

ÉCOLE POLYTECHNIQUE - PROMOTION X-98

## RAPPORT DE STAGE D'OPTION SCIENTIFIQUE

Yves BRONOËL

# Étude d'un système d'acquisition d'images stéréoscopiques à base de miroirs

Option : Informatique  
Directeur de l'option : M. Jean-Marc STEYAERT

Organisme d'accueil :  
Équipe MOVI - Laboratoire GRAVIR-IMAG-INRIA ZIRST  
655, avenue de l'Europe 38330 Montbonnot

Directeur de stage : M. Radu HORAUD

Stage effectué du 2 avril 2001 au 6 juillet 2001



## Résumé

L'utilisation de miroirs permet l'acquisition de différents points de vue à l'aide d'une seule caméra. Dans certains cas on obtient, par exemple, deux demi-images représentant deux photos d'une même scène et il est donc possible d'obtenir des informations 3D avec une seule prise de vue. Ce rapport présente les résultats obtenus dans ce domaine en étudiant les caractéristiques générales de tels systèmes et un capteur particulier utilisant quatre miroirs et développé au sein de l'INRIA.

On présente tout d'abord les avantages de tels outils par rapport aux méthodes plus classiques de vision 3D. Un certain nombre de systèmes présents dans la littérature sont ainsi décrits avec leurs avantages et leur domaine privilégié d'utilisation. Quelques résultats plus théoriques sont aussi présentés.

Une description physique et pratique est ensuite fournie pour le cas particulier du capteur à quatre miroirs qui a été utilisé. On y décrit notamment comment évolue le champ de vision commun en fonction des paramètres physiques de construction et quels ont été les premiers résultats obtenus en utilisant ce système comme ensemble virtuel de deux caméras.

On effectue ensuite une modélisation dans le formalisme de la géométrie projective pour aboutir à des résultats concernant la calibration. Il est montré comment, à partir de deux paires d'images, on arrive à extraire tous les paramètres physiques du capteur. On voit aussi pourquoi une seule paire d'images ne suffit pas, tout en permettant toutefois de bien affiner la calibration. Un logiciel a été développé en application de ces résultats et il fournit des données numériques permettant d'évaluer la qualité de cette calibration.

Enfin une étude est faite de la possibilité d'utilisation d'un tel système en mouvement afin de mettre en oeuvre une reconstruction 3D. Les thèmes abordés dans une telle démarche sont nombreux puisqu'on évoquera le suivi de points, la reconstruction 3D à partir d'une séquence d'images et la fusion de deux nuages de points 3D obtenus à l'aide de deux capteurs rigidement liés.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Motivation et méthodologie</b>	<b>3</b>
2.1	Les raisons d'un tel système . . . . .	3
2.2	Etat de l'art . . . . .	3
2.2.1	Systèmes à base de miroirs plans . . . . .	4
2.2.2	Systèmes à base de miroirs courbes . . . . .	4
2.2.3	Quelques résultats plus généraux . . . . .	5
2.3	Méthodologie employée . . . . .	6
2.4	Conventions . . . . .	6
2.4.1	Le modèle sténopé . . . . .	6
2.4.2	L'erreur de reprojection . . . . .	7
2.4.3	La contrainte épipolaire . . . . .	7
2.4.4	Les différents types de calibration . . . . .	7
<b>3</b>	<b>Description physique du système</b>	<b>8</b>
3.1	Présentation du système . . . . .	8
3.2	Champ de vision obtenu . . . . .	9
3.3	Premières applications . . . . .	10
3.3.1	Prises de vue . . . . .	10
3.3.2	Calibration simple . . . . .	11
3.3.3	Reconstruction 3D . . . . .	11
<b>4</b>	<b>Modélisation projective du système</b>	<b>15</b>
4.1	Modélisation d'un miroir en géométrie projective . . . . .	15
4.2	Calcul des matrices de projection . . . . .	15
4.3	Reconstruction 3D dans le repère caméra . . . . .	17
4.4	Calibration du système . . . . .	17
4.4.1	Calibration avec une paire d'images . . . . .	17
4.4.2	Calibration avec deux paires d'images . . . . .	18
4.4.3	Implémentation pratique . . . . .	19
4.4.4	Résultats . . . . .	19
<b>5</b>	<b>Utilisation du capteur en mouvement</b>	<b>21</b>
5.1	Problématique . . . . .	21
5.2	Méthode développée . . . . .	21
5.2.1	Détection de points . . . . .	21
5.2.2	Suivi de points . . . . .	21
5.2.3	Extraction du mouvement d'une caméra . . . . .	22
5.2.4	Alternative : utilisation de l'approximation par une homographie . . . . .	23
5.2.5	Mise en correspondance stéréo . . . . .	23
5.2.6	Reconstruction finale . . . . .	25
5.3	Perspectives . . . . .	27
<b>6</b>	<b>Conclusion</b>	<b>28</b>
<b>7</b>	<b>Remerciements</b>	<b>29</b>



