

**GDR Robotique  
GT2 – Véhicules terrestres  
journée du 23 janvier 2009**

**ENSAM  
151 boulevard de L'Hôpital  
75013 PARIS**

**Amphi Bézier**

–

**Thématique de la journée : Localisation**

***Ordre du jour***

**10h00** : Introduction

**10h15** : Rafael Toledo (LCPC)

Améliorations in navigation and map matching by means of fusing GNSS, dead-reckoning and enhanced maps

**10h55** : Keerthi Narayana (ENSMP, CAOR)

Localisation for Mobile Mapping Systems, Experimental results and analysis on rural and urban environments

**11h35** : F. Souvannavong (Magellium/CNES)

A visual odometry system for short term self localization.

**12h15** : pause déjeuner

**13h45** : J. Courbon (LASMEA)

Localisation dans une mémoire visuelle topologique pour la navigation de robots mobiles

**14h25** : Cindy Capelle (LORIA)

Modèle virtuel 3D urbain pour l'égo-localisation de véhicules intelligents.

**15h05** : Pause

**15h20** : Joan Solà (LAAS)

localisation et la cartographie pour la robotique mobile par moyens purement visuels

**16h00** : Patrick Rives (INRIA)

Méthodes de SLAM visuel

**16h40** : Conclusions

---

## RESUMES DES PRESENTATIONS

### Rafael Toledo (LCPC)

**title:** ameliorations in navigation and map matching by means of fusing gnss, dead-reckoning and enhanced maps

**Auteurs :** Rafael Toledo, F Peyret, D. Bétaille, (LCPC)

#### Résumé :

Nowadays, it is common that road vehicle navigation systems employ maps to represent the vehicle positions in a local reference. The most usual process to do that consists in the estimation of the vehicle positioning by fusing GNSS (Global Navigation Satellite System) and some other aiding sensors data, and the subsequent projection of these values on the map by applying map-matching techniques. However, it is possible to benefit from the map information also during the process of fusing data for positioning. In this presentation we will show our experiences with an algorithm for lane level road vehicle navigation that integrates GNSS, dead-reckoning (odometry and gyro) and map data in a unique fusion/matching process. Apart from the improvements in navigation, the proposed method brings some benefits for map-matching at lane level because, on the one hand, it allows the tracking of multiple hypothesis and on the other hand, it provides probability values of lane occupancy for each candidate segment. However, in order to employ the map information for the purpose of navigating at the lane level, maps must be enhanced. An enhanced map (Emap) must contain, on the one hand, more detailed information about the road geometry (to solve the problem of lack of contents), and on the other hand, its geometrical description of the road must be more precise than the one stored in a standard map. A new paradigm to describe lanes as piece-wise sets of clothoids was applied in the elaboration of our Emap. The process of its creation will also be presented.

### Keerthi Narayana (CAOR)

**Titre :** Localisation for Mobile Mapping Systems, Experimental results and analysis on rural and urban environments

#### Authors:

Keerthi Narayana K. S, Sungwoo Choi, François Goulette  
CAOR – Centre de Robotique, Mathematiques et Systemes, Mines ParisTech  
Email : prenom.nom@mines-paristech.fr

#### Summary:

Context of MMS and localisation function Mobile mapping systems (MMS) exist since early 1990s and at least 40 such systems are available around the world as either research prototypes or as commercial systems. They provide the mean to digitize the surrounding environment of the trajectory used by the moving mapping platform. The moving platform or the vehicle comprises of several sensors, which help to localise the vehicle (localisation) and then map the geo spatial data (geo referencing).

In the domain of MMS, the functionality of determining the position and the attitude of the moving vehicle at each given instant is performed by the Position Orientation System (POS). Several localisation sensors primarily GPS receivers and Inertial Sensors are used to obtain this functionality.

#### Experimental works and Results

We have been designing and developing an MMS platform called LARA-3D. Currently we are

working on improving and enhancing the POS functionality of this MMS platform.

In this talk we present some of the data acquisitions conducted using LARA-3D, in both rural and urban environments of France. We carried out the analysis of the trajectories in the two dimensional space and also with the altitude. The results show the problems faced by most MMS when the data collections are carried in the GPS signal lacking environments such as the streets with tall buildings (urban) and the roads covered by foliage (rural) etc. The problems of obtaining a smoother altitude and correcting the drift in the orientation are the two primary issues to address. Further on, we also present a post processing technique designed to correct these drawbacks and then we discuss the results.

