

# Un système de guidage de personnes utilisant des images omnidirectionnelles

## An omnidirectional image based guidance system for persons

Bertrand VANDEPORTAELE, CATTOEN Michel, MARTHON Philippe  
IRIT & LEN7  
ENSEEIH7, 2 rue Camichel, 31071 TOULOUSE  
bvdp@enseeiht.fr

### Résumé

*Nous développons un système de localisation et de guidage de personnes dans un site modélisé en 3D utilisant des méthodes de vision par ordinateur et des images omnidirectionnelles.*

*Ce système fonctionne en deux phases. Tout d'abord, un expert doit constituer une base de données du site. Ensuite, un utilisateur souhaitant se localiser dans le site acquiert des images avec un dispositif portable constitué d'une caméra et d'un capteur d'orientation dans l'espace. Ces images sont comparées avec le contenu de la base de données et, si une forte corrélation est détectée, l'utilisateur peut être localisé dans le site. Un chemin à l'intérieur du site allant de la position de l'utilisateur jusqu'à la destination souhaitée doit alors être trouvé afin de permettre le guidage. La base de données est constituée d'un modèle 3D texturé du site acquis à l'aide d'un capteur 3D omnidirectionnel et d'images panoramiques.*

*Une représentation des lieux invariante vis à vis des rotations est calculée en utilisant un capteur omnidirectionnel virtuel. Deux étapes sont nécessaires pour apparier l'image de l'utilisateur avec un échantillon de la base de données. Premièrement, la base de données doit être projetée dynamiquement par le capteur virtuel sur des images panoramiques à l'aide d'une méthode optimisée. Deuxièmement, l'image de l'utilisateur doit être projetée par le même capteur en utilisant les informations fournies par le capteur d'orientation.*

*Dans cet article nous décrivons le système et plus particulièrement les deux parties suivantes : synthèse d'image panoramique et rectification d'image.*

### Mots Clefs

Vision omnidirectionnelle, synthèse et rectification d'images, localisation.

### Abstract

*We are developing a system to allow localization and guidance of people inside a 3D modelled site using computer vision techniques and omnidirectional images.*

*The system operates in two phases. Firstly, a database of the site has to be built by an expert. Then a user wishing*

*to locate himself in this site takes pictures with a portable device which includes a standard camera and an orientation sensor. These pictures are compared with the contents of the database and if a high correlation is detected, the user can be located inside the site. Finally, the desired path through the site from the actual user's location to the place where he wants to go is found, allowing the guidance.*

*The database of the site consists in a 3D textured model and a few panoramic images.*

*A rotation invariant representation of locations inside the site is computed using a virtual omnidirectional sensor. Two steps are necessary to match the user's images with the 3D database. Firstly, the 3D database has to be dynamically projected by this virtual sensor onto 2D omnidirectional images using an optimized method. Secondly, the user's images have to be projected by the same virtual sensor using information from the orientation sensor.*

*In this paper, we describe the system and focus on the following two tasks : omnidirectional image synthesis and image rectification.*

### Keywords

Omnidirectional vision, image synthesis and rectification, localization.

## 1 Introduction

Le GPS (Global Positioning System) est aujourd'hui largement utilisé pour localiser et guider des personnes à l'extérieur. L'appareil de localisation doit pouvoir recevoir les signaux provenant des satellites pour trianguler la position du récepteur de l'utilisateur. Pour les applications à l'intérieur des bâtiments, le LPS (Local Positioning System) est utilisé. Au lieu de trianguler la position à partir de satellites, ce système utilise des bornes émettrices installées à l'intérieur du site. Ce système nécessite l'installation de bornes et peut subir des problèmes d'interférences.

Nous travaillons actuellement sur un système permettant la localisation et le guidage de personnes à l'intérieur d'un site connu, utilisant des techniques d'imagerie. L'intérêt de cette approche est qu'il n'est pas nécessaire d'installer d'appareils particuliers à l'intérieur du site. En contrepartie, le capteur et les algorithmes sont plus complexes.

Le schéma de la figure 1 montre comment un tel système permet de se localiser. La figure présente un site avec trois pièces distinctes. Un certain nombre d'images ont été acquises à l'intérieur de ce site et constituent la base de données. L'espacement maximal entre les différentes images est de l'ordre de 30 centimètres. Pour localiser une image acquise par un utilisateur, il faut trouver dans cette base de données l'image la plus "ressemblante".

