodiode.



(a) Le calibrage extrinsèque consiste à déterminer la transformation X , permettant d'exprimer les points Lidar dans le repère de la centrale inertielle. Connaître X est nécessaire pour reprojeter des impacts Lidar géoréférencés.

(b) Handeye : la flèche violette représente le mouvement de la centrale inertielle I_t de t à $t+1$, et la flèche orange le mouvement du Lidar L_t pendant le même interval de temps. La centrale inertielle et le Lidar étant montés sur un support rigide, la transformation X peut être exprimée deux fois : en t et $t+1$.

FIGURE 7 – Calibrage extrinsèque Lidar/IMU par la méthode Handeye.

Nous avons déplacé la cible tout autour du lidar, nos résultats sont invariants au secteur angulaire. En reprenant l'exemple de la section 3.1, soit un véhicule se déplaçant à une vitesse linéaire de 15m/s, l'erreur moyenne passe de 1.5m à 0.12mm dans le cas du VLP16 et à 0.35mm pour le HDL64. Pour une vitesse de rotation de $25^\circ/s$, l'erreur de projection d'un point mesuré à 50m du Lidar est réduite de 2.19m à 0.18mm pour le VLP16 et de 0.51mm pour le HDL64. Notre méthode nous a permis de corriger pour le HDL64 une erreur de front sur le signal PPS du Lidar, et un décalage temporel de la trame RMC pour la bonne interprétation de l'heure par le Lidar, ainsi qu'une erreur de calcul de l'offset des tirs Lidar dans le composant *conversion sphérique vers cartésien* qui impactait les deux équipements (VLP16 et HDL64e).

4 Calibrage géométrique entre le Lidar et la centrale inertielle

4.1 Présentation du problème

Le calibrage extrinsèque consiste à estimer la pose relative entre le Lidar et l'IMU (cf Figure 7). La connaissance de cette transformation géométrique est indispensable pour fusionner les points 3D issus du Lidar entre deux poses. Les poses sont en effet fournies par la centrale inertielle, mais celle-ci n'a pas le même référentiel que le Lidar car ces deux éléments ne sont pas géométriquement confondus.

4.2 Revue des principales méthodes

Le problème de calibrage entre deux capteurs fournissant des poses dans des repères différents est connu sous le nom de *Handeye calibration*. Ce problème est apparu initialement dans la communauté robotique où une caméra ("eye") était montée sur l'extrémité du bras d'un robot industriel ("hand"). La caméra était calibrée en utilisant une mire de calibrage, il était par conséquent possible d'avoir la pose de la caméra par rapport à cette mire.

La pose de l'extrémité du bras dans le référentiel du robot était également connue en fonction des commandes envoyées à ce bras manipulateur. Il fallait par conséquent estimer simultanément les transformations entre la caméra et l'extrémité du bras, ainsi que la transformation entre les deux repères de référence (le repère associé à la mire et le repère associé au robot). Il y a eu de nombreuses recherches sur le sujet [6] [7] [8], et la plupart des méthodes proposées décomposent généralement le problème en deux étapes : la rotation est estimée dans un premier temps, puis la translation est optimisée.

Les solutions génériques au problème *Handeye* ne fonctionnent pas toujours dans le cas du véhicule se déplaçant dans un plan. Cela est dû au fait que la hauteur de la caméra et certains angles (roulis, tangage) sont dans ce cas inobservables. Des solutions spécifiques aux robots mobiles ou aux véhicules ont donc été proposées. Dans [9], les paramètres extrinsèques sont estimés pour un robot différentiel, cependant il est nécessaire de placer des amers connus dans la scène. Dans [10], ces amers ne sont pas nécessaires mais seulement trois degrés de liberté sont estimés. Nous avons choisi de travailler avec une solution générique, mais nous faisons attention lors de nos expérimentations à réaliser des trajectoires faisant varier tous les degrés de liberté, par exemple en montant deux des roues sur un trottoir ou un talus.

4.3 Implémentations de la méthode Handeye

Le calibrage extrinsèque par la méthode *Handeye* est basé sur l'existence d'une relation circulaire spatiale entre les différentes transformations (cf Figure 7b). Soient \mathbf{L}_t la transformation du Lidar entre les instants t et $t + 1$, \mathbf{I}_t la transformation de la centrale inertielle entre ces deux mêmes instants, et \mathbf{X} la transformation rigide entre la centrale inertielle et le Lidar, si toutes ces transformations sont exprimées en coordonnées homogènes par des matrices 4×4 , alors cette relation circulaire peut s'écrire :

$$\mathbf{L}_t \cdot \mathbf{X} \cdot \mathbf{I}_t^{-1} \cdot \mathbf{X}^{-1} = \mathbf{I}_4 \quad (1)$$

où \mathbf{I}_4 est la matrice identité.

Nous avons implémenté et évalué deux méthodes différentes pour la résolution de ce problème. La première méthode, appelée "Handeye" dans la suite de ce papier, est la méthode linéaire proposée par l'INRIA dans [7]. La seconde solution est une optimisation non-linéaire pour résoudre l'équation (1) que nous avons implémenté grâce à la librairie Ceres [11]. Dans notre implémentation, les rotations sont exprimées grâce à la paramétrisation axe-angle (*axis-angle*). Cette paramétrisation permet d'éviter les problèmes de blocage de cardan et de stabilité numérique. Nous optimisons donc les paramètres de rotation et de translation pour minimiser l'erreur résiduelle constituée de deux critères :

- La norme de la translation résultante.
- La norme du vecteur de rotation (représentation axis-angle), c'est-à-dire l'angle solide.

Le principal avantage par rapport à la première méthode est de pouvoir fixer les paramètres que nous ne voudrions pas optimiser. Par exemple, la translation T_z , qui dans le cas des trajectoires de véhicule majoritairement dans le plan est difficile à estimer.

Après synchronisation des données de la centrale inertielle et Lidar, une pose de la centrale est interpolée pour chaque *timestamp* Lidar. Les mouvements de la centrale inertielle sont obtenus directement par différentiation des positions : $I_t = (P_{imu_t}^{-1} \cdot P_{imu_{t+1}})^{-1}$. Les mouvements du Lidar sont calculés par un algorithme ICP entre les deux nuages de points issus des instants t et $t + 1$ [12]. Nous avons utilisé l'algorithme d'*Iterative Closest Point* (ICP) généralisé [13] réalisant une correspondance entre points et plans plus robuste que la version originale dans le cas des nuages de points peu denses.

4.4 Résultats et discussion

Nous avons commencé par générer des données de simulation, figure 8, avec une transformation Lidar/Centrale inertielle parfaitement connue afin de valider les deux méthodes. Les données de la centrale inertielle et de l'ICP étant bruitées, nous avons introduit un bruit équivalent dans les données simulées. La figure 9 présente le résultat de l'influence du bruit de rotation et de translation de mesure des mouvements sur l'estimation du calibrage. Ces données ont été obtenues avec la méthode *Ceres*, la génération des données et le calibrage sont exécutés 10 fois pour chaque cas de bruit de manière à calculer une moyenne et un écart-type par rapport au calibrage idéal. Nous pouvons noter que le bruit de rotation, du moins dans la plage de variation étudiée (0 à 0.8°) n'influence pas énormément les résultats. Le bruit de translation semble nettement plus impactant, il a une forte influence sur la précision et répétabilité du calibrage.

Nous avons également mené une étude sur la précision du calibrage en fonction du nombre d'échantillons de mouvement utilisés dans le calcul, figure 10. Comme nous pouvions nous y attendre avec des données bruitées, l'influence est importante. La dimension du bras de levier à calibrer ne semble pas influencer l'erreur absolue. L'erreur de translation sur le calibrage est proportionnelle au bruit de translation des mouvements.

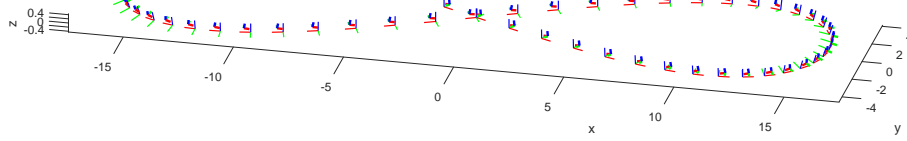


FIGURE 8 – Trajectoire en forme de huit générée en simulation. Grands trièdres : positions de la centrale inertielle, petits trièdres : positions Lidar. La simulation tient compte de la courbure de la trajectoire afin de générer des variations d'angle de roulis.

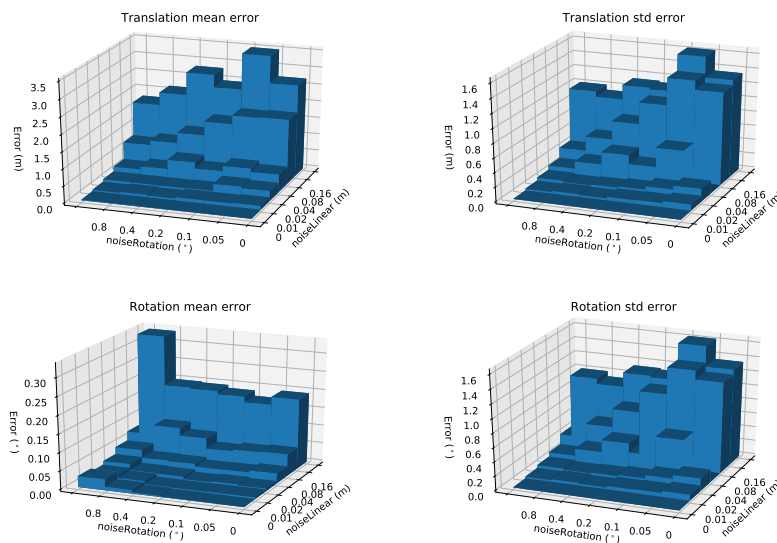
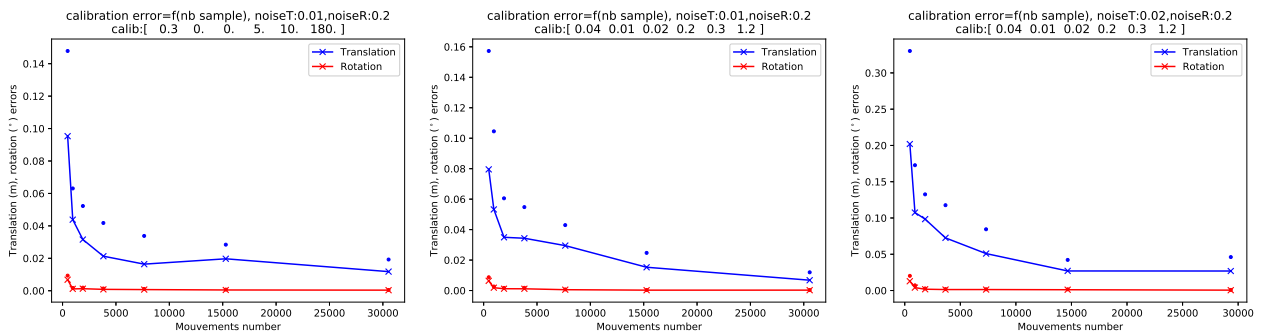


FIGURE 9 – Influence du bruit en translation et rotation, sur les résultats du calibrage. Ligne du haut, moyenne et écart type de la norme en translation de l'erreur de calibrage. Ligne du bas, moyenne et écart type de l'angle solide de l'erreur de calibrage.



(a) X à calibrer important $T_x = 30\text{cm}$, bruit de translation 1cm , de rotation 0.2°

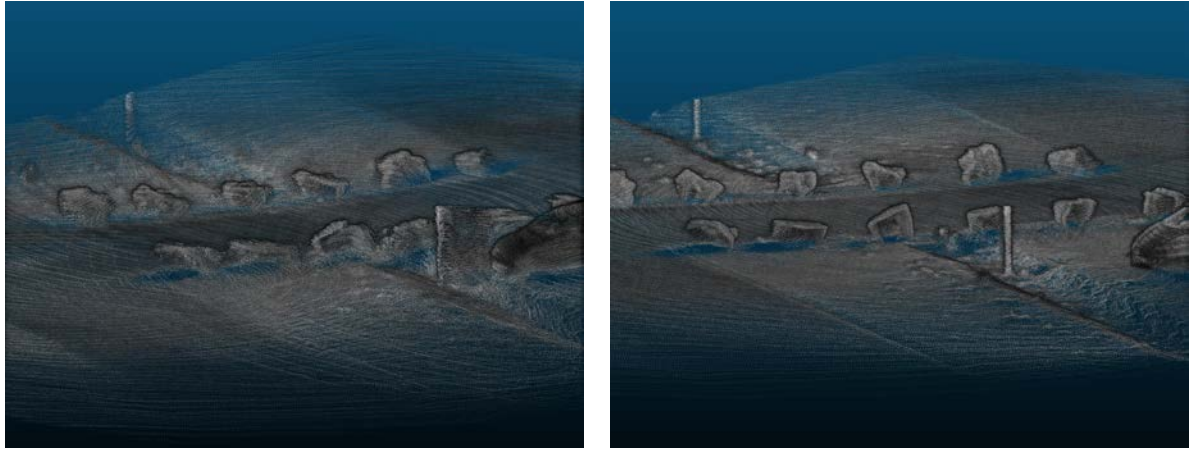
(b) X à calibrer faible $T_x = 4\text{cm}$, bruit de translation 1cm , de rotation 0.2°

(c) X à calibrer faible $T_x = 4\text{cm}$, bruit de translation 2cm , de rotation 0.2°

FIGURE 10 – Étude de l'influence du nombre de mouvements utilisés pour le calibrage, pour 3 cas de calibrage et niveaux de bruit de translation. Traits pleins : erreurs en translation et rotation, points : écart type.

Dataset		Mécanique	Handeye	CERES	CERES $T_z = 0$
A	T_{XYZ}	(0 0 -0.08)	(-0.038 -0.027 0.246)	(-0.034 -0.027 0.101)	(-0.032 -0.028 0)
	R_{XYZ}	(0 0 0)	(0.023 -0.233 -0.139)	(0.0393 0.046 -0.749)	(0.034 0.045 -0.745)
	Nb mvts	-	1451	1623	1623
	Error	-	0.0123	0.0162	0.0162
B	T_{XYZ}	(0 0 -0.08)	(-0.072 0.015 0.286)	(-0.062 0.012 0.049)	(-0.061 0.021 0)
	R_{XYZ}	(0 0 0)	(0.099 -0.203 -2.52)	(0.028 0.076 -2.137)	(0.026 0.076 -2.136)
	Nb mvts	-	1853	2198	2198
	Error	-	0.0128	0.0168	0.0168
C	T_{XYZ}	(0 0 -0.08)	(-0.059 -0.023 0.008)	(-0.040 -0.015 -0.206)	(-0.045 -0.016 0)
	R_{XYZ}	(0 0 0)	(0.062 -0.239 -2.81)	(0.078 -0.029 -0.734)	(0.076 -0.026 -0.736)
	Nb mvts	-	1349	1585	1585
	Error	-	0.0155	0.0175	0.01175

TABLE 3 – Résultat de calibration avec le HDL64e. mécanique : mesure manuelle de la transformation, Handeye : première méthode linéaire, CERES : méthode non linéaire, CERES $T_z = 0$: méthode non linéaire avec T_z fixée à 0. Translations en m , rotations en angles d’Euler ZXY °



(a) Avant

(b) Après

FIGURE 11 – Zoom sur une carte reconstruite à partir de notre coffre de toit, avant et après calibration et correction de la synchronisation.