# 1 Introduction

La vision 3D dont nous sommes naturellement dotés par nos deux yeux, est un défi difficile à relever dans le cadre de la vision par ordinateur. De nombreuses techniques sont possibles mais aucune ne semble pouvoir s'imposer devant les autres. D'une part les systèmes à plusieurs caméras sont relativement difficiles à mettre en oeuvre dans le sens où il n'est pas toujours évident de gérer leurs flux vidéo de façon synchronisée et de compenser les différences inévitables de perception entre les deux capteurs. D'autre part les systèmes utilisant une seule caméra en mouvement présentent de grandes difficultés pour extraire complètement le mouvement effectué par la caméra. Or celui-ci est nécessaire pour faire une reconstruction 3D euclidienne. Il n'y a donc pas dans ce domaine de solution idéale. D'un côté se trouvent les résultats, de l'autre la facilité de manipulation du capteur.

L'idée d'utiliser des miroirs semble apporter les avantages des deux camps et promet ainsi un capteur presque idéal. On a, en effet, de nombreux avantages mais qui ne vont pas bien évidemment sans certaines contraintes. Le sujet sur lequel j'ai travaillé pendant ces trois mois est un système de vision stéréoscopique utilisant quatre miroirs. Ce système a été développé en 1993 au sein de l'INRIA. Le but de mon travail a été l'étude de l'intérêt d'un tel capteur dans le domaine de la vision 3D. Cela passe par une approche tout d'abord assez physique et expérimentale. Il faut ensuite tenter de modéliser ce système dans le formalisme de la géométrie projective, cadre traditionnel de la vision par ordinateur. Enfin on doit envisager des applications adaptées et efficaces pour utiliser ce capteur.

Ce stage s'est effectué au sein de l'équipe MOVI de l'unité de recherche Inria Rhône-Alpes. Le thème principal de recherche y est la modélisation d'objets géométriques avec pour objectifs : leur reconnaissance, leur localisation et leur interprétation. MOVI est une des équipes de recherche du laboratoire GRAVIR (GRAphisme, VIsion, Robotique). C'est aussi un projet commun à l'IMAG et à l'INRIA Rhône-Alpes.

## 2 Motivation et méthodologie

### 2.1 Les raisons d'un tel système

Les systèmes à base de miroirs, appelés aussi catadioptriques, présentent une approche intéressante de la vision 3D. Là où on a besoin de deux prises de vue, ils se contentent d'une seule image. Ils apportent donc un certain nombre d'avantages et d'inconvénients par rapport aux systèmes traditionnels et on va commencer par identifier ces caractéristiques afin de mieux comprendre pourquoi de tels systèmes peuvent être utilisés et quelles sont leurs contraintes.

Une prise de vue simultanée de plusieurs caméras permet d'effectuer la vision 3D et le suivi d'un mouvement. On peut en effet avec un tel système effectuer une reconstruction 3D de la scène à chaque instant et suivre l'évolution dans le temps de cette reconstruction. Cette tâche est cependant relativement difficile dans la mesure où on ne dispose que de très peu de matériels permettant la synchronisation de plusieurs caméras et pour effectuer des acquisitions simultanées. Une astuce couramment employée consiste à utiliser des caméras monochromes et à les brancher sur les différents canaux d'un système d'acquisition couleur. Un système de miroirs résoud ce problème puisqu'on a, en recevant une image, les deux images parfaitement synchronisées. On pourra ainsi faire du suivi de mouvement sans risquer un décalage dans la synchronisation temporelle des caméras.