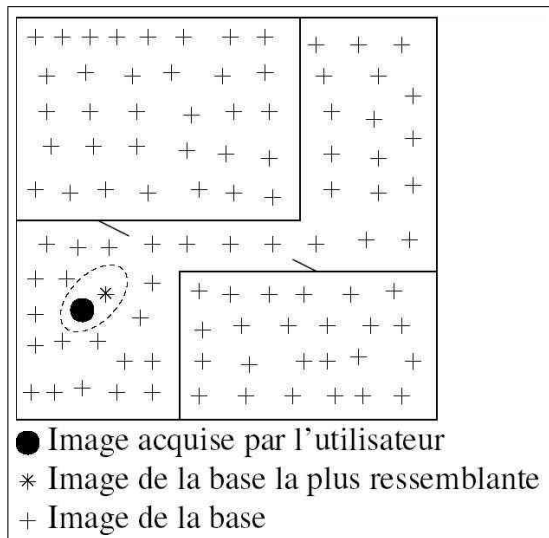


FIG. 1 – Localisation par l'image

Cette étude pourra déboucher sur des systèmes intégrés par exemple dans des téléphones portables équipés de caméra. Dans ce cas, les calculs pourront être déportés vers un serveur recevant les images par téléphone, le guidage pouvant quant à lui être réalisé en temps réel par la voix. Quelques améliorations devraient être réalisées au niveau des téléphones afin de les doter d'un capteur d'orientation angulaire que nous décrivons plus loin.

De nombreuses applications sont possibles, particulièrement le guidage à l'intérieur des lieux publics (musées, administrations...) et l'assistance aux déficients visuels.

Notre système utilise des méthodes déjà appliquées à la robotique [1], [2], [3], [4], où des images panoramiques provenant de capteurs panoramiques catadioptriques [5] sont appariées afin de déterminer l'endroit depuis lequel elles ont été acquises comme montré sur la figure 2.

Notre objectif est d'améliorer ce principe et de l'adapter au guidage de personnes. Premièrement, il doit être possible d'utiliser des capteurs différents pour l'acquisition du site et la localisation de l'utilisateur afin de permettre l'emploi de caméras standards. Ces caméras sont plus petites, moins chères et moins fragiles que les capteurs catadioptriques utilisés pour l'acquisition du site. Deuxièmement, le système doit pouvoir s'adapter facilement à des changements à l'intérieur du site grâce à une représentation appropriée du site. Enfin, contrairement aux robots, un utilisateur ne tient pas la caméra avec une orientation constante et d'une

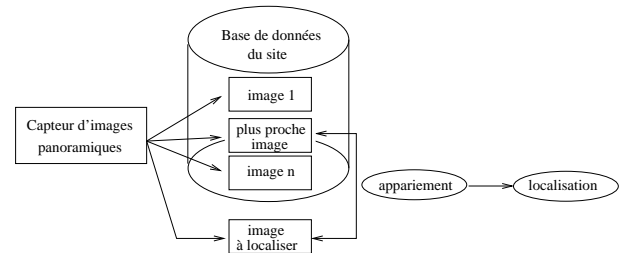


FIG. 2 – Schéma de principe de systèmes existants

personne à l'autre, la caméra n'est pas tenue à la même hauteur.

Dans cet article, nous présentons le principe du système et détaillons plus particulièrement deux étapes nécessaires à son fonctionnement : la synthèse d'images omnidirectionnelles et la rectification d'images.

## 2 Principe du système

Le problème de localisation d'une caméra tenue par un utilisateur dans un site est complexe. Sa résolution nécessite un certain nombre de contraintes.

La caméra de l'utilisateur réalise une projection de la scène 3D vers le plan image différente de celle du capteur utilisé pour l'acquisition du site, projections connues grâce à un calibrage géométrique. Cette caméra se trouve en une position  $(X, Y, Z)$  à l'intérieur du site et possède une orientation  $(\alpha, \beta, \theta)$ . La localisation consiste à estimer ces paramètres. L'utilisation d'un capteur d'orientation fixé à la caméra et étalonné fournit  $(\alpha, \beta, \theta)$ .

L'altitude de prise de vue  $Z$  peut être mesurée a priori. Afin d'éviter l'emploi d'un capteur de mesure spécifique (télé-mètre par exemple), cette altitude est supposée constante pour un utilisateur donné.

Le problème de localisation est ainsi ramené à une estimation des paramètres  $(X, Y)$ . Dans le cas d'un site à plusieurs étages, la localisation est effectuée en 2D dans un plan où les différents étages sont juxtaposés.