La table 3 présente les résultats sur des données réelles, obtenues avec notre coffre de toit (figure 1a et le Lidar Velodyne HDL64e. La colonne *Mécanique* représente la transformation recherchée obtenue par mesure manuelle. Les repères du Lidar et de la centrale inertielle sont presque confondus car nous utilisons cette mesure de la transformation comme donnée d’entrée du composant de correction 3D de scan (figure 3). Le bras de levier recherché par calibration est donc faible. La ligne *Nb mvts* correspond au nombre de mouvements conservés dans le jeu de données pour effectuer le calibration. La ligne *Error* est la moyenne de la norme de l’erreur de re-projection à partir de l’équation 1 après calibration sur l’ensemble du jeu de données.

Actuellement nos jeux de données ne contiennent pas suffisamment de mouvements ce qui entraîne une forte variabilité des résultats de calibration. Ce point a été mis en évidence en simulation comme montré précédemment (figure 10). Nous pouvons tout de même remarquer que T_z sera plus difficile à estimer car le véhicule se déplace majoritairement dans le plan.

5 Conclusions

Cet article a présenté les problèmes que posent la mise en correspondance de données Lidar avec un système de positionnement GPS-RTK. Cette question devient centrale de nos jours en particulier avec les développements pour le véhicule autonome qui exigent des précisions de perception et localisation inférieures au décimètre. Dans ce papier nous avons exposé les principaux points de passage pour mettre en oeuvre un système de reconstruction 3D à partir d’un Lidar 3D et d’un GPS-RTK combiné à une centrale inertielle. Comme nous l’avons présenté, la qualité de la reconstruction 3D ne dépend pas uniquement de la performance intrinsèque de chaque composant ; il est crucial de garantir une parfaite synchronisation des données et d’effectuer avec soin l’étape de calibration extrinsèque entre le Lidar et la centrale inertielle. A partir des méthodes proposées dans la littérature, nous

avons mis en oeuvre un banc de test complet afin d'établir la performance globale de la chaîne d'acquisition. Les résultats de reconstruction 3D avant et après correction présentés sur la figure 11 montrent les gains possibles et l'importance de bien maîtriser ces deux étapes.

Plusieurs pistes d'amélioration restent à explorer :

- l'utilisation d'un boîtier de synchronisation matérielle pour réduire l'erreur de datation commise sur les signal issu de la photodiode ;
- la réalisation de jeux de données plus complets pour mener une étude de variabilité du calibrage similaire à ce que nous avons présenté en simulation.

6 Références

- [1] J. Levinson, J. Askeland, J. Becker, J. Dolson, D. Held, S. Kammel, J. Z. Kolter, D. Langer, O. Pink, V. Pratt, *et al.*, "Towards fully autonomous driving : Systems and algorithms," in *Intelligent Vehicles Symposium (IV), 2011 IEEE*, pp. 163–168, IEEE, 2011.
- [2] J. M. Anderson, K. Nidhi, K. D. Stanley, P. Sorensen, C. Samaras, and O. A. Oluwatola, *Autonomous vehicle technology : A guide for policymakers*. Rand Corporation, 2014.
- [3] R. W. Wolcott and R. M. Eustice, "Robust lidar localization using multiresolution gaussian mixture maps for autonomous driving," *The International Journal of Robotics Research*, vol. 36, no. 3, pp. 292–319, 2017.
- [4] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "Robust robot localization in a complex oil and gas industrial environment," *Journal of Field Robotics*, 2017.
- [5] P. Merriaux, Y. Dupuis, R. Boutteau, P. Vasseur, and X. Savatier, "Correction de nuages de points lidar embarqué sur véhicule pour la reconstruction d'environnement 3d vaste," (Clermont Ferand, France), RFIA, june 27-01 2016.
- [6] R. Tsai and R. Lenz, "Real time versatile robotics hand/eye calibration using 3d machine vision.," in *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 554–561, 1988.
- [7] R. Horaud and F. Dornaika, "Hand-eye calibration.," *International Journal of Robotics Research (IJRR)*, vol. 14, no. 3, pp. 195–210, 1995.
- [8] K. Daniilidis, "Hand-eye calibration using dual quaternions.," *International Journal of Robotics Research (IJRR)*, vol. 18, no. 3, pp. 286–298, 1999.
- [9] G. Antonelli, F. Caccavale, F. Grossi, and A. Marino, "Simultaneous calibration of odometry and camera for a differential drive mobile robot.," in *IEEE International Conference on Robotics and Automation (ICRA)*, (Anchorage, USA), pp. 5417–5422, May 2010.
- [10] A. Censi, A. Franchi, L. Marchionni, and G. Oriolo, "Simultaneous calibration of odometry and sensor parameters for mobile robots," *IEEE Transactions on Robotics*, vol. 29, no. 2, pp. 475–492, 2013.
- [11] S. Agarwal, K. Mierle, and Others, "Ceres solver." <http://ceres-solver.org>.
- [12] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*, pp. 586–606, International Society for Optics and Photonics, 1992.
- [13] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp.," in *Robotics : science and systems*, vol. 2, p. 435, 2009.

Géolocalisation HF et ionosphère / HF Geolocation and ionosphere

Passive HF Geolocation Using TDoA Based Receiver Network

Géolocalisation passive des émetteurs HF par un réseau de récepteurs TDoA

Ankit Jain¹, Pascal Pagani¹, Rolland Fleury¹, Michel Ney¹, Patrice Pajusco¹

¹ IMT Atlantique, Lab-STICC UMR CNRS 6285, Brest, France, {ankit.jain; pascal.pagani; rolland.fleury; michel.ney; patrice.pajusco}@imt-atlantique.fr

Keywords: HF propagation, passive geolocation, ionosphere, time difference of arrival.

Mots clés : Propagation HF, géolocalisation passive, ionosphère, différence de temps d'arrivée

Abstract

This paper presents the geolocation of a High Frequency (HF) transmitter (Tx) located at distances of about 700 to 1300 km from a network of receivers (Rx's). HF signals propagate through the ionospheric medium using skywaves and are reflected back to the Earth, thereby allowing HF communications over long distances. In our study, the location of the HF Tx is found using time difference of arrival (TDoA) method. Four remotely controllable Rx's which are built to capture HF signals synchronously are installed in four different cities in France. HF radio broadcast signals are captured simultaneously on the four deployed Rx's from the HF Tx situated in Nauen, Germany. The TDoA's between the signals received on all the four synchronized Rx's are obtained using the cross-channel sounding method. Using the obtained TDoA's, the Tx location is found using the geolocation algorithm based on TDoA. Experimental results show that it is possible to locate the HF Tx in Nauen with a geolocation error of about 5 km.

Résumé

Cet article présente la géolocalisation d'un émetteur HF situé à des distances d'environ 700 à 1300 km d'un réseau de récepteurs. Les signaux HF se propagent à travers le milieu ionosphérique en utilisant des ondes de ciel et sont réfléchis vers la Terre, permettant ainsi des communications HF sur de longues distances. Dans notre étude, la position de l'émetteur HF est obtenue en utilisant la méthode "Time Difference of Arrival" (TDoA). Des récepteurs contrôlables à distance, conçus pour capturer des signaux HF de manière synchrone, sont installés dans quatre villes différentes en France. Ce réseau de quatre récepteurs permet l'acquisition simultanée de signaux de radiodiffusion HF transmis par un émetteur situé à Nauen, en Allemagne. La différence de temps d'arrivée entre les signaux reçus par les quatre récepteurs est obtenue par une technique appelée sondage de canaux croisés. En utilisant les différences de temps d'arrivée déduites, la position de l'émetteur est estimée en utilisant un algorithme de géolocalisation basé sur la technique TDoA. Les résultats expérimentaux montrent qu'il est possible de localiser l'émetteur HF à Nauen avec une erreur de géolocalisation d'environ 5 km.

1 Introduction

HF radio signals propagating using skywaves are reflected once or multiple times between the Earth and the ionospheric medium. Thus, long-range communication can be established through radio signals transmitted in the HF band. HF communication is extremely useful in defense and civil applications as a complement or fallback solution to satellite communication. Owing to the importance of HF communication in defense and civil applications, geolocation of HF transmitters are of prime importance. In particular, in the context of signal intelligence, passive geolocation methods consist of detecting the source of an electromagnetic signal using a non-intrusive system i.e. without the transmitter being able to detect the localization process [1].

Numerous methods to measure the Direction of Arrival (DoA) of the incoming HF signals are explained in [2]. Most of these direction finding (DF) systems require deployment of large antenna arrays at multiple locations which can be quite expensive. HF geolocation can also be achieved using the single site location (SSL) method which measures the DoA of the HF signal in terms of azimuth and elevation angles [3]. SSL method requires installation of a large antenna array only on a single site which is an advantage over traditional DF systems. But, the ionospheric reflection profile along the propagation path must be known to estimate the ionospheric reflection height which is required to transform the elevation angle to ground range.

HF geolocation can also be performed using time-domain techniques. The location of a HF transmitter can be found from the time of arrivals (ToA's) obtained from multiple single antenna receivers under some assumptions and without the knowledge of the ionospheric profile along the HF propagation path [4]. Also, the cost of the receiver system is much less compared to SSL or DF systems. In passive geolocation systems, the signal

transmission time is not known. Therefore, geolocation cannot be performed using the ToA method. But, the location of the transmitter could be found by multilateration technique using the TDoA's obtained from different receiver pairs. Further, the TDoA technique requires a minimum of four receivers to perform localization [5].

This paper aims to explain that it is possible to perform HF geolocation using the TDoA method. To validate this concept, multiple receivers capable of capturing HF signals synchronously are built using Software Defined Radio (SDR). Section 2 explains the HF propagation using skywaves. In Section 3, the HF receiver network capable of synchronously capturing HF radio signals is presented. In Section 4, a capture from the HF transmitter in Nauen, Germany is analyzed and the geolocation result is presented. Finally, conclusions are drawn in Section 5.

2 HF Propagation Principles

The ionosphere lies from about 50 km to 2000 km above the surface of the Earth [6]. The neutral atoms present in the atmosphere are ionized from the sun's radiation which results in the formation of free electrons which affects HF radio propagation. The ionosphere is split into D, E and the F layer in the order of increasing altitude, respectively. Moreover, the electron density is also variable in each of the layers making the ionosphere a dispersive medium and capable of refraction of the HF radio signals. The D layer mainly accounts for the absorption of the HF radio signals and has the least electron density. The electron density is the maximum in the F layer which is present between 150 to 600 km and it accounts for large propagation paths. The ionosphere is also a highly dynamic medium which varies diurnally, seasonally, geographically and with respect to the solar activity. This variability in the ionosphere is one of the major issues in HF geolocation of signals propagating through skywaves, as it leads to time delay variations [7].

HF radio propagation is possible using ground waves and the skywaves. Long distance HF radio transmissions occur via the skywaves where the radio signals are reflected back to the Earth through the ionosphere. HF radio signals which are reflected once from the ionosphere are referred to as single-hop mode whereas radio signals which are reflected multiple times are known as multi-hop modes. It is possible to communicate up to distances in the range of 500-3000 km using single-hop propagation mode. Due to the presence of Earth's magnetic field, an additional propagation mode due to extraordinary waves exist whereas ordinary waves are present even in the absence of the magnetic field. These propagation modes can be distinguished with respect to the direction of rotation of the waves.

In our geolocation study, we assume that all the received HF radio signals follow single-hop propagation mode considering the distances at which the HF transmitter is located from the receivers. Figure 1 presents the geometry for HF single-hop propagation mode. The transmitter is located at point Tx and receiver is located at point Rx₁. Due to the varying refractive index in the ionosphere, the HF radio signals propagate along a curved path and the real reflection height in the ionosphere is represented by h_m as seen in Figure 1. As per the Breit and Tuve's theorem [6], the time required for the HF radio signal to travel along the red curved path through the ionosphere is equal to the time required to travel along the black triangular path in vacuum. As seen in Figure 1, the reflection height for the triangular path is represented by h_v and it is referred to as virtual reflection height.

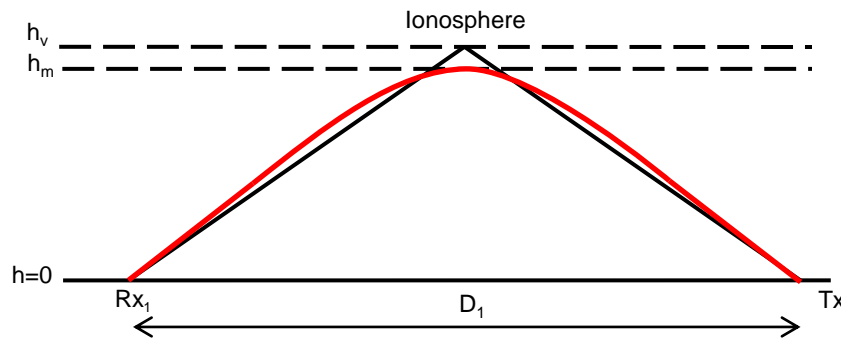


Figure 1 : Single-hop propagation mode geometry

Moreover, the HF geolocation model assumes a uniform ionosphere due to which the reflection height is the same for all the HF paths. The Earth is assumed to be flat by neglecting its curvature. Under these assumptions, it is possible to account for the excess delay due to the ionospheric propagation in the geolocation process, even without information about the ionospheric reflection height [4].

3 HF Receiver Network

In order to perform HF geolocation using the TDoA method, four remotely controllable receivers are built to synchronously capture HF radio signals.

3.1 Receiver Design

The receiver system design is illustrated in Figure 2. It is composed of both hardware and software. The main component of the receiver setup is the SDR Ettus USRP N200/N210 box whose RF frontend is connected to an active HF antenna. The SDR box consists of a direct sampling analog-to-digital converter (ADC) capable of sampling up to 100 MHz, coupled with an FPGA that performs downsampling and frequency conversion. The LFRX daughter card is capable of accepting signals from DC to 30 MHz. Using the GPS antenna, the GPS disciplined oscillator (GPSDO) can be locked to global GPS standards and the signal fluctuation is within ± 50 ns [8]. The GPSDO allows multiple USRP's which are located far away from each other to synchronize their captures with the same clock.

The receiving setup communicates to their respective computers via the Ethernet connection. The computer consists of several programs that are used to schedule and capture data according to user requirement. The scheduling of the captures is done by a Java interpreter. It also manages the launch of the Python scripts which act as an interface between hardware and software. Python scripts are written to synchronize the capture time with the GPS time, initialize the capture and capture the data as per GNU Radio standards [9]. A capture generates two files: a DAT file and a HDR file. The DAT files contain complex samples of each capture whereas the HDR files consist of metadata corresponding to the respective DAT files.