Dans le prolongement de l'argument précédent on peut aussi noter que la présence des deux images sur le même support permet d'effectuer dès l'acquisition des traitements sur les images par paires. On n'a pas besoin d'attendre la reconstitution des paires d'images pour effectuer par exemple une correction de luminosité ou toute sorte de traitement simple s'effectuant sur les deux images de manière identique. On pourrait même envisager dans une application pratique du système d'effectuer un certain nombre de traitements en "hardware" avant même l'acquisition des données par un ordinateur.

Intéressons-nous maintenant plus particulièrement aux images obtenues. Elles bénéficient d'un point commun très appréciable : elles proviennent d'une même caméra. La luminosité, le zoom sont les mêmes et on n'a pas de correction particulière à apporter aux images avant de les comparer. On dit qu'on aura d'une image à l'autre les mêmes paramètres intrinsèques<sup>1</sup> pour la caméra. Cela aura un impact très important lorsqu'on cherchera à améliorer la calibration du système.

Cependant le système a aussi des inconvénients à ne pas négliger. Le premier qui vient à l'esprit est que le fait de stocker dans une seule image les deux vues entraîne nécessairement que chacune des vues contient deux fois moins d'information que si on avait employé deux caméras. De plus certaines zones de l'image ne sont pas utilisables : la limite entre les deux images est un peu floue et certaines zones ne sont pas vues en commun par les deux objectifs. En résumé on a une perte de résolution importante pour les images.

Ces systèmes ne sont donc bien évidemment pas idéaux et présentent des caractéristiques favorables et d'autres moins. La suite du rapport permettra de voir comment, notre travail s'est orienté pour tirer partie au mieux des avantages de ces systèmes.

### 2.2 Etat de l'art

Même si le système sur lequel j'ai travaillé n'a pas été beaucoup étudié jusqu'ici, la littérature fournit des exemples de systèmes de vision 3D à base de miroirs ainsi que quelques résultats généraux dans ce domaine. Je vais évoquer tout d'abord les autres systèmes qui ont été étudiés : en premier lieu ceux avec des miroirs plans puis ceux avec des miroirs courbes. Enfin je donnerai quelques résultats généraux qui ont été apportés dans ce domaine.

---

<sup>1</sup>On précisera par la suite dans les conventions ce que sont exactement dans notre modèle les paramètres intrinsèques d'une caméra.



### 2.2.1 Systèmes à base de miroirs plans

D'autres systèmes de vision 3D ont été établis à partir de miroirs plans. Le nombre et la disposition des miroirs sont très variables. On trouvera dans la figure 1 trois exemples de ces systèmes. Un grand nombre de miroirs peut être utilisé mais cela rend la construction du système assez fastidieuse. Les principales propriétés peuvent être de plus mises en avant en étudiant les systèmes ayant de un à quatre miroirs.