La figure 3 montre le schéma complet du système proposé. Ce système opère en deux étapes. Premièrement, une base de données du site est constituée par un expert. Ensuite, un utilisateur souhaitant se localiser utilise un appareil portatif pour prendre une image. Cette image est comparée au contenu de la base de données et si une image proche est trouvée, le système peut localiser l'utilisateur dans le site. L'appariement des images est réalisé sur des représentations 2D de la scène qui sont des images omnidirectionnelles constituées de deux images panoramiques montrant les parties supérieures et inférieures d'une sphère. Un grand nombre d'images est nécessaire afin de bien échantillonner spatialement le site en  $(X, Y)$ . Elles sont toutes orientées de la même manière et donc cette représentation est invariante vis à vis des rotations du capteur utilisé pour l'acquisition des images. Les données provenant de la base et du capteur de l'utilisateur sont formatées pour corres-

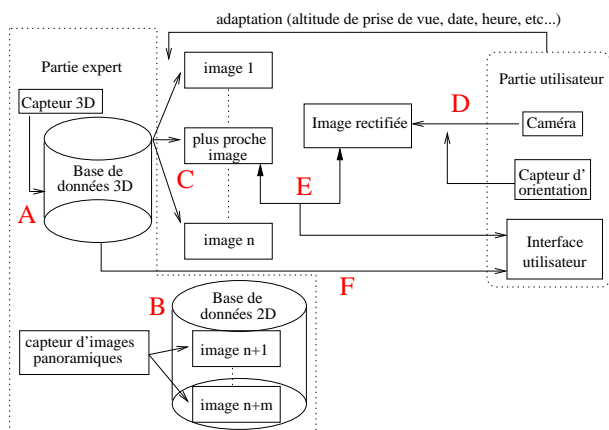


FIG. 3 – Principe du système. A : Acquisition 3D du site ; B : Acquisition et rectification d'images panoramiques ; C : Synthèse d'images omnidirectionnelles ; D : Rectification de l'image de l'utilisateur ; E : Appariement ; F : Guidage

pondre à cette représentation. Dans la suite, la création de ces représentations sera appelée synthèse lorsqu'elles seront obtenues à partir d'un modèle 3D et rectification lorsqu'elle seront obtenues à partir d'une image provenant d'un capteur.

Un modèle 3D texturé du site est utilisé comme majeure partie de la base de données. Un capteur 3D omnidirectionnel est actuellement en cours de développement afin de réaliser l'acquisition de ce modèle.

L'utilisation d'un modèle 3D pour représenter le site présente de nombreux avantages par rapport à une base de données d'images 2D statique parmi lesquels :

- Possibilité de générer les images en accord avec les conditions réelles de l'utilisateur qui souhaite se localiser, notamment à l'altitude de prise de vue de l'utilisateur  $Z$ .
- Compacité au niveau du stockage.
- Facilité de mise à jour par l'expert. En effet, il n'est pas nécessaire d'acquérir des centaines d'images du site chaque fois qu'il change. Une simple modification au niveau du modèle suffit.
- Possibilité d'ajouter de l'information une fois pour toutes au niveau du modèle 3D au lieu d'avoir à le faire sur chaque image.

Si une partie du site ne peut pas être modélisée en 3D, il est toujours possible d'acquérir des images panoramiques à l'aide d'un capteur catadioptrique [5], [6], [7]. Notre capteur permettant l'obtention de ces images est composé d'un miroir paraboloidal, d'une lentille orthographique et d'un appareil photo numérique. Il dispose d'un centre de projection unique situé au foyer du paraboloïde du miroir. Cette propriété est très intéressante puisqu'elle permet de reprojeter l'image du capteur sur une surface quelconque. Il est donc possible de générer des représentations 2D omnidirectionnelles partielles du lieu où se trouve le foyer.

Une base d'images omnidirectionnelles de synthèse est gé-

nérée dynamiquement à partir du modèle 3D et calculée de manière adaptative.

La hauteur de prise de vue, fixe pour un utilisateur donné, permet de générer des images de la base de données correspondant à la même altitude, étape nécessaire pour permettre l'appariement.

La date et l'heure peuvent être utilisées pour éventuellement générer les images de la base de données à partir d'un modèle 3D acquis dans des conditions proches (même saison ou même période de la journée). Ces paramètres permettent de prendre en compte des variations d'éclairage ou même de prévoir la présence ou non de certains objets dans le site.

L'image de l'utilisateur est rectifiée pour être appariée avec l'une de ces images omnidirectionnelles. Une représentation invariante vis à vis des rotations de la caméra est générée à partir de l'image et des informations provenant du capteur d'orientation monté sur la caméra.

Les images panoramiques de la base de données acquises directement par un capteur catadioptrique sont aussi rectifiées pour correspondre aux représentations 2D omnidirectionnelles et s'affranchir de l'orientation du capteur. Ainsi, toutes les images de la base de données sont du même type. Des méthodes d'appariement d'images sont utilisées afin de trouver l'image la plus proche dans la base de données d'images. Alors, la localisation peut être effectuée si l'image est considérée comme représentative d'un lieu unique à l'intérieur du site.

Une autre approche consistant à générer des images à partir de la base de données qui soient directement appariables avec celles du capteur de l'utilisateur a été étudiée puis écartée car elle ne permet ni le précalcul ni la réutilisabilité de la base de données d'images. Cette démarche peut néanmoins être appliquée dans un second temps, une fois l'utilisateur localisé, en synthétisant une seule image afin de confirmer la localisation.