## F. Souvannavong (Magellium/CNES)

**Titre** : A visual odometry system for short term self localization

**Auteur** : F. Souvannavong

**Résumé** : We present a visual odometry system for short term self localization. It targets autonomous planetary exploration in an unknown rough environment. The singularity of the method is to estimate long elementary displacements at low frequency, 0.1Hz against more than 1Hz usually. Our algorithm is evaluated in a Mars- like visual environment, on short and long trajectories, as a whole and step by step. We provide preliminary characteristics of its performances that can be used to design an autonomous locomotion system.

**Key words**: Robotics Autonomy; Imaging Technologies; Rover Localization; Visual Odometry; ExoMars mission.

## J. Courbon (LASMEA)

**Titre** : Localisation dans une mémoire visuelle topologique pour la navigation de robots mobiles (LASMEA)

**Auteurs** : J. Courbon<sup>\*2</sup>, Y. Mezouar\*, L. Eck<sup>2</sup>, P. Martinet\*

\* GRAVIR-ROSACE, LASMEA, Clermont-Ferrand

<sup>2</sup> SRI/LTC, CEA-List, Fontenay-Aux-Roses

**Résumé** :

Dans cette présentation, nous étudions les stratégies de localisation dans une mémoire d'images organisée de façon topologique et leur mise en œuvre.

Lors d'une phase téléopérée ou lors d'une exploration autonome, les images sont acquises par la caméra embarquée. Une phase de sélection et organisation permet de mémoriser certaines de ces images (les images clés) sous la forme d'une carte topologique. Par localisation nous entendons la localisation initiale c'est-à-dire quelle est l'image de la mémoire correspondant le mieux à l'image courante acquise par la caméra embarquée. Dans le cas d'images omnidirectionnelles, les localisations utilisant un descripteur global de l'image, des descripteurs locaux et une approche hiérarchique sont comparées en terme de temps de calcul, taille mémoire des données mémorisées et résultats.

La mise en œuvre pratique de la navigation par vision nécessite une bonne gestion des données mémorisées, un accès rapide à ces données ainsi que des outils permettant d'utiliser ces données notamment dans la phase de localisation. Un logiciel (SoViN) a été créé afin de répondre à ces besoins et sera brièvement présenté. Des expérimentations sur robot Cycab en environnement extérieur et sur robot Pioneer en environnement intérieur, utilisant les descripteurs précédemment

proposés seront exposées.

Des travaux sont en cours pour utiliser ce principe à la navigation dans les 3 dimensions de l'espace d'un robot aérien de type VTOL, travaux qui permettront entre autres de vérifier la robustesse des algorithmes de localisation et éventuellement de les améliorer.

### **Cindy Capelle (LORIA)**

**Titre :** modèle virtuel 3D urbain pour l'égo-localisation de véhicules

**Auteurs :** Cindy Capelle (LORIA) , Maan El Badaoui El Najjar (LAGIS)

#### **Résumé :**

La localisation d'un véhicule repose souvent sur l'utilisation d'un récepteur GPS. Cependant, en milieu urbain, le service GPS peut s'avérer intermittent et non fiable. Pour cette raison, il est utile de considérer d'autres capteurs.

La tâche de localisation est alors réalisée par le couplage d'un système de géo-référencement satellitaire avec des capteurs de localisation à l'estime (odomètres et gyromètre). L'outil de fusion de données multi-capteurs utilisé est un filtre de Kalman sans parfum (UKF).

Cependant, en milieu urbain, l'indisponibilité du service GPS peut être longue. Sans recalage précis, la navigation à l'estime seule risque alors de diverger à long terme. Un GPS hybridé seul ne peut donc pas être utilisé pour un positionnement précis et robuste.

Nous proposons alors d'étudier l'apport d'un modèle virtuel 3D urbain pour la localisation des véhicules. Pour cela, la position du véhicule est calculée en recalant le modèle 3D sur l'image courante acquise par une caméra embarquée. Dans ce cadre, la localisation repose sur la recherche d'amers qui structurent l'environnement. Les amers (points de Harris) détectés dans l'image courante sont ensuite mis en correspondance avec ceux du modèle 3D. A partir des appariements 2D/3D, la pose du véhicule est calculée en minimisant l'erreur de reprojection des points 3D sur l'image 2D.