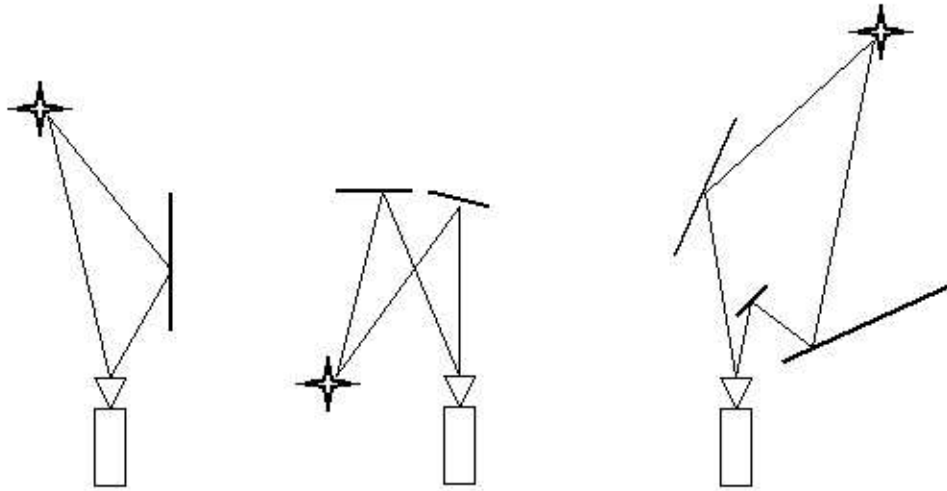


FIG. 1 – Trois exemples de systèmes à base de miroirs plans

On peut remarquer pour les systèmes de de la figure 1 que les images sont inversés puisqu'elles sont obtenues après un nombre impair de réflexions. Cet inconvénient mineur n'est pas présent dans le système à quatre miroirs étudié par la suite. On pourra consulter notamment [2] pour avoir une description plus précise de l'utilisation et de la calibration du système à deux miroirs et [3] fournit des éléments sur l'utilisation et la construction de systèmes à un ou trois miroirs pour l'obtention d'images rectifiées.

### 2.2.2 Systèmes à base de miroirs courbes

L'utilisation première des miroirs courbes n'avait pas pour but la stéréoscopie. Le but était de créer des caméras procurant une vision panoramique. On pouvait ainsi augmenter le champ de vision, ce qui peut s'avérer utile dans des applications comme la création de photos 360° d'un paysage.

Cependant certains se sont lancés dans la création de systèmes de vision 3D à base de miroirs courbes. Deux sont présentés dans la figure 2.

Le système de gauche a un fonctionnement relativement similaire au système à deux miroirs plans. Il utilise en effet deux miroirs courbes placés côte à côte et que regarde la caméra.

Le système de droite est plus original dans sa conception. On prend un grand miroir courbe au centre duquel on pose un petit miroir courbe ayant une courbure plus importante. La caméra est placée en face. La courbure plus importante du deuxième miroir permet la convergence vers un même point de l'espace de deux rayons distincts issus de la caméra et frappant les deux miroirs. Ce système est d'autant plus particulier qu'on a en définitive une image formée par le deuxième miroir au centre de l'image formée par le premier miroir.

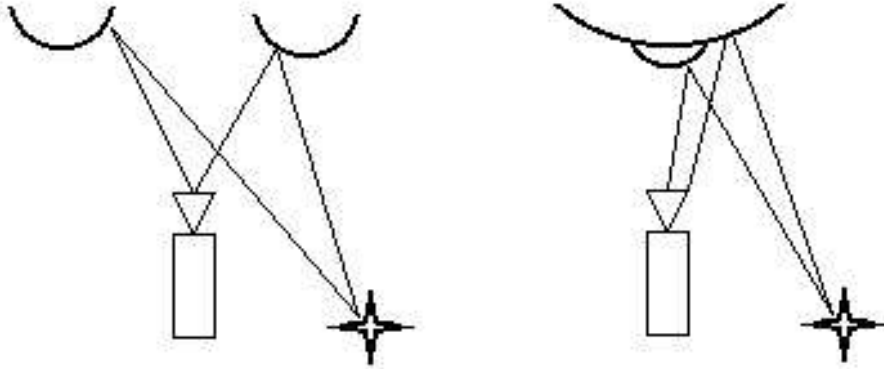


FIG. 2 – Deux exemples de systèmes à base de miroirs courbes

Ces systèmes à base de miroirs courbes ont été étudiés dans la littérature ainsi que la géométrie épipolaire particulière qui leur est associée. On pourra consulter [5] pour plus de détails sur l'utilisation du deuxième système. Toutefois leur utilisation reste assez marginale.

### 2.2.3 Quelques résultats plus généraux

Au delà des divers types de systèmes et de leur description, quelques résultats plus généraux sur les systèmes catadioptriques ont aussi été apportés. Deux d'entre eux m'ont plus particulièrement intéressé.

Le premier concerne la création d'un capteur stéréo permettant d'acquérir directement des images rectifiées. Une paire d'images rectifiées est une paire d'images obtenue en ayant des plans de projection parallèles à la droite reliant les deux centres de projection. Leur intérêt est de permettre une recherche plus rapide pour la mise en correspondance de point entre les deux images. En effet dans le cas où la droite reliant les deux centres de projection est horizontale, un point de la première image se trouve sur une ligne horizontale dans la deuxième image alors que pour une paire d'images non rectifiées, cette ligne peut être quelconque ce qui entraîne une recherche plus fastidieuse.