Le guidage est réalisé en trouvant un chemin à l'intérieur du modèle 3D du site et en relocalisant régulièrement l'utilisateur afin de vérifier qu'il ne prend pas une mauvaise direction.

Le prototype du capteur omnidirectionnel 3D étant encore en cours de développement, aucun modèle 3D d'un site réel n'est actuellement disponible. C'est pourquoi les résultats présentés dans cet article sont obtenus à partir d'un modèle 3D synthétique même si la rectification d'image fonctionne avec des images provenant d'une caméra réelle.

### 3 Synthèse d'images omnidirectionnelles

Les images omnidirectionnelles 2D doivent être calculées à partir du modèle 3D du site afin d'être comparées avec les images de l'utilisateur. Une image omnidirectionnelle peut être décomposée en deux images panoramiques, chacune montrant l'hémisphère supérieur ou inférieur.

Comme le calcul d'une image panoramique en utilisant des méthodes standards de raycasting [8] est relativement coûteux

teux en temps, un algorithme rapide est nécessaire pour accomplir cette tâche. Nous proposons une méthode basée sur une décomposition de la projection de la scène vers le capteur virtuel en deux étapes.

La figure 4 montre le principe de la synthèse.

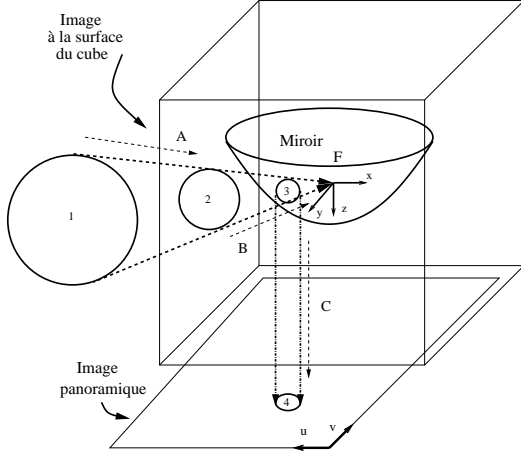


FIG. 4 – Principe de la synthèse d’image panoramique. A : Projection de la scène vers les faces d’un cube. B : Projection des faces du cube vers le miroir. C : Projection du miroir vers le plan image panoramique.

Tout d’abord, la scène (Objet 1) est projetée sur les faces d’un cube (Image 2) au travers de son centre  $F$ .

Ensuite, la projection des faces du cube vers le capteur catadioptrique virtuel avec un centre de projection  $F$  est calculée via raycasting. Le capteur virtuel utilise les mêmes paramètres intrinsèques que le capteur catadioptrique réel utilisé pour l’acquisition directe d’images dans la base de données. Les images à la surface du cube (Image 2) sont projetées à la surface du miroir paraboloidal (Image 3) et réfléchies vers le plan image panoramique (Image 4).

Grâce à cette décomposition, la partie non linéaire de la projection des faces du cube à l’image panoramique peut être précalculée et stockée dans une table de correspondance. Seule la partie linéaire reste à calculer pour chaque image panoramique.

Les détails du calcul sont décrits ici :

Soient  $\alpha_u$  et  $\alpha_v$  les distances focales horizontales et verticales exprimées en unité pixelique et  $h$  le paramètre du paraboloïde du miroir. Soient  $u_0$  et  $v_0$  les coordonnées de la projection du centre optique  $F$  sur le plan image.

Pour calculer la table de correspondances, chaque pixel de l’image panoramique est associé avec un pixel d’une image d’une des faces du cube.

La surface du miroir inférieur est définie par :

$$z_1 = \frac{h^2 - (x_1^2 + y_1^2)}{2h} \quad (1)$$

La surface du miroir supérieur est définie par :  $z_2 = -z_1$ .

La projection orthographique peut s’écrire :

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} -\alpha_u & 0 & 0 & u_0 \\ 0 & -\alpha_v & 0 & v_0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \\ 1 \end{pmatrix} \quad (2)$$

Donc le centre de chaque pixel  $(u, v)$  de l’image panoramique est projeté à la surface du miroir en un point  $P_1 = (x_1, y_1, z_1)^T$ . Pour le miroir inférieur :

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} = \begin{pmatrix} -\frac{u-u_0}{\alpha_u} \\ -\frac{v-v_0}{\alpha_v} \\ \frac{h^2 - x_1^2 - y_1^2}{2h} \end{pmatrix} \quad (3)$$

La première intersection du rayon de  $(0, 0, 0)$  vers  $P_1$  avec l’une des six faces du cube détermine quel pixel de quelle image correspond à  $(u, v)$  dans l’image panoramique.