### **Joan Solà (LAAS-CNRS)**

**Titre :** localisation et la cartographie pour la robotique mobile par moyens purement visuels

**Auteurs :** Joan Solà, Michel Devy, Simon Lacroix (LAAS)

Dans cet exposé nous allons aborder le sujet de la localisation et la cartographie pour la robotique mobile par moyens purement visuels. Depuis déjà plusieurs années, la problématique de la localisation a été principalement résolue par deux techniques en origine lointaines, mais en fait, lors de sa réalisation par vision, très liées entre elles: optométrie visuelle (OV) et SLAM (cartographie et localisation simultanées) visuel (V-SLAM).

Pour l'OV, l'on prend une tête stéréo, et par optimisation non linéaire (Ajustement de Faisceaux ou d'autres méthodes), on retrouve les transformations 3D subies par la tête entre deux (ou plus) prises d'image stéréo. L'intégration de tous ces incréments 3D produit une trajectoire 3D, laquelle, si les conditions visuelles le permettent (scène principalement rigide avec des motifs texturés), peut arriver à être d'une très haute précision -- moins de 10m d'erreur après 10km de course en milieux urbain [Comport et al., 2007] ou extérieur [Konolige et al., 2007].

D'autre part, le V-SLAM prends son point de partie des méthodes de filtrage non linéaire. Un filtre non linéaire est dédié à estimer la position du robot et la position de tous les amers extérieurs perçus. La structure de l'environnement (la liste des positions de tous les amers) est inconnue à

priori, et le SLAM doit donc rajouter aux opérations classiques de filtrage (prédiction pour le mouvement du robot; et correction lors des observations des amers existants) une opération d'**initialisation**, c'est à dire, d'inclusion de nouveaux amers dans le vecteur d'état. Ceci permet au robot d'entreprendre des tâches de vraie exploration. Cette étape d'initialisation est triviale dans le cas où les capteurs extéroceptifs rendent des mesures inversibles (avec mesurabilité complète, par exemple par télémétrie laser), mais représente des gros casse-têtes pour le cas d'observations non-inversibles ou partielles: le niveau d'incertitude dans les degrés de liberté non mesurables est infini, mettant en échec les approximations linéaires utilisées dans les filtres SLAM. C'est le cas de la vision, capteur à nature projective qui entraîne la perte d'une dimension dans les observations: la profondeur des motifs perçus dans la scène. Il faut noter que cette perte est compensée par la capacité d'observation jusqu'à l'infini, avec des avantages importantes comme la stabilité des estimations d'orientation [Sola et al., 2008]. Je vais donner la solution au problème d'initialisation la plus utilisée en ce moment, pour des systèmes SLAM basés sur le filtre de Kalman étendu et similaires. Elle s'agit d'une **initialisation immédiate** [Solà et al., 2005], qui permet au système SLAM de prendre en compte des amers présentant des parallaxes très faibles, voire nuls. Dans la solution standard d'aujourd'hui, les points 3D sont codés par le rayon visuel initial (la partie mesurée, composée du centre optique au moment de l'initialisation et l'orientation du rayon visuel) et l'**inverse de la distance** à l'amer le long de ce rayon (la partie non mesurable). Cet "inverse de la distance" possède un rang d'incertitude fini et rend les équations d'observation quasi-linéaires dans ce rang [Civera et al., 2008].

On peut remarquer des nombreux points en commun et en divergence entre l'OV et le V-SLAM. Souvent, des différences de type méthodologique ont été les premières à être soulignées (optimisation versus filtrage), mais la réalisation récente de techniques V-SLAM par la force des optimisations non-linéaires (attribuées normalement à l'OV) a fait disparaître cette division. Je vais focaliser mon attention sur des aspects bien plus conceptuels, à savoir, l'aspect mémoire et l'aspect de recherche active, car ils sont ceux qui peuvent affecter d'une forme plus claire les performances limites des technologies visées.

L'aspect **mémoire** fait clair appel au fait d'avoir l'environnement cartographié. La cartographie permet au robot de re-visiter les endroits, et si la technique d'estimation utilisée est suffisamment bien conçue, son incertitude en position retournera toujours à celle que le robot avait au moment du premier passage par l'endroit. Ce phénomène, appelé "fermeture de boucle", est inaccessible aux méthodes OV (pour lesquels l'incertitude en position est en croissance monotone), et semble naturel pour le SLAM. La navigation en convoi demande aussi le partage de points de repère parmi les différents véhicules et profitera bien des aspects mémoire du SLAM. En plus, une cartographie multi-robot permet aussi des observations inter-véhicule (événements dits rendez-vous), qui peuvent avoir un effet très positif sur la consistance globale des estimations, notamment en proportionnant fermetures de boucle et observabilité du facteur d'échelle en applications monoculaires [Vidal-Calleja et al., 2009].