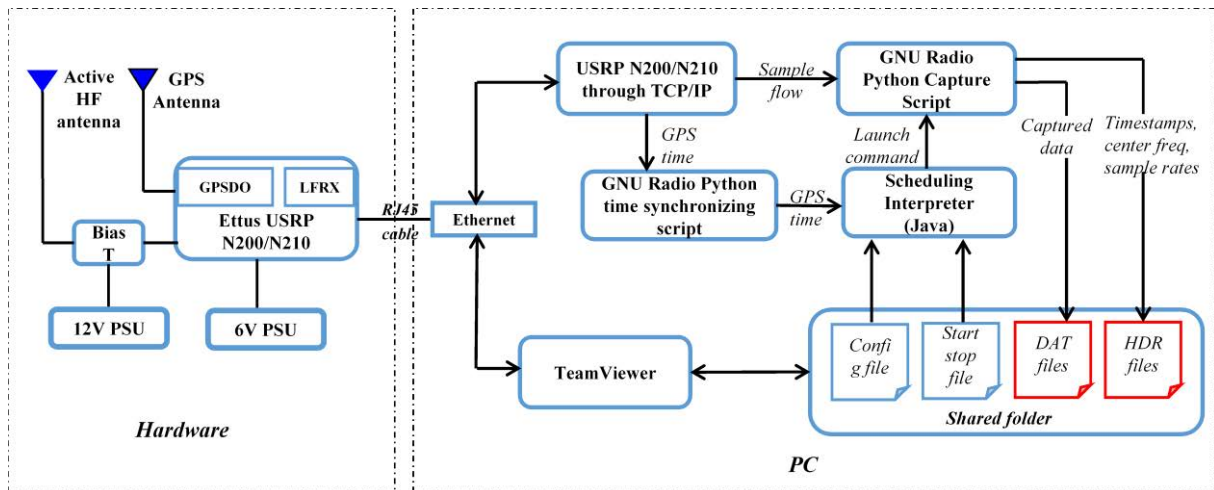


Figure 2 : Receiver system design

3.2 TDoA Receiver Network

The receiver system explained above is deployed in four different cities in France. As illustrated in Figure 3, the receivers are deployed in Brest, Bordeaux, Grenoble and Lille. HF radio signals were captured simultaneously from the HF transmitter located in Nauen, Germany. The geographic location of the HF transmitter in Nauen is presented in Figure 3. All the receivers in the network are managed from a central machine located in Brest. The ground ranges for the different HF links is summarized in Table 1.

HF radio link (Rx-Tx)	Ground Range (km)
Brest-Nauen	1320
Bordeaux-Nauen	1313
Grenoble-Nauen	976
Lille-Nauen	710

Table 1 : Ground distance between the HF transmitter and all the receivers

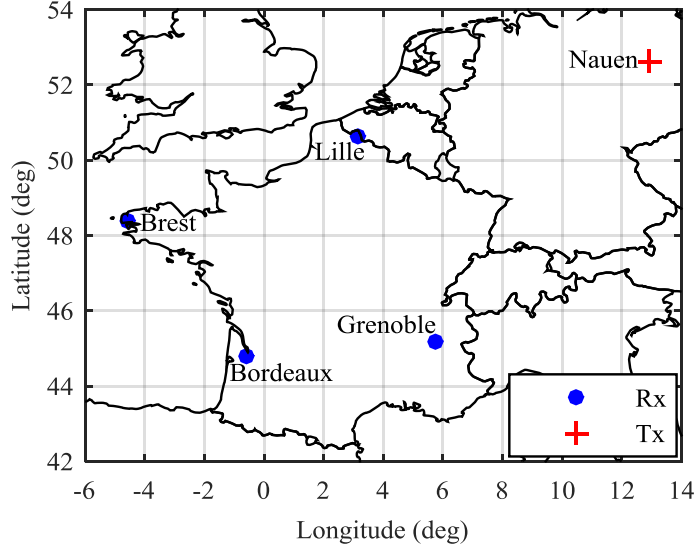


Figure 3 : GPS coordinates of the transmitter site and all the receiving sites in the network

3.3 Measurement Details and Signal Processing

All the HF radio broadcast signals are captured for a span of 5 seconds and sampled at the rate of 200 kHz. The captured signal is filtered in the frequency domain to remove undesired parts of the spectrum. As most of the HF broadcasts are transmitted over a 10 kHz band, only signals within this bandwidth around the carrier frequency of the transmitter are used, the rest is filtered out. The signal is transformed back to the time domain and received signal average power is calculated. Finally the message signal samples are obtained by demodulating the transformed signal.

3.4 Reference Receiver Selection

The geolocation algorithm used to estimate the geographic location of the HF transmitter is dependent on the reference receiver [10]. The received signal average powers for all the receivers were compared and the receiver with the highest average power was considered as the reference receiver. Using the cross-channel sounding technique [11], the message signal samples from the reference receiver are cross-correlated with the message signal samples from the other three receivers to obtain the three TDoA estimates. The TDoA estimate (Δt) corresponds to the maximum value of the cross-correlation and is expressed as follows:

$$\Delta t = \max_t R_{s_1 s_i}(t) = \max_t \int_{-\infty}^{\infty} s_1(\tau) s_i(\tau + t) d\tau \quad (1)$$

where $R_{s_1 s_i}(t)$ is the cross-correlation between the message signal samples $s_1(t)$ obtained from the reference receiver and the message signal samples $s_i(t)$ obtained from the other three receivers ($i = 1,2,3$), respectively.

4 Experimental Results

Multiple HF broadcast radio signals were captured simultaneously by all the receivers in the network between 19:11–19:20 UTC on 18th July 2017 from the HF transmitter located in Nauen. The interval between the start times of each capture was equal to one minute. The transmitter was emitting at a frequency of 11.790 MHz. The GPS locations of the broadcast transmitter in Nauen and all the four receivers were transformed to x and y coordinates using the azimuthal equidistant projection [12] and considering the selected reference receiver as the origin of the coordinate system. An analysis of a data capture is presented in the following section.

4.1 Data capture analysis

A HF radio broadcast signal captured from Nauen simultaneously by all the four Rx's on 18th July 2017 at 19:17 UTC is analyzed. The signal received in Grenoble was considered as the reference signal on the basis of the received signals power. The signal captured in Grenoble is presented in Figure 4.

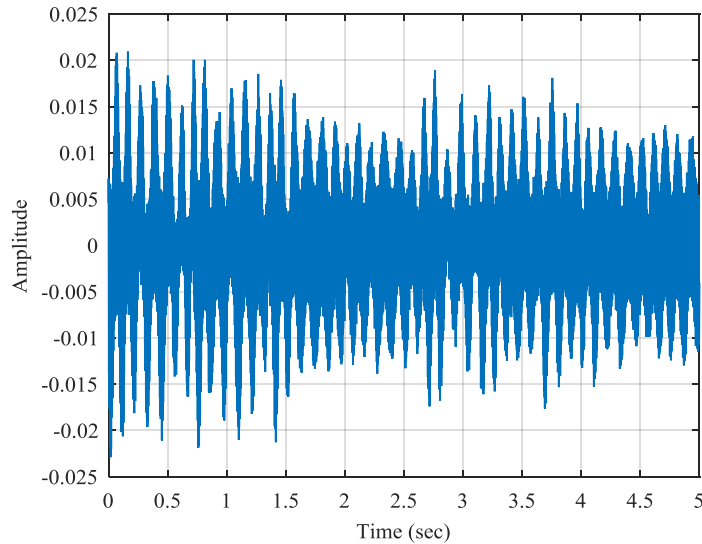


Figure 4 : HF signal received in Grenoble from the HF transmitter in Nauen on 18th July at 19:17 UTC

The message signal samples were obtained from the four receivers as explained earlier. The propagation duration differences for the signals received at different receiver pairs (i.e. Grenoble-Brest, Grenoble-Bordeaux and Grenoble-Lille) were obtained using the cross-correlation method. An example of the cross-correlation output for the signals captured in Grenoble and Lille is shown in Figure 5. Using the obtained TDoA estimates, the Tx coordinates were found by solving a system based on quadratic equations [10]. The estimated coordinates were transformed to geographical coordinates and the geolocation error was found to be 5.7 km.

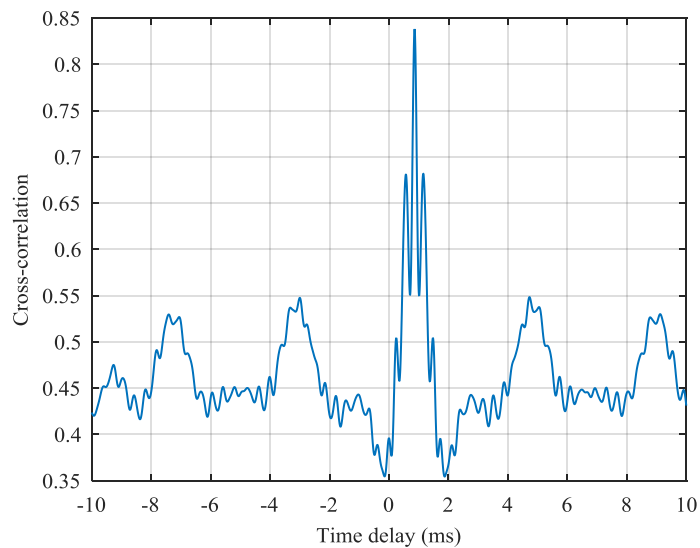


Figure 5 : Cross-correlation of radio signal captured in Grenoble and Lille from the HF transmitter in Nauen

Table 2 provides the summary of the geolocation error for the 10 different captured signals from the HF transmitter in Nauen on 18th July between 19:11-19:20 UTC. For all ten captures, the received signal in Grenoble was selected as the reference receiver on the basis of received signal intensity. The signal-to-noise ratio (SNR) for all ten signals captured in Grenoble was at least 12 dB. The minimum geolocation error of 5.7 km corresponds to a relative error of about 0.6% computed with respect to the maximum ground range among the four possible HF paths whereas the maximum geolocation error of 112.6 km corresponds to a relative error of about 5.9%.

Number of captures considered	10
Minimum geolocation error (km)	5.7
Maximum geolocation error (km)	112.6
Median geolocation error (km)	30.7

Table 2 : Summary of geolocation error for all HF signals captured from Nauen

5 Conclusion

In this paper, we presented the results of geolocation of the HF transmitter in Nauen, performed using the TDoA method. The TDoA method is well suited for passive geolocation and can be performed using light and portable equipment. Four receivers capable of synchronously capturing distant HF signals have been built and deployed in four different cities in France. We have also presented the receiver design in detail. Multiple signals are captured simultaneously using the receiver network from the HF transmitter in Nauen. The measurement data are analyzed and the geolocation errors are presented. The analysis of the geolocation output allows us to believe that HF geolocation can be performed using the TDoA method.

In future, we plan to capture large amount of data using different HF broadcast transmitters around Europe and perform a statistical study with respect to the different parameters of the received signal. In order to improve the geolocation accuracy, several aspects related to the use of ionospheric information along with the estimated TDoA's need to be studied further in detail.

Acknowledgements

This work is supported in part by the Carnot Institute and the Brittany region in France. The authors thank the Atmospheric Optics Laboratory in University of Lille 1 and the Electrical Engineering Department in IUT1-University Grenoble Alps for hosting our receiver setup.

References

- [1] D. Heurquier, "La localisation d'émetteurs de radiocommunication en zone urbaine," Journées Scientifiques URSI France, Mar. 2014.
- [2] P. J. D. Gething, "HF direction finding," Proceedings of the Institution of Electrical Engineers, vol. 113, no. 1, pp. 49–61, Jan. 1966.
- [3] G. Fabrizio and A. Heitmann, "A multipath-driven approach to HF geolocation," Signal Processing, vol. 93, no. 12, pp. 3487–3503, Dec. 2013.
- [4] P. Pagani, R. Fleury, Y. L. Roux, and D. L. Jeune, "A study of HF transmitter geolocation through single-hop ionospheric propagation," 8th EuCAP, 2014, pp. 2689–2693.
- [5] D. J. Torrieri, "Statistical theory of passive location systems," IEEE Trans. on Aerospace and Electronic Systems, vol. AES-20, no. 2, pp. 183–198, Mar. 1984.
- [6] K. Davies, "Ionospheric radio," IEE Electromagnetic Waves Series 31, pp. 124-158, n° ISBN: 0 86341 186 X, May 1989.
- [7] R. W. LaBahn and R. B. Rose, "Time delay variations in HF propagation," Radio Science, vol. 17, no. 5, pp. 1285–1299, Sep. 1982.
- [8] "GPSDO Selection Guide - Ettus Knowledge Base." Available: <https://kb.ettus.com/GPSDO>.
- [9] "USRP-GNURadio-gnuradio.org." Available: <http://gnuradio.org/redmine/projects/gnuradio/wiki/USRP>.
- [10] A. Jain, P. Pagani, R. Fleury, M. M. Ney, and P. Pajusco, "Efficient time domain HF geolocation using multiple distributed receivers," 11th EuCAP, 2017, pp. 1852–1856.
- [11] A. Jain, P. Pagani, R. Fleury, M. M. Ney, and P. Pajusco, "Cross-Channel Sounding for HF Geolocation: Concepts and Experimental Results," The 12th European Conference on Antennas and Propagation (EuCAP), 2018, accepted for publication.
- [12] J. P. Snyder, "Map projections: A working manual," U.S. Government Printing Office, Washington, D.C., Series 1395, pp. 191–202, 1987.

Propagation des ondes hautes fréquences dans l'ionosphère terrestre *High frequencies wave propagation in the Earth ionosphere*

Foucault E. ⁽¹⁾, **Blelly P-L.** ⁽²⁾, **Marchaudon A.** ⁽³⁾

¹ IRAP, CNRS, Toulouse, France, etienne.foucault@irap.omp.eu

² IRAP, CNRS, Toulouse, France, pierre-louis.blelly@irap.omp.eu

³ IRAP, CNRS, Toulouse, France, aurelie.marchaudon@irap.omp.eu

Mots- clefs : Ionosphère, propagation, modélisation

Keywords : Ionosphere, propagation, modeling

Résumé:

L'ionosphère est un domaine critique de l'environnement spatial de la Terre, qui joue un rôle clef dans la dynamique du système magnétosphérique. De fait, la météorologie de l'espace, qui est le cadre applicatif des études autour des relations Soleil-Terre orientée vers l'impact de ces relations sur les activités humaines, s'intéresse particulièrement à l'ionosphère, tant ses effets sont visibles et immédiats et peuvent s'avérer critiques, notamment pour les communications. Afin de mieux comprendre comment ces relations peuvent impacter les communications, il a été développé un code de tracé de rayon, en géométrie sphérique et à trois dimensions, pour simuler les trajectoires d'ondes radios dans une ionosphère réaliste. Les premiers résultats de cette étude sont encourageants, les trajectoires simulées respectant le principe de Fermat, et suivant les propriétés du milieu (variation de la densité électronique). L'objectif de la thèse étant de coupler ce tracé de rayon avec le module de simulation d'ionosphère IPIM, développé à l'IRAP, pour créer un modèle de prévision de la propagation des ondes radios et de l'ionosphère. Il est également prévu d'injecter des données issues d'observations pour contraindre le modèle numérique.

Resume:

The ionosphere is a critical area of the Earth's space environment, which plays a key role in the dynamics of the magnetospheric system. In fact, Space Weather, which is the applicative (?) framework of studies around Sun-Earth relations oriented towards the impact of these relations on human activities, is particularly interested in the ionosphere, as its effects are visible and immediate and can be critical, especially for communications. In order to better understand how these relationships can impact communications, a three-dimensional, spherical geometry ray tracing code has been developed to simulate radio wave trajectories in a realistic ionosphere. The first results of this study are encouraging, the simulated trajectories respecting the Fermat principle, and the properties of the medium (variation of the electron density). The aim of the thesis is to couple this ray tracing tool with the ionosphere simulation module, IPIM, developed at IRAP, to create a forecast model for radio waves propagation and ionosphere estimation. It is also planned to inject data from observations to constrain the numerical model.

1 Introduction

Le développement des moyens de communications à grandes échelles (télécommunications, systèmes de positionnement type GPS, ...) et le besoin de précision et de continuité de tels systèmes rendent nécessaire une meilleure compréhension du milieu dans lesquels ces moyens sont utilisés. En particulier, l'ionosphère est certainement le milieu qui interagit le plus fortement avec les ondes radios et cette étude s'intéresse à quelques effets de ce milieu sur les ondes en bande HF. L'ionosphère est composée d'un plasma partiellement ionisé (milieu électriquement neutre, composé d'électrons et d'ions), résultat de l'ionisation par le rayonnement solaire des constituants neutres de l'atmosphère terrestre. Ce plasma modifie sensiblement l'indice de réfraction du milieu, qui devient notamment dépendant de la fréquence, avec une forte capacité à modifier la trajectoire d'une onde électromagnétique se propageant dans l'ionosphère. En fonction de la densité électronique présente sur le chemin

de celle-ci, la perturbation peut aller d'un simple retard temporel, à la disparition totale du signal.