Pour calculer les six projections de la scène sur le cube, six caméras perspectives virtuelles sont utilisées. Ces caméras ont un champ de vision de 90 degrés et sont toutes placées en  $F$ . Leurs orientations sont réglées afin de correspondre aux six faces du cube comme montré dans la figure 5.

Deux images panoramiques sont calculées à partir de ce jeu d’images par simple raycasting pour obtenir une représentation 2D omnidirectionnelle d’un lieu donné. Les figures 6 et 7 montrent les images obtenues pour des miroirs orientés soit vers le haut soit vers le bas. Nous avons décidé d’utiliser le pixel le plus proche dans les images à la surface du cube pour chaque rayon.

Étant donné que les images panoramiques sont calculées à partir d’un échantillonnage de la scène à la surface des faces d’un cube, le rayon utilisé pour calculer la couleur d’un pixel ( $R2$ ) n’est pas exactement dans la même direction que celui obtenu par réflexion ( $R1$ ) comme le montre la figure 8. Les expérimentations ont montré que l’erreur angulaire maximale entre  $R1$  et  $R2$  est de l’ordre de  $0.1^\circ$  pour une caméra panoramique avec  $\alpha_u = \alpha_v = 5$ ,  $h = 32$  et une taille d’image de  $512 \times 512$  pixels pour l’image panoramique et les images à la surface du cube.

Les six projections de la scène à la surface du cube sont calculées efficacement avec une API de synthèse d’image telle qu’OpenGL. Le rendu est réalisé par des cartes graphiques accélératrices 3D. Grâce à notre méthode, le temps de rendu peut être réduit par un facteur 30 par rapport à du simple raycasting pour une seule image. De plus, comme deux images panoramiques sont calculées en même temps pour obtenir une représentation omnidirectionnelle 2D, le gain effectif est encore quasiment multiplié par deux, les six projections à la surface du cube pouvant être utilisées deux fois.

## 4 Rectification d’images

Les images de l’utilisateur doivent être rectifiées pour être appariées avec les images panoramiques de synthèse générées à partir de la base de données 3D. Le capteur virtuel utilisé pour générer ces images panoramiques ayant

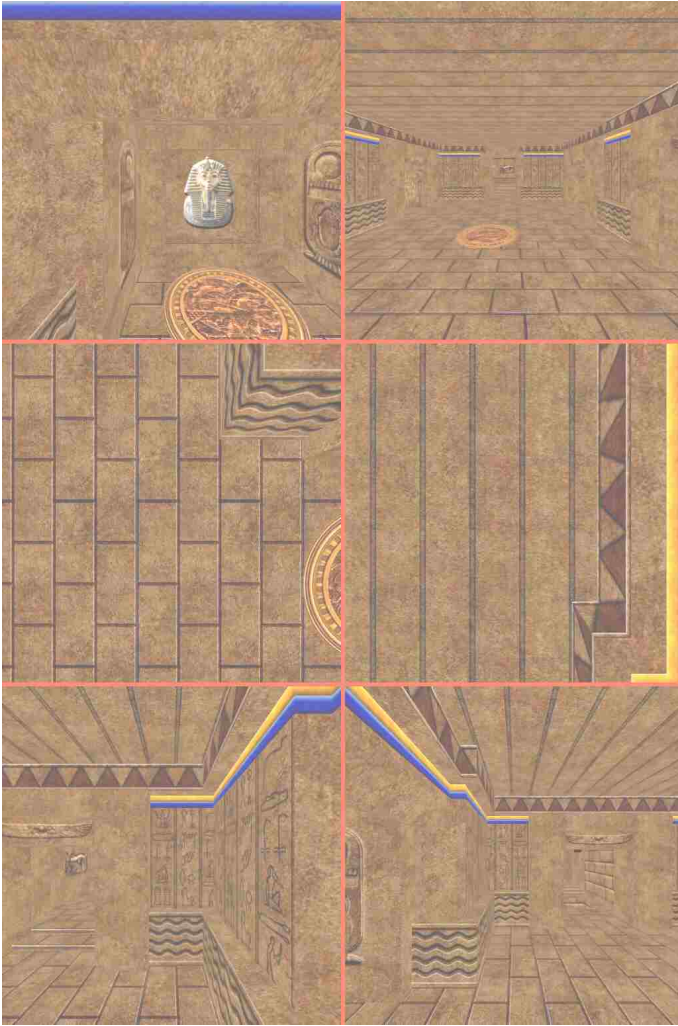


FIG. 5 – Projection de la scène sur les faces du cube

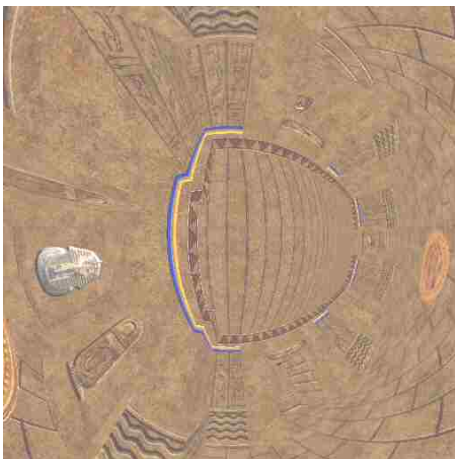


FIG. 6 – Image panoramique de synthèse supérieure

un centre de projection unique, il est possible de projeter l'image d'un autre capteur disposant aussi d'un centre de