L'aspect de **recherche active** est vaguement lié au précédent. Il concerne la stratégie d'extraction d'information des images pour optimiser l'usage de l'information tout en minimisant l'effort de calcul. Il s'agit d'utiliser toute l'information dont on dispose pour cibler l'opération des algorithmes d'extraction et d'appariement d'amers dans les images (ces tâches sont souvent les plus fortes consommatrices de ressources de calcul). Cet aspect est aussi un avantage pour le SLAM. En particulier, l'exploitation de l'information SLAM permet de sélectionner un nombre réduit d'amers à mesurer (ceux qui vont donner le plus d'information), et de prédire dans quelle région de l'image on va les trouver, et avec quelle apparence. De manière similaire, nous serons aussi capables de rechercher des nouveaux amers à initialiser seulement dans les régions de l'image où l'on détecte un manque d'information.

Nous allons aussi discuter, si le temps le permet, sur les parties **L** (de Localisation) et **M** (de Mapping: cartographie) du SLAM. Nous avons dernièrement poussé le M du SLAM en avant en rajoutant des amers rectilignes dans la représentation de l'environnement. Je montrerai quelques résultats préliminaires sur ce sujet. Néanmoins, je serai obligé de souligner que, si bien le L et le M sont intégrés dans le SLAM, le L peut en fait se résoudre par moyen de très peu d'amers bien

sélectionnés (cartes très éparées), parmi que le M, s'il se veut utile à d'autres tâches (planification, commande, reconnaissance d'objets) demande la construction de cartes bien plus denses et complexes. Le SLAM proprement dit ne donnera pas une solution globalement satisfaisante à ces deux sous-problèmes du fait que les ressources de calcul seront toujours limitées, et la résolution de L avec une carte M très dense est absurde. Je crois que le SLAM devra rester comme une méthode pour le L avec des capacités de mémoire, et le M demandera d'efforts supplémentaires de densification, lesquels devront s'appuyer sur le SLAM complète afin de garantir la consistance des solutions L et M. On pourrait dire, SLAM+M.

[Civera et al., 2008] Civera, J., Davison, A., and Montiel, J. (2008). Inverse depth parametrization for monocular SLAM. *IEEE Trans. on Robotics*, 24(5).

[Comport et al., 2007] Comport, A. I., Malis, E., and Rives, P. (2007). Accurate quadrifocal tracking for robust 3D visual odometry. In *Int. Conf. on Robotics and Automation*, pages 40–45.

[Konolige et al., 2007] Konolige, K., Agrawal, M., and Solà, J. (2007). Large-scale visual odometry for rough terrain. In *Int. Symp. on Research in Robotics*, Hiroshima, Japan.

[Solà et al., 2005] Solà, J., Monin, A., Devy, M., and Lemaire, T. (2005). Undelayed initialization in bearing only SLAM. In *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 2499–2504, Edmonton, Canada.

[Solà et al., 2008] Solà, J., Monin, A., Devy, M., and Vidal-Calleja, T. (2008). Fusing monocular information in multi-camera SLAM. *IEEE Trans. on Robotics*, 24(5):958–968.

[Vidal-Calleja et al., 2009] Vidal-Calleja, T., Berger, C., Solà, J., and Lacroix, S. (2009). Multiple robot mapping with heterogeneous landmarks. *Journal of Field Robotics*, In preparation.

## **Patrick Rives (INRIA)**

**Titre:** Méthodes de SLAM visuel

### **Résumé :**

Du fait de la richesse d'information qu'elle procure et du coût négligeable des caméras, la vision est devenue dans les dix dernières années une modalité sensorielle incontournable pour l'autonomie des systèmes robotiques. Parallèlement aux développements technologiques permettant d'envisager maintenant des applications temps réel, les avancées méthodologiques autorisent aujourd'hui le développement de solutions efficaces et robustes au problème de localisation et de Slam. Dans cet exposé, nous formaliserons tout d'abord le problème générique du SLAM et présenterons quelques outils méthodologiques permettant de l'aborder. Puis, nous focaliserons notre propos sur les capteurs de vision en montrant les problèmes spécifiques liés à l'utilisation de tels capteurs. Nous présenterons ensuite quelques résultats récents obtenus en SLAM visuel pour différentes configurations de vision monoculaire ou stéréo.