Gluckman et Nayar étudient dans [3] les conditions requises pour qu'un système génère de telles images. Ils montrent notamment comment en créer à l'aide d'un ou trois miroirs. Il y est aussi expliqué qu'un système comme le notre peut permettre d'obtenir des images rectifiées mais dans ce cas le champ de vision commun est nul ce qui présente un intérêt assez limité.

Le deuxième résultat concerne la géométrie épipolaire de certains systèmes à base de miroirs. Il est démontré dans [2] qu'on a parfois une contrainte supplémentaire sur la matrice fondamentale  $F$ . Cette contrainte s'exprime sous la forme suivante :

$$\det(F + F^T) = 0$$

Gluckman et Nayar utilisent un résultat de Maybank [4]. Celui-ci montre que dans la cadre d'un mouvement planaire entre les deux caméras, la matrice essentielle  $E$  vérifie :

$$\det(E + E^T) = 0$$

Appliquant ce résultat à leur système avec deux miroirs et utilisant le fait que les deux caméras virtuelles ont les mêmes paramètres intrinsèques, ils en déduisent la contrainte évoquée précédemment. Ce résultat est aussi applicable à notre système puisqu'on verra plus tard qu'on a aussi un mouvement planaire<sup>2</sup> entre les deux caméras virtuelles.

L'application pratique de ce résultat est un renforcement du calcul de la géométrie épipolaire. On réduit en effet avec cette contrainte le nombre de degrés de liberté à 6 au lieu de 7 dans le cas général.

## 2.3 Méthodologie employée

Le travail autour de ce système de vision 3D peut se décomposer en trois approches successives.

Tout d'abord, la prise en main et la familiarisation avec le système permettent de commencer à appréhender ses spécificités. Cela consiste à étudier la conception du capteur et à faire des acquisitions et des manipulations sur les paramètres pour obtenir des champs de vue autorisant la 3D. Cette première étape a conduit naturellement à l'utilisation du capteur comme un système virtuel de deux caméras, faisant ainsi abstraction de ses spécificités. On a dès lors pu effectuer la calibration du système et des reconstructions 3D.

Ensuite on effectue une modélisation projective du système. Le but d'un tel travail est de pouvoir améliorer toutes les étapes du travail de vision qui suivront. Cela concerne la calibration, la reconstruction 3D et de nombreuses autres applications. On mettra ainsi en avant les paramètres physiques du système et la possibilité, voire l'impossibilité, de les extraire lors de la calibration.

Enfin on envisage une possibilité d'application d'un tel capteur dans le domaine de la vision 3D en mouvement. On étudie en effet une méthodologie de reconstruction 3D à partir d'une séquence d'images. Notre but est qu'elle ne nécessite pas de mise en correspondance entre les deux images. L'idée mise en oeuvre est qu'on peut effectuer avec un système stéréo rigide une reconstruction 3D dans chacune des deux séquences d'images. La question à laquelle on tente alors de répondre est la suivante : comment fusionner ces reconstructions et les améliorer en utilisant les deux séquences parfaitement synchronisées de notre système.

## 2.4 Conventions

On donne ici quelques conventions qui seront utilisées dans ce rapport ainsi que quelques notations.

On se placera pour la modélisation dans le cadre de la géométrie projective et on notera par  $\simeq$  l'égalité à un facteur non nul près.

### 2.4.1 Le modèle sténopé

Les caméras employées sont supposées suivre le modèle sténopé : chaque caméra a un centre de projection et l'image d'un point de l'espace est l'intersection avec le plan image de la droite passant par ce point et par le centre de projection.

Une caméra se modélise donc par une matrice de taille 3x4, appelée matrice de projection. Celle-ci se décompose généralement de la manière suivante :

$$M = \begin{pmatrix} K & 0 \end{pmatrix} \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

$K$  correspond aux paramètres intrinsèques de la caméra et se met sous la forme<sup>3</sup> :

<sup>2</sup>Un mouvement planaire est un déplacement se composant d'une rotation et d'une translation dans le plan normal à l'axe de rotation.