FIG. 7 – Image panoramique de synthèse inférieure

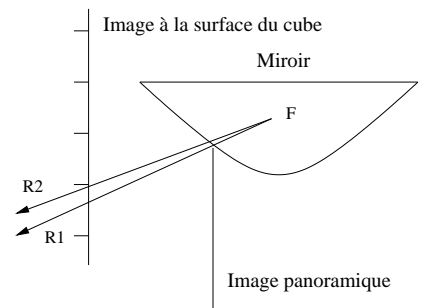


FIG. 8 – Erreur de synthèse d'image.

projection unique sur une image panoramique. Nous avons développé un capteur d'orientation pour connaître l'orientation de la caméra de l'utilisateur autour de trois axes. Ce capteur est composé d'une boussole numérique et d'un accéléromètre deux axes utilisé comme inclinomètre.

$\alpha_u, \alpha_v, h, u_0, v_0$  sont les paramètres du capteur catadioptrique virtuel avec lequel les images sont synthétisées comme défini dans le chapitre 3.

La caméra de l'utilisateur est modélisée en utilisant les paramètres suivants :

$\alpha_{P_u}$  et  $\alpha_{P_v}$  = distances focales exprimées en unité pixelique dans les directions  $u$  et  $v$ .

$u_{P_0}, v_{P_0}$  = position du point principal (projection du centre optique sur le plan image de la caméra).

$k_{P_1}, k_{P_2}$  = paramètres de distorsions radiales.

La valeur de chaque pixel de l'image panoramique est calculée comme montré dans la figure 9.

Le centre de chaque pixel  $(u, v)$  est projeté à la surface du miroir du capteur virtuel vers un point  $P_1 = (x_1, y_1, z_1)^T$  en utilisant l'équation (3).

Pour exprimer  $P_1$  dans le système de coordonnées de la caméra  $P$ ,  $P_1$  est multiplié par une matrice de rotation qui représente la rotation de  $C$  vers  $P$  en accord avec les angles  $\alpha, \beta$  et  $\theta$  fournis par le capteur d'orientation.

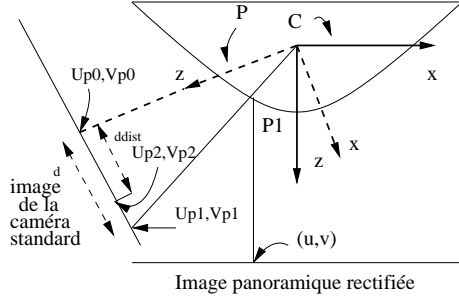


FIG. 9 – Principe de rectification d’une image d’une caméra standard

$$P_p = \begin{pmatrix} x_P \\ y_P \\ z_P \end{pmatrix} = R_{(\alpha, \beta, \theta)} \cdot \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (4)$$

Ce point est alors projeté à la surface du plan image de la caméra.

$$\begin{pmatrix} u_{P1} \\ v_{P1} \end{pmatrix} = \begin{pmatrix} \frac{\alpha_{Pu} \cdot x_P}{z_P} + u_{P0} \\ \frac{\alpha_{Pv} \cdot y_P}{z_P} + v_{P0} \end{pmatrix} \quad (5)$$

Pour appliquer la distorsion radiale à deux paramètres, la distance  $d$  par rapport au point principal est calculée :

$$d = \sqrt{\left( (u_{P1} - u_{P0})^2 + (v_{P1} - v_{P0})^2 \right)} \quad (6)$$

Une nouvelle distance est calculée en utilisant  $k_{P1}$  et  $k_{P2}$  :

$$d_{dist} = d \cdot (1 + k_{P1} \cdot d^2 + k_{P2} \cdot d^4) \quad (7)$$

La position rectifiée du point est donnée par cette nouvelle distance, en gardant la même direction :

$$\begin{pmatrix} u_{P2} \\ v_{P2} \end{pmatrix} = \begin{pmatrix} u_{P0} + \frac{d_{dist}}{d} \cdot (u_{P1} - u_{P0}) \\ v_{P0} + \frac{d_{dist}}{d} \cdot (v_{P1} - v_{P0}) \end{pmatrix} \quad (8)$$

La couleur du point est obtenue en utilisant la couleur du plus proche pixel dans l’image à rectifier. La figure 10 montre une image synthétique générée avec une caméra virtuelle en projection perspective. Dans ce cas, il n’y a pas de distorsion optique ( $k_{P1} = k_{P2} = 0$ ). La figure 11 montre une image rectifiée à partir des données de la figure 10.