Afin de mieux comprendre et appréhender les changements de trajectoires des ondes radios dans l'ionosphère, nous présentons ici les premiers résultats d'un code de tracé de rayons (estimation de la trajectoire, en minimisant l'énergie le long de celle-ci) dans une ionosphère simulée numériquement à partir d'un modèle physique réaliste de ce milieu. Dans la suite, on présentera en premier l'architecture du code de tracé de rayons, puis l'application à une ionosphère modélisée. Une discussion des résultats obtenus sera ensuite proposée, avec ensuite une conclusion sur l'état des travaux de recherche.

2 Tracé de rayon

L'objectif du développement de ce code de tracé de rayons, est de permettre une estimation de la trajectoire d'une onde radio (de haute (HF) à très haute fréquence (VHF)), dans un milieu possédant un indice de réfraction variable, mais supposé stable durant le temps de propagation de l'onde. Pour cela, le code de tracé de rayons résout la trajectoire d'une onde de fréquence f dans un système de coordonnées sphérique (r, ϕ, θ) , où r est la distance au centre de la Terre, ϕ et θ respectivement la longitude et latitude géographiques, en se basant sur le principe de Fermat. Pour cela on utilise un algorithme de type Runge-Kutta, à pas variable, pour intégrer ces quantités le long du trajet. Le schéma d'incrémentatation est le suivant :

$$\begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix}_{S+\delta S} = \begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix}_S + \delta \begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix}_{\delta S} + \delta \begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix}_{\delta S} (n) \quad (1)$$

$$\begin{pmatrix} a_z \\ e_l \end{pmatrix}_{S+\delta S} = \begin{pmatrix} a_z \\ e_l \end{pmatrix}_S + \delta \begin{pmatrix} a_z \\ e_l \end{pmatrix}_{\delta S} + \delta \begin{pmatrix} a_z \\ e_l \end{pmatrix}_{\delta S} (n) \quad (2)$$

Où S = l'abscisse curviligne

δS = l'incrément d'abscisse curviligne

n = l'indice de réfraction, définit comme suit :

$$n = \sqrt{1 - \frac{N_e e^2}{m_e \epsilon_0 (2\pi f)^2}} \quad (3)$$

Où N_e = la densité électronique

e = la charge électrique élémentaire

m_e = la masse d'un électron

ϵ_0 = la permittivité du vide.

La résolution des équations se fait le long de l'abscisse curviligne S . Les termes d'incrémentations peuvent se mettre sous la forme :

$$\delta \begin{pmatrix} r \\ \phi \\ \theta \end{pmatrix} = \begin{pmatrix} \frac{\sin(e_l)}{\sin(a_z) \cos(e_l)} \\ \frac{r \cos(\theta)}{\cos(a_z) \cos(e_l)} \\ r \end{pmatrix} \delta S + \begin{pmatrix} \cos(e_l) \delta e_{l_i} \\ \frac{-\sin(a_z) \sin(e_l) \delta e_{l_i} + \cos(a_z) \cos(e_l) \delta a_{z_i}}{r \cos(\theta)} \\ \frac{-\cos(a_z) \sin(e_l) \delta e_{l_i} - \cos(e_l) \sin(a_z) \delta a_{z_i}}{r} \end{pmatrix} \delta S \quad (4)$$

$$\delta \begin{pmatrix} a_z \\ e_l \end{pmatrix} = \begin{pmatrix} \frac{\tan(\theta) \sin(a_z) \cos(e_l)}{r} \\ \frac{\cos(e_l)}{r} \end{pmatrix} \delta S + \begin{pmatrix} (\vec{k} \times \vec{e}_s) \cdot \vec{e}_\phi \delta i + \frac{\delta e_{l_i} \sin(a_z) \sin(e_l)}{\cos(a_z) \cos(e_l)} \\ (\vec{k} \times \vec{e}_s) \cdot \vec{e}_r \frac{\delta i}{\cos(e_{l_i})} \end{pmatrix} \delta S \quad (5)$$

Où a_z = azimuth de l'angle de visée, vu depuis le Nord géographique

e_l = l'élévation de l'angle de visée, pris depuis le sol

δi = l'incrément d'angle dû à la présence de l'indice de réfraction, définit comme :

$$\frac{\delta i}{\delta S} = \frac{\|\vec{\nabla} n\|}{n} \vec{v} \cdot \vec{e}_s^\perp = -\frac{\|\vec{\nabla} n\|}{n} \vec{t} \cdot \vec{e}_s \quad (6)$$

Où $\vec{\nabla} n$ = le gradient spatial de l'indice de réfraction

\vec{e}_s = le vecteur de propagation, normé¹

\vec{v} et \vec{t} les vecteurs de la base locale orthonormée tel que $\vec{v} = \frac{\vec{\nabla} n}{\|\vec{\nabla} n\|}$ et $\vec{v} \times \vec{t} = 1$, repris dans la Figure 1.

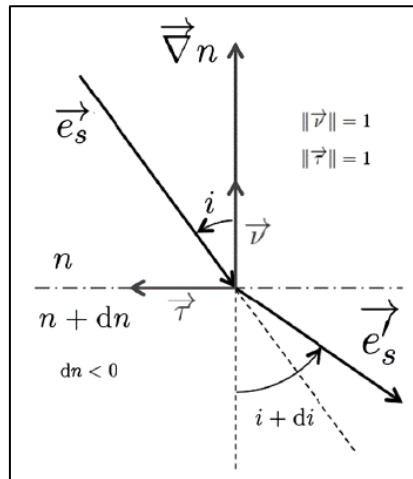


Figure 1 : Rotation local de la direction de propagation, dans le repère local

Le gradient d'indice de réfraction ∇n provoque une rotation du vecteur \vec{e}_s , qui définit la direction de propagation de l'onde, au niveau du dioptre caractérisé par ce gradient d'indice.

¹ Grandin et al., « Mars Express Radio-Occultation Data ».

L'utilisation d'un algorithme à pas adaptatif permet de suivre précisément cette rotation lors de la présence de forts gradients de densité.

3 IPIM

Afin de mieux rendre compte de la propagation des ondes radios dans un milieu de type ionosphérique, il a été choisi pour cette étude de s'appuyer sur les résultats issus du code IPIM² (IRAP Plasmasphere-Ionosphere Model), qui est un modèle numérique de simulation de transport du plasma ionosphérique, suivant la géométrie des lignes de champ magnétiques. Après une rapide présentation du modèle, nous décrivons le type de simulation utilisée dans le cadre de cette étude.

3.1 Présentation du modèle IPIM

Le modèle IPIM est un modèle ionosphérique inter hémisphérique qui résout les équations de transport le long des lignes de champ magnétiques pour les différentes espèces ionisées présentes dans le milieu (ions et électrons). Les paramètres résolus sont la concentration, la vitesse parallèle au champ magnétique, les températures parallèles et perpendiculaires au champ, ainsi que les flux de chaleur associés, selon une approche basée sur une représentation à 16 moments des fonctions de distributions des espèces. Un schéma mixte lagrangien-eulérien permet de considérer le transport du plasma sous l'effet du champ électrique de convection et ainsi d'avoir une approche 3D (1D+2D) du transport du plasma ionosphérique. Ce modèle associe plusieurs modules décrivant l'atmosphère, la chimie, le rayonnement solaire, les effets des particules énergétiques (électrons suprathermiques) sur les constituants du milieu et le transport. La Figure 2 présente un synopsis du modèle IPIM et des couplages internes, qui permettent au modèle de décrire de manière quantitative la structure et la dynamique de l'ionosphère dans ses couplages avec la magnétosphère et le vent solaire.

² Marchaudon et Bletly, « A New Interhemispheric 16-Moment Model of the Plasmasphere-Ionosphere System ».

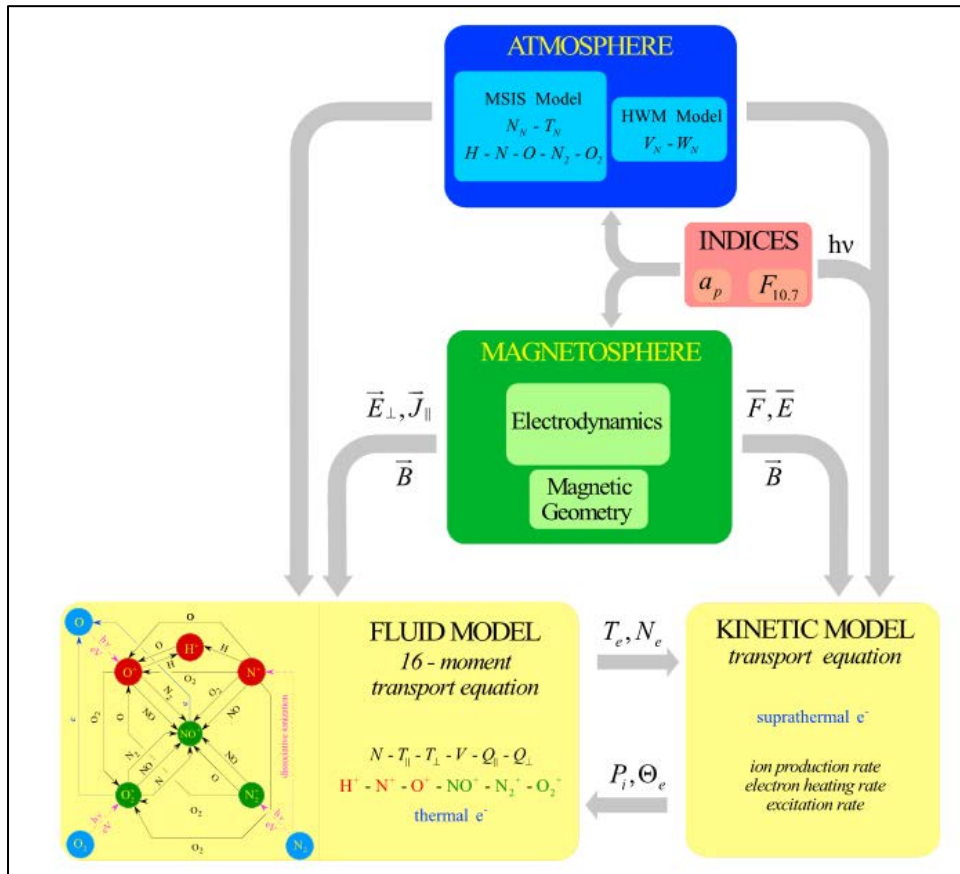


Figure 2 : Synopsis du fonctionnement du modèle IPIM

3.2 Présentation de la simulation, type SuperDARN

Dans le cadre de cette étude, nous avons choisi d'étudier le comportement d'ondes radio de fréquences de l'ordre de 10MHz, afin de simuler les signaux observés par des radars ionosphériques de type SuperDARN³ (Super Dual Auroral Radars Network). Le réseau consiste en plus de 35 stations, permettant le suivi de l'ionosphère depuis les latitudes moyennes jusqu'aux cercles polaires, en mesurant la convection du plasma ionosphérique au niveau des zones aurorales et polaires.

Dans cette étude, nous avons simulé le champ de vue du radar de Stokkseyri (63.83N, -22.02E) pour le 20 Mars 2015 à 8h00 UT. La simulation a permis de produire 16 coupes ionosphériques, représentant les 16 faisceaux de mesure du radar. Dans chacune de ces coupes, différents paramètres du milieu sont accessibles comme par exemple la densité électronique, la température des ions et des électrons, la densité des neutres, la vitesse de convection du plasma ionosphérique associée au transport perpendiculaire aux lignes de champ magnétique. La grille ionosphérique correspondante ainsi placée en entrée du code de tracé de rayons est à trois dimensions (concaténation des 16 coupes dans l'espace), et non régulière. En effet celle-ci s'étend de 80 à 3000 km en altitude en 128 points, de 0 à 3555 km en portée au sol (portée maximale d'un radar type SuperDARN) en 160 points, selon 16 coupes, correspondant aux 16 faisceaux sélectionnés. La Figure 3 propose un aperçu de cette

³ Chisham et al., « A Decade of the Super Dual Auroral Radar Network (SuperDARN) ».

géométrie à une altitude donnée, avec codé en couleur la densité électronique issue du modèle IPIM à 250 km d'altitude.

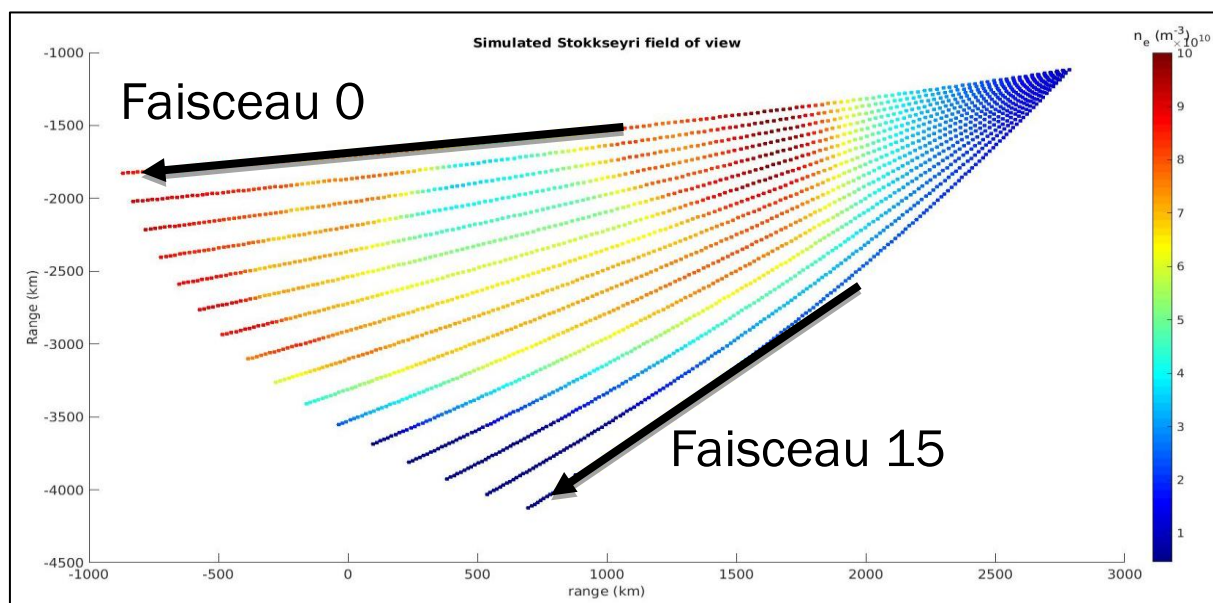


Figure 3 : Grille de densité électronique obtenue après simulation du champ de vue du radar de Stokkseyri

On notera sur la Figure 3, les gradients horizontaux de concentration électronique, qui résultent de l'effet des précipitations dans les régions aurorales. Ces gradients vont jouer un rôle important dans la réfraction des ondes.

4 Résultats

La validation de l'algorithme de tracé de rayon en utilisant une ionosphère synthétique (calculée à base de loi mathématique très simple) sera d'abord présentée, puis nous nous concentrerons sur les premiers résultats du tracé de rayon dans l'ionosphère modélisée par le code IPIM.