Une étape d’étalonnage du capteur de l’utilisateur est nécessaire. Tout d’abord les paramètres de la caméra sont estimés. Ensuite, un étalonnage mécanique est effectué pour aligner les axes du capteur d’orientation avec les axes de la caméra.

Il est possible d’intégrer les données de plusieurs images de l’utilisateur sur la même image panoramique si le centre de projection ne bouge pas. Grâce à cela, une plus grande partie de la scène peut être vue comme le montre la figure 12. Ceci permet une localisation plus robuste, l’environnement de l’utilisateur étant plus contraint.

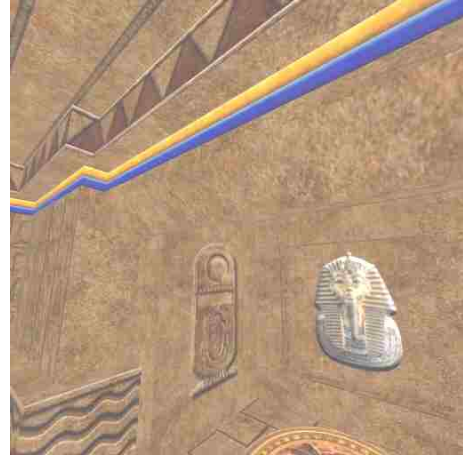


FIG. 10 – Image en provenance de la caméra standard virtuelle



FIG. 11 – Image rectifiée (partie panoramique supérieure)

Le même principe de rectification peut être appliqué à des images acquises avec un capteur panoramique afin de générer des représentations 2D omnidirectionnelles à partir des images panoramiques de l’expert. Le détail des calculs est présenté ci-après.

$\alpha_{Cu}, \alpha_{Cv}, h_C, u_{C0}, v_{C0}$  sont les paramètres du capteur catadioptrique réel utilisé pour l’acquisition d’images panoramiques comme défini dans le chapitre 3. Le modèle du capteur catadioptrique est enrichi pour prendre en compte les distorsions radiales (paramètres  $k_{C1}$  et  $k_{C2}$ ) dues à la caméra.

La valeur de chaque pixel de l’image panoramique rectifiée est calculée comme le montre la figure 13. Le centre du pixel  $(u, v)$  est projeté à la surface du miroir du capteur virtuel vers un point  $P_1 = (x_1, y_1, z_1)^T$  en utilisant l’équation (3).

Le point subit la rotation en utilisant l’équation (4).

La projection sur le capteur catadioptrique est un peu plus



sentation omnidirectionnelle comme montré au chapitre précédent. Il est aussi possible de réaliser plusieurs localisations (avec une ou plusieurs images pour chacune), en des lieux proches les uns des autres. Les données issues des différentes localisations sont ensuite corrélées pour trouver la position réelle de l'utilisateur comme montré dans la figure 16.

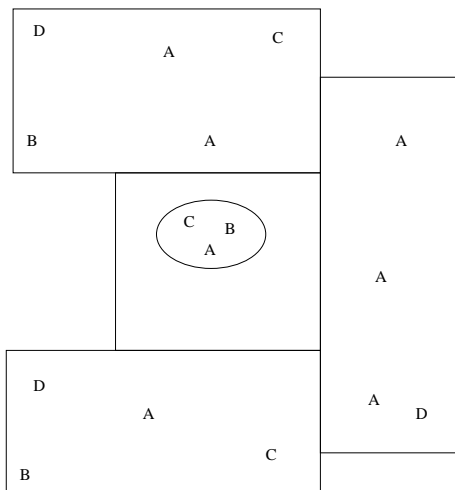


FIG. 16 – Détermination de la position par plusieurs appariements

Sur cette figure, quatre appariements sont effectués, dont les positions correspondantes sont notées *A*, *B*, *C* et *D*. La solution la plus probable est entourée par une ellipse. Cette ellipse contient *A*, *B*, *C* mais pas *D*, parce que l'appariement a échoué sur cette image.

Finalement, le guidage est effectué en utilisant un graphe de connexité des différentes parties du site. Une fois l'utilisateur localisé, il demande la direction pour aller en un autre lieu du site. Un chemin est trouvé et l'utilisateur est guidé au travers du site. L'utilisateur peut être localisé à intervalle de temps réduit, ce qui permet de s'assurer qu'il ne prend pas une mauvaise direction.