4.1 Validation

Afin de valider notre algorithme de tracé de rayon, il a été choisi de se placer dans un premier temps dans un cas simple d'ionosphère parfaitement connue. Pour cela nous avons défini une ionosphère avec un profil de densité électronique défini comme suit :

$$N_e(r) = \left(1 + \tanh\left(\frac{r - r_d}{dr}\right) \tanh\left(\frac{r_u - r}{dr}\right)\right) N_e^o \quad (7)$$

Où r = la distance au centre de la Terre

r_d = l'altitude basse de l'ionosphère

r_u = l'altitude haute de l'ionosphère

dr = l'épaisseur de la couche de transition

N_e^o = la concentration au pic de région F_2 défini à $1e^{11}$.

La Figure 4 présente un exemple de tracé de rayon dans un plan ionosphérique tel que décrit précédemment. On observe bien la forte réfraction obtenue lorsque l'onde traverse le fort gradient vertical de concentration électronique.

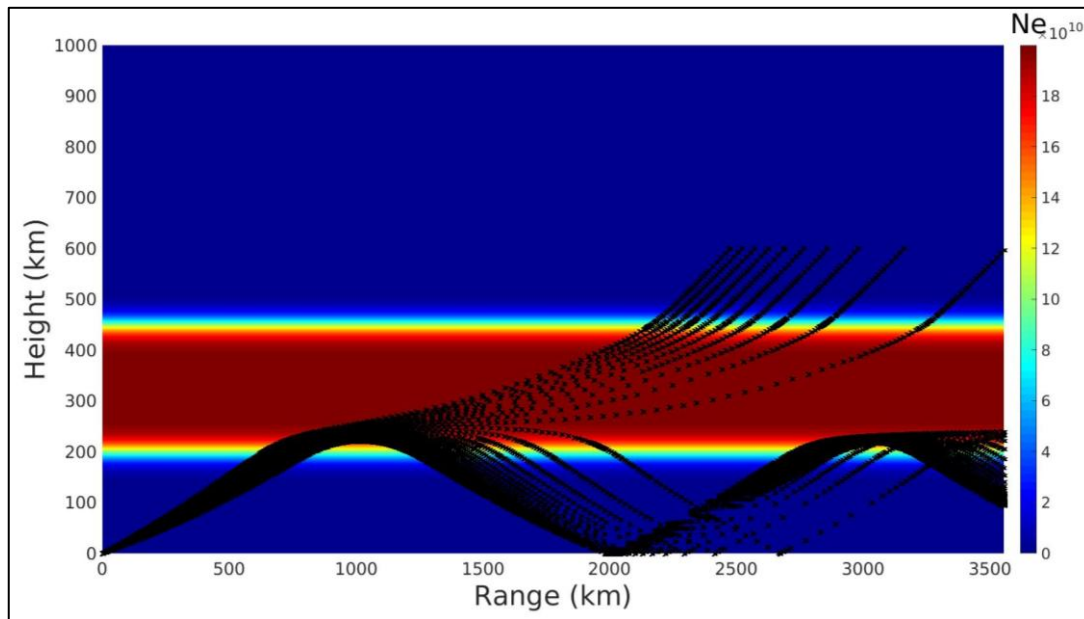


Figure 4 : Exemple de tracé de rayon dans une ionosphère synthétique. Les trajectoires sont représentées en pointillés noirs.

Selon l'angle d'élévation, on voit que les ondes peuvent être diffractées jusqu'à être réfléchi par le sol et se propager à grande distance dans le guide d'onde constitué par la région en dessous de l'ionosphère. A plus forte élévation, l'onde suit le guide d'onde ionosphérique et sort du milieu pour être perdue dans l'espace.

4.2 Simulation IPIM pour un radar type SuperDARN

Dans un second temps, nous nous sommes intéressés aux performances du tracé de rayon dans l'ionosphère simulée par IPIM et présentée en section 3.2. Les résultats présentés dans cette étude utilisent l'ionosphère simulée suivant le faisceau pointant le plus au Nord du radar SuperDARN de Stokkseyri (faisceau 0). La Figure 5 présente les trajectoires obtenues à l'aide du tracé de rayons. Les trajectoires d'élévations faibles subissent une réfraction vers le sol (comme dans le cas précédent), indiquant ainsi que le gradient de concentration électronique de la région E est suffisamment élevé pour réfracter fortement la trajectoire des ondes. Les élévations plus élevées subissent une déviation plus faible sous l'effet de l'ionosphère, mais sans revenir au sol.

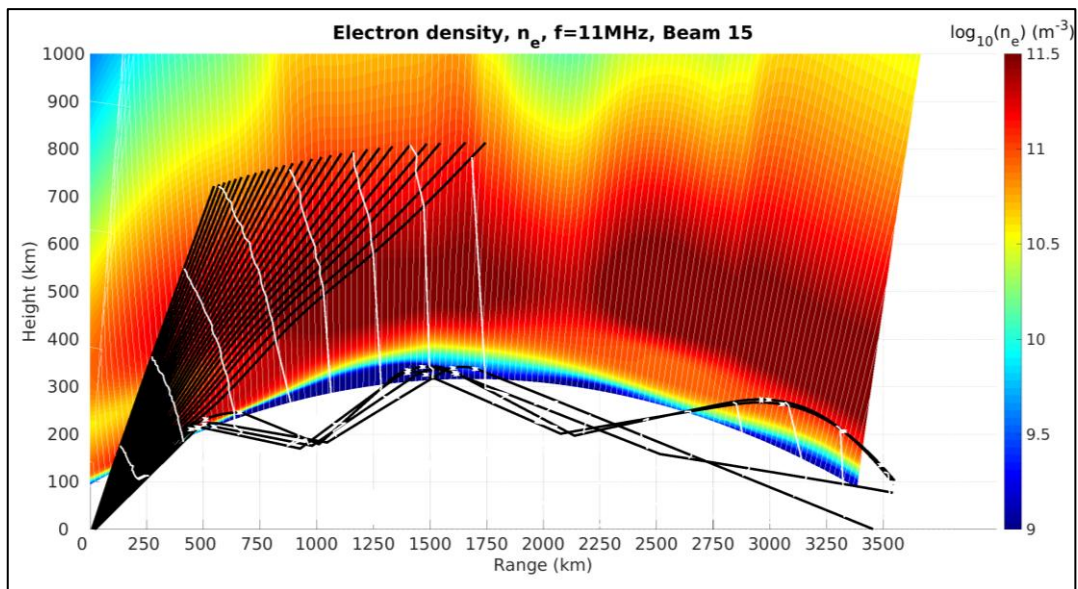


Figure 5 : Tracé de trajectoire (en noires) dans une ionosphère simulée numériquement. En abscisse est représentée la portée au sol (en km) et en ordonnée, l'altitude (en km). Sont représentés en blanc les fronts d'ondes, et les croix blanches sont des points vérifiant les critères de réflexion

La simulation numérique d'IPIM, couplé à ce tracé de rayon à trois dimensions, permet de rendre compte d'un milieu et donc de trajectoire plus réaliste. En effet le code de simulation ionosphérique autorise le transport des particules le long des lignes de champ et permet de prendre en compte les effets des précipitations, qui participent au renforcement des régions D et E de l'ionosphère. A partir de ces résultats, on voit qu'il est possible d'utiliser le couplage entre le tracé de rayons et le code numérique de simulation d'ionosphère IPIM pour estimer localement les zones à viser ou à éviter, si on veut respectivement effectuer une communication longue distance, ou bien communiquer, sans risquer de subir d'interférence, avec un objet à plus haute altitude, comme un satellite.

5 Conclusion

A partir de ces premiers résultats il a été possible de valider les équations de propagation du code de tracés de rayon, en regardant une ionosphère parfaitement définie. Ce résultat nous a permis par la suite de commencer une vague de tests sur une ionosphère simulée numériquement par un modèle physique (IPIM), reproduisant le champ de vue type d'un radar HF SuperDARN. Les résultats de cette première étude sont encourageants. A courte et moyenne portée le code semble robuste et donne des trajectoires cohérentes avec l'environnement rencontré. Il a cependant été noté des problèmes à longue portée, liés aux méthodes d'interpolation choisies. Un travail d'amélioration de celles-ci est actuellement en cours afin de pallier ce problème et garantir la meilleure propagation possible des ondes radios dans une ionosphère la plus réaliste possible. Nous prévoyons également d'introduire les indices de réfractions ordinaires et extraordinaires pour étudier la propagation différenciée dans ondes électromagnétiques selon ces deux modes.

Le but visé une fois le développement de ce code de tracé de rayons finalisé, sera de développer une version simplifiée et plus rapide du code de simulation ionosphérique IPIM. Le couplage de ces deux codes, code de propagation et code de simulation ionosphérique,

aura pour but de prévoir l'état de l'ionosphère terrestre, en incluant dans le modèle des données observées par les instruments. Ainsi le modèle ingérera en entrée des données réalistes, car mesurée et donnera en sortie, une ionosphère réaliste localement autour des points de mesure.

Références bibliographiques

- Chisham, G., M. Lester, S. E. Milan, M. P. Freeman, W. A. Bristow, A. Grocott, K. A. McWilliams, et al. « A Decade of the Super Dual Auroral Radar Network (SuperDARN): Scientific Achievements, New Techniques and Future Directions ». *Surveys in Geophysics* 28, n° 1 (7 juin 2007): 33- 109. <https://doi.org/10.1007/s10712-007-9017-8>.
- Grandin, M., P.-L. Blelly, O. Witasse, et A. Marchaudon. « Mars Express Radio-Occultation Data: A Novel Analysis Approach ». *Journal of Geophysical Research: Space Physics* 119, n° 12 (décembre 2014): 10,621-10,632. <https://doi.org/10.1002/2014JA020698>.
- Marchaudon, A., et P.-L. Blelly. « A New Interhemispheric 16-Moment Model of the Plasmasphere-Ionosphere System: IPIM ». *Journal of Geophysical Research: Space Physics* 120, n° 7 (1 juillet 2015): 2015JA021193. <https://doi.org/10.1002/2015JA021193>.

**Métrieologie des horloges / Metrology
aspects of clocks**

Horloges atomiques - Etat de l'art et enjeux

Atomic clocks - State of the art and challenges

N. Dimarcq¹

*¹ SYRTE – Systèmes de Référence Temps-Espace Unité mixte de recherche Observatoire de Paris, Université PSL, CNRS,
Sorbonne Université, LNE Observatoire de Paris 61 avenue de l'Observatoire, 75014 Paris, France*

[*noel.dimarcq@obspm.fr*](mailto:noel.dimarcq@obspm.fr)

*Mots clés (en français et en anglais) : Horloges atomiques, Références de temps et de fréquence
Atomic clocks, Time & Frequency standards*

Présentation en session plénière / Plenary session communication

Résumé / Abstract

Depuis leur invention au milieu du XXe siècle, les horloges atomiques n'ont cessé de s'améliorer pour atteindre aujourd'hui des niveaux exceptionnels de stabilité et d'exactitude dans la gamme des 10-18. Même si ce niveau de précision n'est pas encore nécessaire pour le positionnement et la navigation par satellite, le fonctionnement même des GNSS repose sur un ensemble d'horloges atomiques, embarquées dans les satellites ou fonctionnant au sol pour la construction de l'échelle de temps du système de positionnement considéré. L'exposé présentera un état de l'art des horloges atomiques fonctionnant dans les domaines micro-onde et optique, en décrivant de nouveaux types d'horloges atomiques en train d'émerger dans le monde industriel, dans le but d'améliorer les performances en fréquence ou la miniaturisation des dispositifs. Avec l'amélioration de la précision des horloges atomiques se pose la question de la synchronisation d'horloges distantes, qui sera aussi abordée au cours de l'exposé.

Since their invention in the mid-twentieth century, atomic clocks have continuously improved to reach outstanding levels of stability and accuracy in the 10-18 range. Even if this level of precision is not yet necessary for the localization and the navigation by satellite, the GNSS operation rests on a set of atomic clocks, embarked in the satellites or working on the ground for the construction of considered positioning system time scale. The presentation will give the state of the art of atomic clocks operating in the microwave and optical domains, describing new types of atomic clocks emerging in the industrial world, with the aim of improving the frequency performance or the miniaturization of the devices. With the improvement of the atomic clock precision, there is the question of the synchronization of distant clocks, which will be also addressed during the presentation.

Atomic clocks State of the art and challenges

Noël DIMARCO

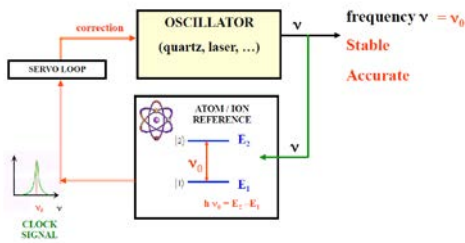
SYRTE - Systèmes de Référence Temps-Espace



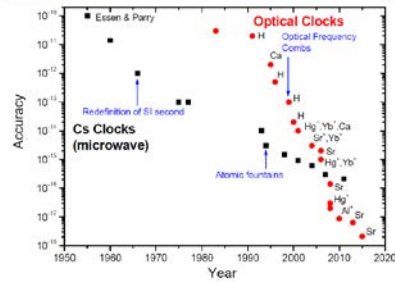
Atomic clocks - State of the art and challenges

- Introduction
- Clocks for space segment
- Clocks for ground segment
- Conclusions

Basic principle of atomic clocks / atomic frequency standards



Improvement of clock frequency accuracy



Frequency / Phase / Time errors

$$\text{Signal} = A \cdot \cos(2\pi \nu(t) t) = A \cdot \cos(q(t)) = A \cdot \cos(2\pi \nu_0 T(t))$$

Frequency: $\nu(t) = \nu_0 \times (1 + \varepsilon + y(t))$

Phase: $q(t) = 2\pi \int_0^t \nu(t') dt' = 2\pi \nu_0 \left[(1 + \varepsilon)t + \int_0^t y(t') dt' \right]$

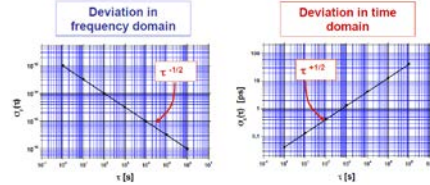
Time: $T(t) = \frac{q(t)}{2\pi \nu_0} = \frac{q(t)}{2\pi \nu_0} \cdot \frac{1}{\nu_0} = (1 + \varepsilon)t + x(t)$

with $y(t) = \frac{dx(t)}{dt} \Leftrightarrow x(t) = \int_0^t y(t') dt'$

Frequency / Time errors

$$y(t) = \frac{dx(t)}{dt} \Leftrightarrow x(t) = \int_0^t y(t') dt'$$

Case of white frequency noise:



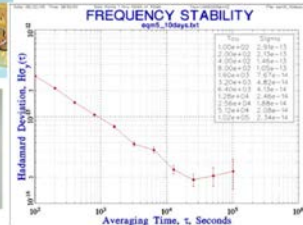
10^{-13} relative frequency error = 10 ns time error @ 1 day = 3 m position error

Clocks for space segment - Requirements

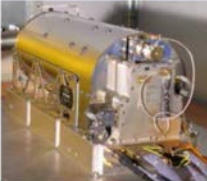
- Requirements for clocks on-board Galileo Satellites:
- < few ns at max Age of Data = 100 min (future: extended to 1 - several days)
- Predictable frequency drift
- Robust and reliable (12+ years, radiation...)