Pour permettre au système de fonctionner sur des données réelles, plusieurs problèmes doivent être résolus. Les changements d'éclairage peuvent rendre les images très différentes. Des déplacements d'objets ou des occlusions peuvent avoir lieu, rendant certaines zones de l'image non appariable. Un algorithme d'appariement robuste vis à vis de ces phénomènes doit donc être développé.

- Pour résoudre le problème des changements à l'intérieur du site, nous proposons d'ajouter dans la base de données 3D une information de probabilité de changement. Cette information sert ensuite de pondération lors de l'appariement, afin de rendre moins importantes les zones ayant une forte probabilité de changement. L'information peut être stockée dans les textures de la scène (en utilisant un canal alpha en plus des canaux RVB) et donc être projetée sur les représentations omnidirectionnelles 2D en même temps que le modèle 3D.

- Pour résoudre le problème d'occlusion de la scène, nous proposons de décomposer l'image en portions qui sont individuellement appariées. Cette décomposition permet à la majeure partie de l'image qui n'est pas occultée d'être appariée.
- Pour obtenir une représentation relativement invariante vis à vis des changements d'éclairage, nous proposons de travailler sur des informations de contours et de chrominance au lieu de travailler directement sur les images en RVB.

## 6 Conclusion

Un nouveau système de localisation et de guidage à l'intérieur d'un site et une représentation omnidirectionnelle 2D d'un lieu ont été présentées. L'utilité de cette représentation est de permettre l'appariement de données de la base et de l'utilisateur.

Les images de l'utilisateur et de l'expert sont rectifiées en utilisant un capteur portatif que nous avons développé composé d'un capteur d'image (panoramique ou standard) et d'un capteur d'orientation.

Les images omnidirectionnelles sont synthétisées à la volée depuis la base de données en utilisant une méthode de synthèse optimisée, permettant au système de disposer très rapidement des images à appairer.

Ce système est opérationnel sur des données synthétiques mais différentes parties du système doivent être améliorées pour qu'il puisse fonctionner sur des données réelles, principalement :

- Le capteur 3D omnidirectionnel va être validé et il sera possible d'acquérir un modèle 3D d'un site réel.
- La méthode de synthèse d'image optimisée va être implémentée différemment. L'inconvénient de l'implémentation actuelle est que les images doivent être transférées de la carte graphique à la mémoire centrale de l'ordinateur et la bande passante associée est relativement petite. Le temps nécessaire à ce transfert est par conséquent assez important. Plus de 60% du temps total est consacré à ce transfert. Le nouveau standard de carte vidéo, le PCI Express devrait résoudre ce problème, grâce à une bande passante plus grande. Le rendu devrait donc être encore accéléré. Le calcul de l'image panoramique à partir des six projections à la surface du cube peut aussi être décentralisé dans la carte vidéo, en utilisant par exemple l'extension OpenGL "fragment program". Ceci évite complètement le transfert et laisse au processeur central le temps d'exécuter d'autres tâches.
- Les méthodes d'appariement vont être étudiées et testées pour effectuer la localisation dans des sites réels.

## Références

- [1] T.Pajdla., V.Hlavac, *Image-based self-localization by means of zero phase representation in panoramic images*, In Proceedings of the 2nd International Conference on Advanced Pattern Recognition, volume 2013 of Lecture Notes in Com-

puter Science, pages 24- 33, Heidelberg, Germany, March 2001. IAPR, Springer-Verlag

- [2] J.J. Gonzalez-Barbosa et S. Lacroix, *Localisation d'un Robot Mobile par Indexation d'Images Panoramiques*, 16èmes Journées des Jeunes Chercheurs en Robotique. Lyon France 2002
- [3] A. Argyros, K. Bekris and S. Orphanoudakis, *Robot homing based on corner tracking in a sequence of panoramic images*, Computer Vision and Pattern Recognition, 2001. Pages : II-3- II-10 vol.2
- [4] L. Paletta, S. Frintrop, J. Hertzberg, *Robust localisation using context in omnidirectional imaging*, Proc IEEE International Conference on Robotics and Automation, ICRA 2001. Pages 2072-2077
- [5] R. Benosman and S. Bing Kang, *Panoramic Vision, Sensors Theory and Applications*, Ed Springer 2001
- [6] T. Svoboda, T. Pajdla, V. Hlavac, *Central panoramic cameras : Geometry and design*, Research report, Czech Technical University, Faculty of Electrical Engineering, 1997
- [7] C. Geyer and K. Daniilidis, *Catadioptric Projective Geometry*, International Journal of Computer Vision 45(3), 2001 pp. 223-243
- [8] A. S. Glassner, A. Arvo, R. L. Cook, E. Haines, P. Hanrahan, P. S. Heckbert, and P. Hanrahan. *An Introduction to Ray Tracing*, Harcourt Brace Jovanovich Published : 1989