RAFS - Rubidium clocks for GALILEO

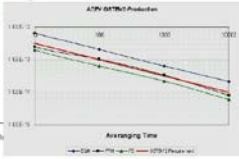


PHM – Passive H-Maser for GALILEO



Application: GALILEO
 1ns (10^{-14}) time error
 ↓
 30 cm position error

Goal: 10^{-14} stability @ 10'000 s (keeping 1 ns over one orbit)
 ↓
 10^{-12} @ 1 s





18 kg, 28 L, $7 \cdot 10^{-13}$ @ 1 s

Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Envisaged atomic clocks for next generations

→ Cesium beam clocks (with optical pumping)
 ~ $3 \cdot 10^{-12}$ $\tau^{-1/2}$ + good long term stability

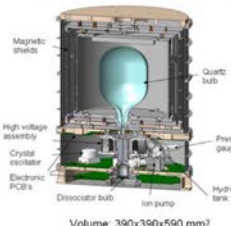
→ Upgraded Rubidium cell clocks: spectral lamps replaced by lasers (LTF, Switzerland) or pulsed operation (INRIM, Italy)
 ~ $1 \cdot 10^{-13}$ $\tau^{-1/2}$ + long term drifts (collisions with buffer gas)

→ Ion clocks in microwave domain
 ~ $1 \cdot 10^{-13}$ $\tau^{-1/2}$ + good long term stability

Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Very high stability space atomic clocks

Active hydrogen maser (ACES mission on board ISS)



SHM role in ACES
 - ACES flywheel oscillator
 - PHARAO characterization

Technical challenges
 - Low mass, volume, and power consumption
 - Full performances:
 • $1.5 \cdot 10^{-13}$ @ 1 s
 • $1.5 \cdot 10^{-15}$ @ 10^4 s

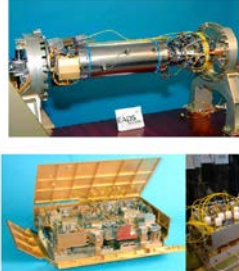
Design solution
 - Full size Al cavity
 - Automatic Cavity Tuning System (ACT)

Volume: $390 \times 390 \times 590$ mm³
 Mass: 42 kg


Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Very high stability space atomic clocks

Cold atom clock PHARAO (ACES mission on board ISS)

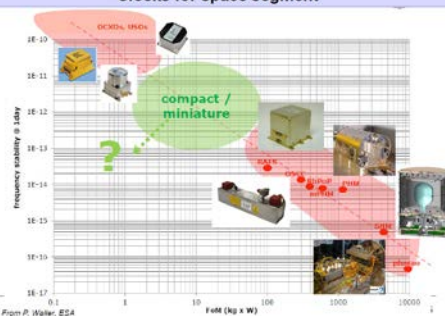


$\alpha_y(t) < 10^{-13} \tau^{-1/2}$
 Inexactitude $\delta\alpha$ au niveau de 10^{-16}



Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Clocks for space segment

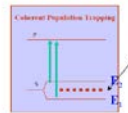
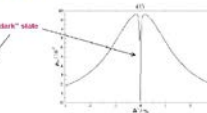


compact / miniature


From P. Walker, ESA

An efficient way for miniaturization

→ CPT – Coherent population trapping





Potential advantages of using CPT:
 • No microwave cavity
 • Reduced light-shift

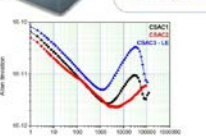


Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Miniaturized CPT clocks



16 cm³, 35 g
 120 mW
 Stability: $< 2 \cdot 10^{-16}$ @ 1 s ; $< 10^{-11}$ @ 1 jour
 Accuracy: $5 \cdot 10^{-11}$



+ Current industrial development in France (Syralinks)

Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu

Clocks for ground segment

Galileo System Time (GST)

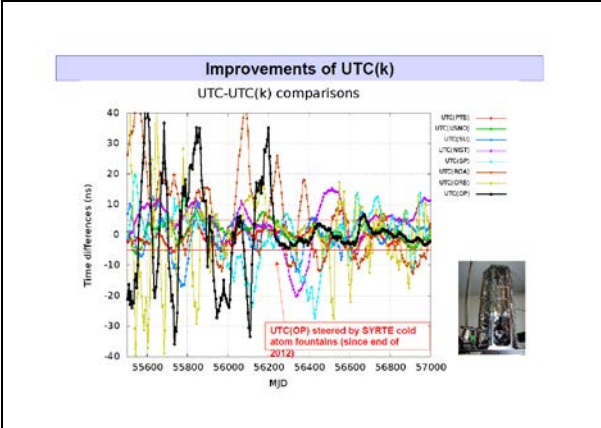
Need to have a stable ground atomic time scale to:

- Monitor et resynchronize space clocks with a time scale close to UTC (difference between GST and UTC < 50 ns, 95% of time)
 Remark: difference between time scale at ground user level and GST < 30 ns, 95 % of time
- Ensure interoperability between GALILEO and GPS (difference between time scales < 20 ns, 95% of time)

→ Realization of GST in 2 Precise Timing Facilities (Germany, Italy) equipped with an ensemble of active H-masers and commercial Cs beam clocks.

→ Link between GST and European UTC(k) realized in T/F metrology institutes to ensure accuracy and provide a long term steering of GST (contributions of cold atom fountains in France and Germany)

Journées Scientifiques CNRS-France, Mars 2015, Moulon, N. Dinu



Industrial cold atom clock: Mu-Clock

→ Objective: Replacement of H-maser + Cs beam clock

► **Frequency stability**

1s	$\pm 3.0 \cdot 10^{-11}$
10s	$\pm 9.5 \cdot 10^{-12}$
100s	$\pm 3.0 \cdot 10^{-12}$
1000s	$\pm 9.5 \cdot 10^{-13}$
10000s	$\pm 3.0 \cdot 10^{-13}$
1 day	$\pm 2.0 \cdot 10^{-13}$
Flicker floor	$\pm 2.0 \cdot 10^{-13}$ (@ 10 days)

► **Power**

Operating power	200W
Peak power	250W

► **Physical characteristics**

Dimensions	
Height	120cm
Width	51cm
Depth	40cm
Weight	75kg

Atomoptics Sciences France L'ESF France, Mars 2018, Meudon, N. Dineen

21

Conclusions and future prospects

Today: only one European provider for GALILEO space clocks (Spectratime)

- Need to maintain and reinforce the development in laboratories and in industry of new kind of clocks for space and ground segment: upgraded cell clocks, cold atoms free falling or trapped on a chip, Coherent Population Trapping, Ion clocks, ...
- Explore the possibility to install ultra stable clocks in geostationary satellites (direct calibration in space of the constellation clocks, realization of a better atomic time scale with a lower uncertainty in the relativistic gravitational correction)

Importance of ultra stable frequency standards in T/F institutes (calibration of UTC, Galileo System Time)

- Calibration with cold atom fountains provides atomic time scales with high accuracy / high long term stability (SYRTE = 40 % of the calibration over years)
- Growing role of optical clocks in the construction of better atomic time scales (Stability: few 10^{-18} @ 1s, Accuracy: 10^{-18} - 10^{-17})
- With the improvement of clocks, need of upgraded T/F links (in microwave and/or optical range) for space and ground clock synchronisation

Atomoptics Sciences France L'ESF France, Mars 2018, Meudon, N. Dineen

23



Galileo gravitational redshift test with eccentric satellites *Test de l'effet de redshift gravitationnel avec les satellite Galileo excentriques*

Delva P.¹, Puchades N.^{2,1}, Meynadier F.¹, Le Poncin-Lafitte Ch.¹, and Wolf P.¹

¹SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, 61 avenue de l'Observatoire 75014 Paris

²Departamento de Astronomia y Astrofísica - Valencia University

Keywords: Galileo; GNSS; General Relativity; Gravitational Redshift Test.

Abstract/Résumé

We are going to present the results of the analysis of the GREAT (Galileo gravitational Redshift test with Eccentric sATellites) experiment from SYRTE (Observatoire de Paris), funded by the European Space Agency. The General Relativity (GR) predicts that time flows differently for two clocks that have a relative speed and are placed in different gravitational potentials. It is therefore possible to test GR by comparing the frequencies of two clock, in a so-called gravitational redshift test. The best test to date was performed with the Gravity Probe A (GP-A) experiment in 1976 with an uncertainty of 1.4×10^{-4} .

An elliptic orbit induces a periodic modulation of the fractional frequency difference between a ground clock and the satellite clock, while the good stability of Galileo clocks allows to test this periodic modulation to a high level of accuracy. Galileo 201 and 202, with their large eccentricity and on-board H-maser clocks, are perfect candidates to perform this test. However, the accuracy of the gravitational redshift test is limited by the systematic uncertainty due to orbital errors, and Satellite Laser Ranging (SLR) measurements are crucial to understand them. SLR data allows us to reduce the effect of the systematics, in particular to partly decorrelate the orbit perturbations from the clock errors. By analyzing several years of Galileo satellites data we have been able to improve on the GP-A test of the gravitational redshift.

Implémentation matérielle du système de Synchronisation d'Horloge à Quatre Canaux pour le Convertisseur A/N du Radiohéliographe de Nançay

Hardware Implementation of Quad-channel Clock Synchronization System for Nançay Radioheliograph A/D Converter

El Houssain Ait Mansour¹, Karl-Ludwig Klein¹, Stéphane Bosse¹ and Severin Barth

¹Observatoire de Paris, Station de radioastronomie de Nançay, Route de souesmes, 18330 Nançay, France

Email: eaitmansour, ludwig.klein, bruno.dasilva, stephane.bosse, severin.barth, ludwig.klein@obspm.fr

Mots-clefs: Synchronisation des Horloges, Calibration, Implémentation Matérielle

Keywords: Clocks Synchronization, Calibration, Hardware Implementation

Résumé:

La synchronisation est un problème majeur dans les réseaux distribués et distants, y compris l'interférométrie radio. La précision du système de synchronisation dans un réseau d'antennes en interférométrie radio dépend de la bande passante analogique du signal acquis. Peu de protocoles de synchronisation existants peuvent assurer une précision inférieure à la nanoseconde. Dans ce document, on va présenter les premiers résultats obtenus d'implémentation matérielle du premier prototype d'un algorithme de synchronisation appliqué à quatre horloges.

Abstract: The synchronization is a major problem in distributed networks, including radio interferometry. The accuracy of the synchronization system in radio interferometry depends on the analog bandwidth of the acquired signal. However, few of previous synchronization protocols can provide sub-nanosecond precision. In this document, we will present the first results obtained from Hardware Implementation of the synchronization algorithm dedicated for four clocks.

1 Introduction

La synchronisation d'horloges est un problème important dans l'infrastructure des réseaux distribués, y compris l'interférométrie radio [1]. De nombreux algorithmes de synchronisation du temps ont été développés auparavant, mais peu d'entre eux peuvent atteindre la précision des sous-ns dans les réseaux distribués [2, 3, 4, 5]. En outre, plusieurs sources d'erreurs de système limitent la précision, tels que les retards asymétriques, les retards de propagation, le nombre de noeuds et d'autres sources de retards induites par la méthode de mesure. Le Radiohéliographe de Nançay [6] est un réseau d'antennes analogiques radio qui se compose de 48 antennes multi-réparties. Il permet de construire des images radio de la couronne solaire dans la bande de fréquence 150-450 MHz [7]. Les signaux Radio Fréquence (RF) sont transportés dans des câbles coaxiaux avec des longueurs différentes (50 m à 3,2 km) des antennes au récepteur. Les principales limitations de l'instrument actuel sont les erreurs de retards d'offsets introduites par l'étalonnage manuel entre chaque antenne et le récepteur. En outre, le temps de commutation entre chaque fréquence analysée introduit une latence dans les images solaires, elle réduit également le rapport signal sur bruit. Le système de synchronisation proposé du nouveau NRH numérique nécessite une fréquence d'horloge du convertisseur A/D de 1 GHz au minimum pour chaque antenne (échantillonnage pour une bande complète). Ce nouveau système permettra d'acquérir simultanément des images radio [8] à différentes fréquences. Avec une résolution configurable. Les principaux problèmes de ce système sont : la distribution de l'horloge haute fréquence requise pour chaque antenne, les différentes longueurs de câbles entre les antennes individuelles et le récepteur, qui nécessitent une synchronisation de temps globale avec une précision de l'ordre de la sous-ns.

L'objectif de ce travail est de parvenir d'illustrer les résultats préliminaires du premier prototype de 4 quatre antennes. Nous proposons, dans un premier temps, les besoins réels en performance de synchronisation demandés par le système actuel en comparaison avec les instruments existants. On terminera avec la description de l'architecture matérielle proposée et les résultats.

2 Spécification de système de synchronisation de nouveau NRH

Dans cette section, on va introduire les principales caractéristiques techniques des imageurs radios existants qui fonctionnent dans des bandes de fréquence d'observations allant jusqu'à 950 GHz (ALMA dans la bande 10). L'étude comparative se focalise sur la bande numérisée et la précision requise de synchronisation des horloges. Le tableau 1 illustre la bande numérisée ainsi des divers instruments. Elle est pour la plupart des imageurs radio inférieure à 100 MHz (Tab. 1). Elle est égale à 2 GHz pour les deux instruments ALMA (Atacama Large Millimeter/submillimeter Array) et EVLA (Expanded Very Large Array) avec une bande d'observation allant de 31,3 GHz à 950 GHz. Ce choix est dû à la complexité technique du système du numériseur, distribution et synchronisation des horloges hautes fréquences.

TABLE 1 : La bande de base des instruments radio

Instrument	Bande numérisée (MHz)
NRH	0,7
LOFAR	30-250
EMBRACE	30-250
GMRT	16
ATA [9, 10]	100
EVLA [11, 12]	2000
ALMA [13, 14, 15]	2000
LWA [16, 17]	80

La précision minimal d'ajustement des horloges de numériseur (Convertisseurs A/D) des instruments existants est illustrée dans la Table. 2. On constate que la précision augmente avec l'augmentation de la bande numérisée : LOFAR (78 ps pour une bande maximale de 250 MHz), ALMA (15,62ps, 2 GHz). On conclut une numérisation large bande consiste une contrainte majeure de la mise en oeuvre du système de synchronisation des horloges multi-reparties. On déduit, d'après la Table. 2 que la précision requise par le nouveau NRH est égale à 50,74 ps.

TABLE 2 : La précision requise des horloges pour les divers instruments

Instrument	Précision (ps)
RH	23364
LOFAR	78
EMBRACE	78
EVLA	15,62
ALMA	15,62
Nouveau RH : Par rapport à ALMA/EVLA	$15,62 \cdot 500/2000 = 62,48$
Nouveau RH : Par rapport à LOFAR/EMBRACE	$78 \cdot 250/500 = 39$
Nouveau RH : Moyenne	50,74

La distribution des horloges haute fréquence à faible gigahertz pour les convertisseurs A/D intégrés aux pieds des antennes nécessite une synchronisation globale du système. La désynchronisation est induite par la différence de longueur entre les antennes (50 m-1200 m) et le système centrale ainsi à l'asymétrie des canaux. Le NRH actuel est constitué d'un réseau de 48 antennes radio multi-reparties [18] dédiée à l'imagerie de la couronne solaire. D'après l'étude [19], les courbes de la figure 2 illustrent l'évolution de l'erreur différentielle $d\tau_g$ en fonction de tau de corrélation pour chaque fréquence. Le tau de corrélation $r(\tau_g)(\%)$ est normalisé par rapport à la puissance. Il diminue avec l'augmentation de l'erreur par contre l'erreur diminue avec l'augmentation de la fréquence. La précision du système de synchronisation (décalage entre les horloges) à 5% de perte de tau de corrélation est égale à 112 ps, elle est à 50 ps à 1 % (figure 2) pour une fréquence maximale de 450 MHz. La précision requis par le système à 1% est quasiment identique à la précision déduite en comparaison avec les imageurs (voir la Table. 2).

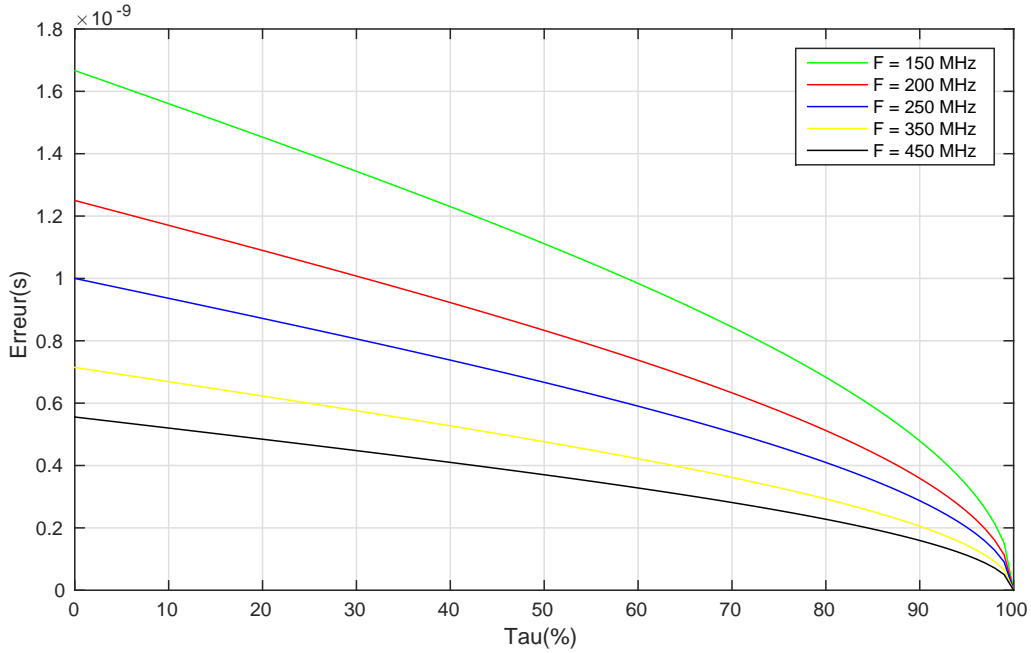


FIGURE 1 : L'erreur différentielle $d\tau_g(s)$ en fonction de tau de corrélation $r(\tau_g)(\%)$ [19]

3 Description de l'architecture proposée à quatre canaux

Le système complet de calcul et de l'ajustement pour 4 canaux (horloges) indépendants (4 antennes/esclaves) en est illustré dans la figure 2. L'objectif du système est la mesure et l'estimation en temps réel des écarts d'ajustement puis l'application de ses derniers pour chaque horloge en se basant sur l'algorithme illustré dans les travaux [20, 21] effectués précédemment. La boucle de mesure et d'ajustement est constituée des éléments suivants :

- un TDC (Time-to-Digital Converter)[22, 23, 24] : Un circuit de mesure du retard aller/retour d_{MM}^i pour chaque canal.
- un DTC (Digital-to-Time Converter)[25] : Plusieurs circuits d'ajustement des offsets des horloges de hautes précisions.
- un bloc d'acquisition : Une Unité d'acquisition des retards mesurés pour chaque canal et le transfert des mesures à l'ordinateur pour l'analyse et le traitement différés. Au même temps, l'interface d'acquisition est utilisée pour la mesure en temps réel des retards.
- un bloc SyncEv : Générateur des intervalles de temps T_i programmable et du signal "Start" (Sync) permet de synthétiser des intervalles du temps fixes durant la phase de calibration.
- un bloc Calibration Calcul & Ajustement : Unité de calibration des retards, des calculs des écarts d'ajustements et du transfert des écarts au circuit à retard programmable (DTC)
- Slave : Noeud dans le réseau modélisant une antenne.

La calibration est la première phase ((1), Figure. 3) qui permet de compenser l'effet des retards accumulés dans le chemin aller-retour des horloges. La figure 4 illustre les différents délais dans un trajet aller-retour entre le maître et l'esclave. On distingue trois délais : T_i : Intervalle du temps générée par FPGA, la durée de ce dernier est connue d'avance elle est générée par le générateur des événements, d_0^i : La somme des retards instrumentales accumulés dans la chaîne, due généralement au temps de propagation des divers composants dans le chemin aller-retour, d_1^i : Retard induit par la longueur du câble, d_{MM}^i : Mesure du retard global (du maître au maître) du circuit TDC et la somme des retards dans le chemin. Durant la phase de calibration, l'effet du retard du câble d_1^i est supprimé en utilisant des câbles avec des longueurs connues et identiques conçus spécialement à cet effet dans le chemin aller-retour des signaux START et STOP. Par conséquent, la mesure du TDC d_{MM}^i (Eq. 1) est réduite à $T_i + d_0^i$, donc on peut facilement déduire la valeur d_0^i à partir de la mesure de d_{MM}^i et par la

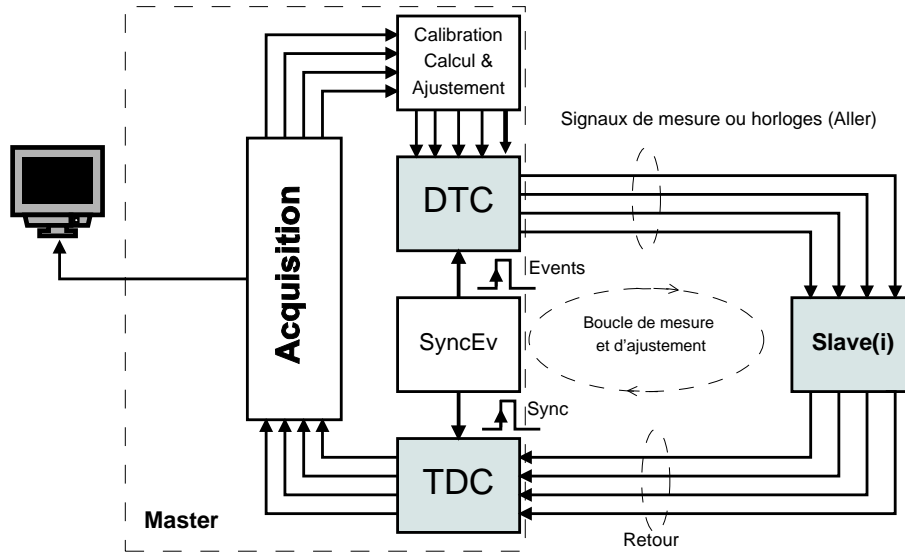


FIGURE 2 : Schéma bloc du système de mesure et d'ajustement des offsets des horloges.

connaissance préalable du T_i . L'opération de tri ((2), Figure. 3) est la première étape durant l'implémentation sur de l'algorithme de calcul des écarts d'ajustement d_{xi} . Pour objectif de trier les mesures de telle manière à obtenir des valeurs différentielles positives pour chaque couple d'antennes $(i, i + i)$. Par conséquent, obtenir des écarts d_{xi} (3) positifs qui seront intégrés dans un circuit à retard programmable de haute précision (DTC)(5).

$$d_{MM}^i \sim T_i + d_0^i \quad (1)$$

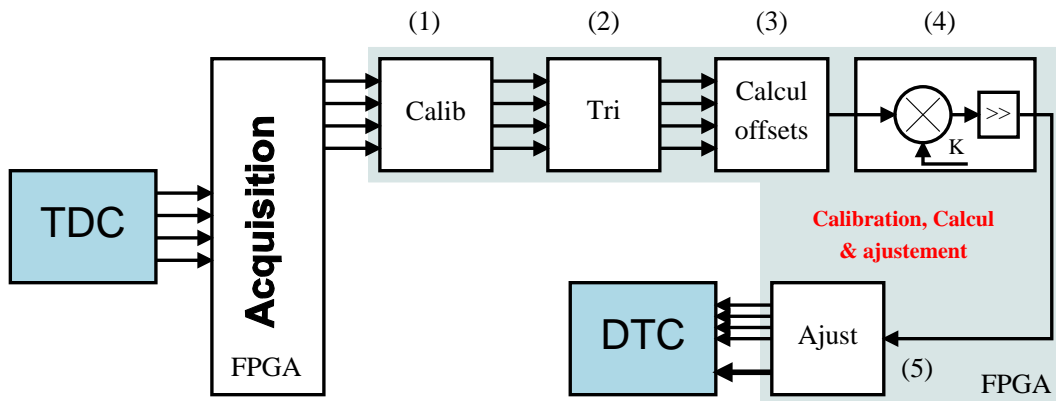


FIGURE 3 : Schéma bloc de circuit de calcul et ajustement des horloges. (1) : Circuit de calibration. (2) : Tri des mesures pour le circuit de calcul d'ajustement afin d'obtenir des écarts d'ajustements positifs. (3) : Algorithme de calcul des offsets fin d'obtenir la synchronisation des canaux. (4) : Conversion des mesures du TDC au DTC (adaptation des résolutions binaires). (5) : Transfert multiplexé des écarts d'ajustement aux circuits DTC.

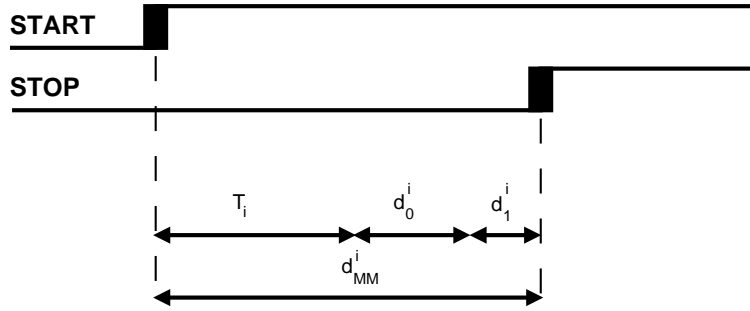


FIGURE 4 : Chronogrammes des retards dans un trajet aller-retour.

4 Résultats

Pour valider le fonctionnement du système illustré dans la Figure. 2, on considère les liaisons de même distances entre le maître et les esclaves. La longueur du câble du signal "Start" (Sync) entre le FPGA et le TDC est égale à la longueur du câble aller+retour du signal Stop afin de réduire l'erreur due aux câbles des deux signaux. Le maître envoie des intervalles de temps T_i (l'intervalle du temps entre le front montant des signaux "Start" et le front montant du "Stop") contrôlés par la carte FPGA (générateur des T_i) pour les quatre canaux, puis on récupère les résultats des mesures via l'interface d'acquisition. On a effectué des fenêtres de mesure sur 10 secondes pour les divers modes de mesures (voir la Table. 4). Le maître utilisée une horloge de référence à 200 MHz avec un "skew" inférieure à 50 ps (6 sigma) comme il es illustré dans la Figure. 5.

La Figure. 6 montre l'estimation moyenne du retard induit par la chaîne en fonction du mode pour chaque canal. Le retard est compris entre 250 ps et 5 ns. Il est minimal pour le canal 1 (A) et maximal pour le canal 4 (D). Le retard croit en fonction du canal et du mode comme il est illustré précédemment. La précision maximale du système après calibration est inférieure à 2 ps et quasiment constant en fonction des modes avec un skew inter-canaux inférieure à 0,5 ps.

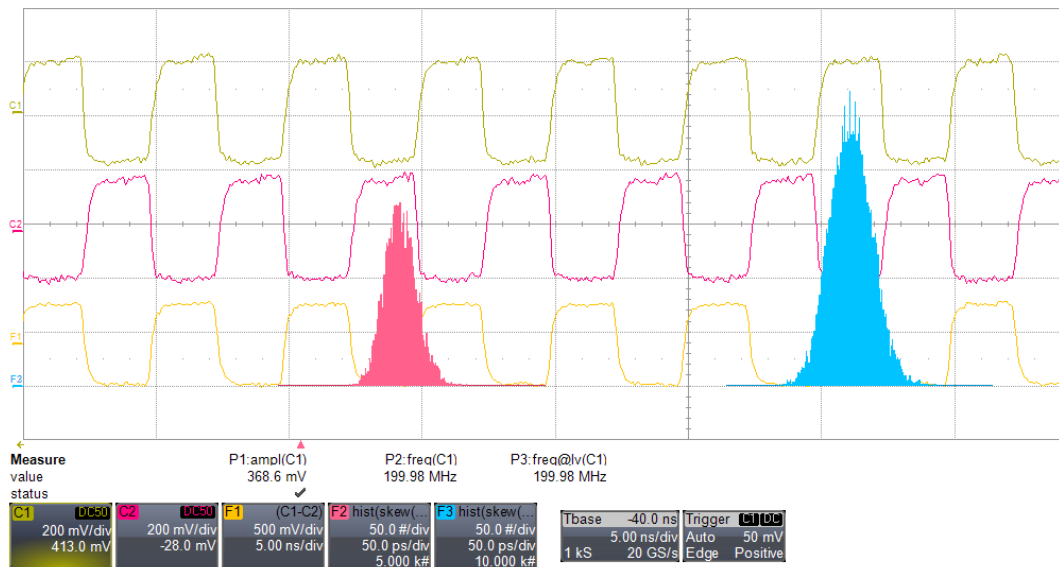


FIGURE 5 : Gigue de l'horloge de référence au niveau du maître à 200 MHz. La distribution (rose) indique le skew entre C1 et F1 (50 ps/div, 200 mV/div). La distribution (bleue) indique le skew entre C2 et F1 (50 ps/div, 500 mV/div).

Les résultats du système de calibration montrent la compensation des divers retards induits par la chaîne de mesure. On a constaté également la compensation des retards induits par le générateur des intervalles de temps pour chaque canal. La précision du système après calibration est inférieure à 3 ps dans le cas de d'utilisation d'une horloge avec une gigue de 50 ps (6 sigma) (voir la figure.5). De plus, le circuit de calibration permet d'obtenir des mesures précises malgré l'utilisation des signaux bruyant (Start et Stop avec une gigue supérieure à 200 ps (6 sigma)) et le non respect de la contrainte sur l'horloge du circuit TDC (gigue inférieure à 3 ps). La Table.

3 résume les performances et limitations de la technique proposée en comparaison avec les autres méthodes de synchronisation. Note méthode est plus compatible et optimisée à la topologie des antennes du NRH (étoile). Les simulations montrent une précision inférieure à 50ps avec la minimisation de l'effet de l'asymétrie du canal. La précision préliminaire mesurée est de l'ordre de la nanoseconde de l'architecture globale du système de synchronisation

TABLE 3 : Caractéristiques de la technique proposée. (*) : Précision Simulée dans le cas symétrique, (**) : Précision Simulée dans le cas asymétrique, (***) : Précision mesurée sans minimisation de l'erreur.

	Complexité (1)	Minimisation (2)	Asymétrie (3)	Type de mesure	Adaptabilité	Précision
Technique proposée	Minimal	Ajustement périodique Avec minimisation	Réduction de l'effet de l'offset du à l'asymétrie	Mesure unique au niveau du maitre	Optimisé pour la topologie étoile des antennes	Simulée(*) : 0 ps Simulée(**) : <30 ps Mesurée(***) : < 2 ns
WR	Élevée Couches 1,2	Ajustement périodique	Supposé symétrique	Deux mesures	Générique	< 500 ps Augmente avec la longueur
PPT	Élevée Couches 1,2	Ajustement périodique	Supposé symétrique	Deux mesures	Générique	350 ns Augmente avec la longueur
PPS	Élevée Couches 1,2	Ajustement périodique	Supposé symétrique	Deux mesures	Générique	100 ns Augmente avec la longueur

TABLE 4 : Intervalles du temps $T_i(\mu s)$ pour les quatre canaux A à D en fonction du mode du générateur de signaux Start et Stop

Mode/Canal	A	B	C	D
2	5.12	10.24	15.36	20.48
3	10.24	20.48	30.72	40.96
4	20.48	40.96	61.44	81.92

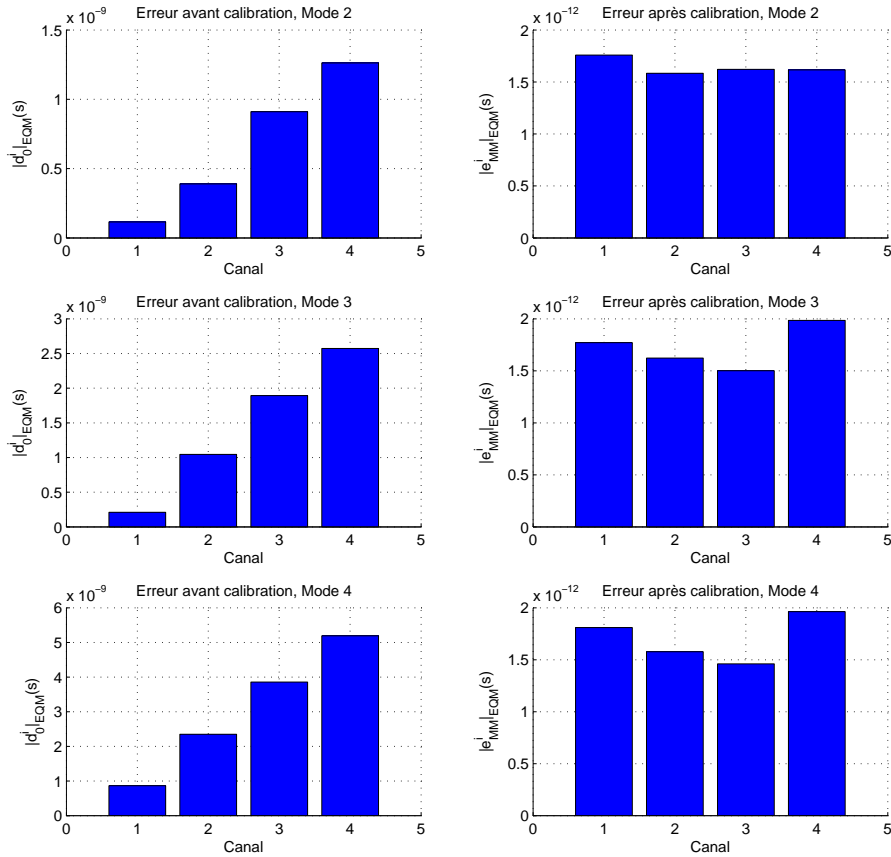


FIGURE 6 : Erreur quadratique moyenne(EQM) avant et après la phase de calibration pour les modes (2,3 et 4) pour les quartes canaux.

5 Conclusion

Les résultats du système de synchronisation en particulier la calibration montrent la compensation des divers retards induits par la chaîne de mesure. On a constaté également la compensation des retards induits par le générateur des intervalles de temps pour chaque canal. La précision du système après calibration est inférieure à 3 ps dans le cas de d'utilisation d'une horloge avec une gigue de 35 ps (1 sigma). Elle est de 2 ps dans le cas d'utilisation d'une horloge ayant une gigue inférieure à 5 ps (1 sigma). On conclut que la précision après calibration augmente avec la qualité de l'horloge utilisée. De plus, le circuit de calibration permet d'obtenir des mesures précises malgré l'utilisation des signaux brouillant (Start et Stop avec une gigue supérieure à 200 ps (6 sigma)) et le non respect de la contrainte sur l'horloge du circuit TDC (gigue inférieure à 3 ps). La précision préliminaire mesurée est de l'ordre de la nanoseconde de l'architecture globale du système de synchronisation sans minimisation [20]. La précision finale sera meilleure avec l'intégration de l'algorithme de minimisation de l'erreur en utilisant la structure LMS (Least Mean Square)[26, 27]. De plus, les circuits de mesure et ajustement (TDC et DTC) ne sont pas optimaux, une Implémentation sur FPGA (IP disponible) avec une précision inférieure à 20 ps peut augmenter la précision et rendre le système plus flexible et performant [28, 29].

6 Références

- [1] A. T. S. T. M. Colegate and R. B. Wayth, "Characterization of a low-frequency radio astronomy prototype array in western australia," *IEEE Transactions on Antennas and Propagation*, 2015.
- [2] H. Weibel, "High precision clock synchronization according to IEEE 1588 - implementation and performance issues," *Embedded World*, pp.22-24, 2005.
- [3] I. F. I. Sousa and A. Klautau, "Analysis and evaluation of end-to-end ptp synchronization for ethernet-based fronthaul," *Global Communications Conference (GLOBECOM), 2016 IEEE*, 2016.

- [4] A. M. R. Exel and T. Sauter, "Delay and jitter characterization for software-based clock synchronization over wlan using ptp," *IEEE Trans. Industrial Informatics* 10 (2) : 1198-1206, 2014.
- [5] C. S. S. Zug and J. Kaiser, "Uncertainly aware hybrid clock synchronization in wireless sensor networks," *UBICOMM2014 : The Eighth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, 2014.
- [6] K.-L. Klein and A. Kerdraon, "Solar physics at nançay radio observatory : Recent developments," *IEEE General Assembly and Scientific Symposium*, 2011.
- [7] A. Kerdraon and J.-M. Delouis, "Coronal physics from radio and space observations," *IEEE General Assembly and Scientific Symposium*, 2011.
- [8] A. K. M. P. S. H. Y. Wang and D. Haggerty, "The coronal and heliospheric 2007 may 19 event : coronal mass ejection, extreme ultraviolet imager wave, radio bursts, and energetic electrons," *The American Astronomical Society*, 2010.
- [9] D. B. A. Bauermeister and al., "The allen telescope array the first widefield, panchromatic, snapshot radio camera," *Radio Astronomy Laboratory, University of California Berkeley CA, USA*, 2009.
- [10] W. Urry, "The ata imager update," *URSI Meeting, Boulder, CO*, 2004.
- [11] R. P. P. N. J. J. B. B. Carlson and al., "The expanded veryz large array," *Proceedings of the IEEE / Vol. 97, No. 8*, 2009.
- [12] R. Perley, "The expanded very large array new capabilities for new science," *National Radio Astronomy Observatory NRAO*, 2012.
- [13] R. E. Hills and A. J. Beasley, "The atacama large millimeter/submillimeter array," *ALMA, Av El Golf 40, Piso 18, Santiago, Chile*, 1995.
- [14] A. W. . K. Y. Lo and al., "The atacama large millimeter/submillimeter array development program," *Activity White Paper for Astro2010 Program Prioritization Panel*, 2010.
- [15] L. R. D'Addario, "Timing and synchronization," *ALMA Memo No. 298*, 2000.
- [16] S. E. T. Clarke and al., "The long wavelength array," *Proceedings of IEEE*, 2009.
- [17] G. B. Taylor, "Science with the long wavelength array," *The Metrewavelength Sky ASI Conference Series, 2014, Vol. 13, pp 349 354*, 2014.
- [18] T. R. Hunter and P. J. Napier, "Antennas and receivers in radio interferometry," *Proceedings of the 2014 Synthesis Imaging Workshop held in Socorro, NM on May 13-20*, 2014.
- [19] E. A. M. K.-L. K. B. D. Silva and al., "Nancay radioheliograph a/d converter specifications for solar corona observation and clocks synchronisation system accuracy computation," *Observatoire de Paris, Station de Radioastronomie de Nancay, Projects : Time Synchronizatio*, 2016.
- [20] E. A. M. B. D. Silva and al., "Sub-nanosecond time synchronization mechanism for radio interferometer array," *2017 IEEE International Workshop on Measurement and Networking (M and N) 27-29 Sept*, 2017.
- [21] E. A. M. B. D. Silva and al., "Global time synchronization mechanism for radio interferometer network," *2017 Joint Conference of the European Frequency and Time Forum and IEEE International Frequency Control Symposium (EFTF/IFCS) 9-13 July*, 2017.
- [22] G. S. Jovanovic and M. K. S. cev, "Verniers delay line time to digital converter," *Scientific Publications of the State University of Novi Pazar ser a Appl Math Inform and Mech vol 1, 1 2009, 11 :20*, 2009.
- [23] C. H. Lin and M. Syrzycki, "Single stage vernier time to digital converter with sub gate delay time resolution," *Circuits and Systems, 2011, 2, 365 :371*, 2011.
- [24] T. Instruments, "Ths788 quad-channel time measurement unit (tmu)," *SLOS616D*, 2015.
- [25] O. Semiconductor, "Mc10ep195, mc100ep195 programmable delay chip,"
- [26] C. S. T. O. et al., "Fpga implementation of lms-based fir adaptive filter for real time digital signal processing applications," *2015 IEEE International Conference on Digital Signal Processing (DSP)*, 2015.

- [27] X. D. H. L. et al., "High-speed fpga implementation of an improved lms algorithm," *2013 International Conference on Computational Problem-Solving (ICCP)*, 2013.
- [28] J. Z. P. C. et al., "Low-cost fpga tdc with high resolution and density," *IEEE Transactions on Nuclear Science, Journal of Instrumentation*, vol. 12, pp. P07028, 2017, ISSN 1748-0221., 2017.
- [29] J. K. Y. W. et al., "A 128-channel high performance time-to-digital converter implemented in an ultrascale fpga," *NSS Poster Session I*, 2017.

Transfert de temps par lien laser spatial

Alexandre Belli, Pierre Exertier

UMR 7329 Geoazur, CNRS-OCA-UNS

Le transfert de temps par laser est une technique qui s'est développée sur la base de la technique spatiale de télémétrie laser d'une part, et sur les développements au sol et pour l'Espace du secteur temps-fréquence d'autre part. Entre 2008 et 2018 plusieurs expériences spatiales ont vu le jour en Europe, en Chine et aux Etats-Unis. Sur la sonde lunaire américaine LRO en mode une voie, sur le satellite Jason-2 à 1336 km d'altitude en combinant les modes une et deux voies (expérience T2L2 du CNES et de l'OCA), et enfin sur les satellites de navigation Beidou entre 20,000 et 36,000 km. En outre, l'Allemagne développe un système équivalent à T2L2 qui fonctionnera en simple photon sur la station spatiale internationale à partir de 2019/20.

Toutes ces expériences spatiales ont des objectifs scientifiques différents (physique fondamentale, géodésie, échelles de temps, etc.) mais sont toutes essentielles au domaine de la métrologie et à la navigation. Nous décrivons et comparons les performances de leur instrumentation en termes de précision et de stabilité. L'expérience T2L2, dont la durée de vie atteint 10 ans en 2018, a permis d'établir plusieurs campagnes de transfert de temps sol-espace et sol-sol via le satellite entre plusieurs sites géodésiques équipés de télémètres laser, d'horloges atomiques (des masers à hydrogène, essentiellement), de systèmes de distribution des signaux de temps et de récepteurs géodésiques GNSS. Nous décrivons les performances atteintes et les applications qui en découlent, eu égard aux avancées récentes sur les horloges et les nouveaux réseaux fibrés au sol de transfert de fréquence.

Author Index

- Adirosi Elisa, 8–14
Ait Mansour El Houssain, 179–187
- Bétaille David, 128–138
Baldini Luca, 8–14
Barbaresco Frederic, 73–80
Barbot Jean-Pierre, 120–127
Barth Severin, 179–187
Belli Alexandre, 188
Benlarbi-Delaï Aziz, 95–101
Berbinau Marion, 128–138
Bernard Loic, 59–61
Bianchi Pascal, 102–106
Blelly Pierre-Louis, 2–5, 164–172
Bonnabel Silvère, 69–80
Bosse Stéphane, 179–187
Bouhleb Nizar, 36–43
Boutteau Rémi, 147–156
Briole Pierre, 54–58
Brousseau Christian, 36–43
- Campbell Mark, 81–88
Capderou Michel, 67
Chandra Madhu, 6, 7
Clark Daniel, 81–88
Collardey Sylvain, 59–61
Coru Guillaume, 147–156
Cuccoli Fabrizio, 8–14
- Da Silva Bruno, 179–187
Daquin Jérôme, 67
De Doncker Philippe, 95–101
Deleflie Florent, 67
Delva Pacôme, 178
Denis Benoît, 115–119
Dimarcq Noel, 174–177
- Elgui Kevin, 102–106
Exertier Pierre, 188
Eynard Goulven, 139–144
- Facheris Luca, 8–14
Fleury Bernard, 90–94
Fleury Rolland, 158–163
Foucault Etienne, 164–172
Fouettilloux Romain, 44–47
Fuerxer Pierre, 62–66
- Gailliard Gregory, 139–144
García Gámez Laura, 59–61
Garzelli Andrea, 8–14
- Härrri Jérôme, 115–119
- Heurguier Dominique, 139–144
Hoang Gia Minh, 115–119
Horlin François, 95–101
- Isnard Jean, 145, 146
- Jain Ankit, 158–163
- Klein Karl-Ludwig, 15–17, 179–187
- Lacrouts Olivier, 44–47
Le Chevalier François, 19–27
Le Poncin-Lafitte Christophe, 178
Lemer Alain, 139–144
Lupidi Alberto, 8–14
- Méric Stéphane, 36–43
Mahdjoubi Kouroch, 59–61
Marais Juliette, 128–138
Marchaudon Aurélie, 164–172
Marque Christophe, 15–17
Menudier Cyrille, 28–35
Merriaux Pierre, 147–156
Meynadier Frédéric, 178
Milbrandt Jorick, 28–35
Monstein Christian, 15–17
Motyl Guillaume, 44–47
Moullec Claude, 36–43
- Ney Michel, 158–163
- Pagani Pascal, 158–163
Pajusco Patrice, 158–163
Perosanz Felix, 49–53, 67
Petit Alexis, 67
Petrov Nikita, 19–27
Pilté Marion, 73–80
Potier Patrick, 59–61
Pouliguen Philippe, 59–61
Puchades Neus, 178

Quitin Francois, 95–101

Roberto Nicoletta, 8–14

Samama Nel, 107–114

Sarrazin Julien, 95–

101Sauleau Ronan, 59–

61 Savatier Xavier, 147–
156

Taillandier-Loize Thierry, 107–114

Tanzi Tullio, 145, 146

Thevenot Marc, 28–35

Van Der Vorst Thomas, 95–101

Van Eeckhaute Mathieu, 95–101

Vauchey Vincent, 147–156

Vervisch-Picois Alexandre, 107–114

Wolf Peter, 178

Ywanne Frederique, 139–144

Zhu Ni, 128–138

AVEC LE SOUTIEN DE :
WITH THE SUPPORT OF:



INSTITUT DE FRANCE
Académie des sciences



COMITÉ NATIONAL FRANÇAIS DE RADIOÉLECTRICITÉ SCIENTIFIQUE
UNION RADIO SCIENTIFIQUE INTERNATIONALE

Siège social : Académie des Sciences, 23 quai de Conti, Paris 6^{ème}

Site Internet : <http://www.ursi-france.org>

Adresse postale : Alain Sibille, **Secrétaire général d'URSI-France**,
Télécom ParisTech, 46 rue Barrault, 75634 Paris Cedex 13

Téléphone : 01 45 81 70 60

Courriel : contact@ursi-france.org