<sup>3</sup>Parfois on ajoute un cinquième paramètre pour prendre en compte le fait que les axes de la caméra ne sont pas toujours exactement orthogonaux, mais la qualité des caméras employées rend ce paramètre généralement superflu.

$$K = \begin{pmatrix} a_u & 0 & u_0 \\ 0 & a_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Le point de coordonnées  $(u_0, v_0)$  est appelé centre optique de la caméra.

$R$  et  $t$  sont appelés paramètres extrinsèques de la caméra et correspondent à la rotation et à la translation qui font passer du repère de la mire de calibration au repère de la caméra.

Lorsqu'on connaît la matrice de projection  $M$ , il est possible de la factoriser pour obtenir les paramètres intrinsèques et extrinsèques. Cela nous permet alors de travailler aussi bien dans le repère de la mire de calibration que dans le repère de la caméra.

### 2.4.2 L'erreur de reprojection

On utilisera au cours de ce rapport l'erreur de reprojection pour évaluer la qualité d'une calibration ou d'une reconstruction. Ce critère correspond à la distance entre un point 2D et la projection du point 3D correspondant dans l'image. Cela nous donne la formule suivante<sup>4</sup> :

$$erreur = \left(u - \frac{(PM)_1}{(PM)_3}\right)^2 + \left(v - \frac{(PM)_2}{(PM)_3}\right)^2$$

où  $(u, v)$  est le point 2D,  $M$  le point 3D et  $P$  la matrice de projection.

Lorsqu'on évalue une matrice de projection, on somme cette erreur pour tous les points considérés. Afin de présenter des résultats numériques plus lisibles, on divise ensuite par le nombre de points et on prend la racine carrée du résultat. Cela nous donne ainsi une distance en pixel.

### 2.4.3 La contrainte épipolaire

Lorsqu'on manipule un système de deux caméras, on a une contrainte entre les projections  $p$  et  $p'$  d'un même point dans les deux images. Connaissant  $p$ ,  $p'$  ne peut se trouver que sur une certaine droite de l'autre image. On l'appelle droite épipolaire associée à  $p$ . Cette contrainte géométrique se traduit algébriquement par l'existence d'une matrice  $3 \times 3$  notée  $F$  vérifiant pour tout couple  $(p, p')$  la relation suivante :

$$p'^T F p = 0$$

$F$  est appelée matrice fondamentale.

### 2.4.4 Les différents types de calibration

Enfin on distingue deux types de calibration pour un même système stéréo. La calibration forte consiste à retrouver les matrices de projection des deux caméras. La calibration faible se contente par contre de l'estimation de la matrice fondamentale. La première permet d'effectuer des reconstructions dans un espace euclidien. On peut choisir comme repère celui d'une des caméras ou bien celui de la mire de calibration. La deuxième calibration ne nous permet qu'une reconstruction projective. Elle a par contre l'avantage de ne pas nécessiter de mire de calibration. On peut se contenter en effet d'une série de points mis en correspondance dans les deux images et estimer  $F$  à l'aide de la relation algébrique vue précédemment.

---

<sup>4</sup>On note  $V_i$  la  $i$ ème composante du vecteur  $V$

### 3 Description physique du système

#### 3.1 Présentation du système

Le système à quatre miroirs a été imaginé et réalisé par Frédéric Devernay et Hervé Mathieu. Les informations nécessaires à sa réalisation pratique peuvent être trouvées dans [1].

On commence par placer en face de la caméra deux miroirs faisant un angle de  $45^\circ$  avec l'axe optique. Ces deux miroirs permettent de diviser en deux parties le champ de vision de la caméra. Ensuite on place, de part et d'autre de ces miroirs, deux autres miroirs qui redirigent vers la cible la vue de la caméra. Ces deux derniers miroirs sont mobiles et permettent, comme on le verra dans la partie suivante, de faire évoluer le champ de vision commun. L'ensemble du système repose sur un support métallique sur lequel on a la possibilité de fixer la caméra. La figure 3 nous donne une photo de l'ensemble du système en condition de fonctionnement.

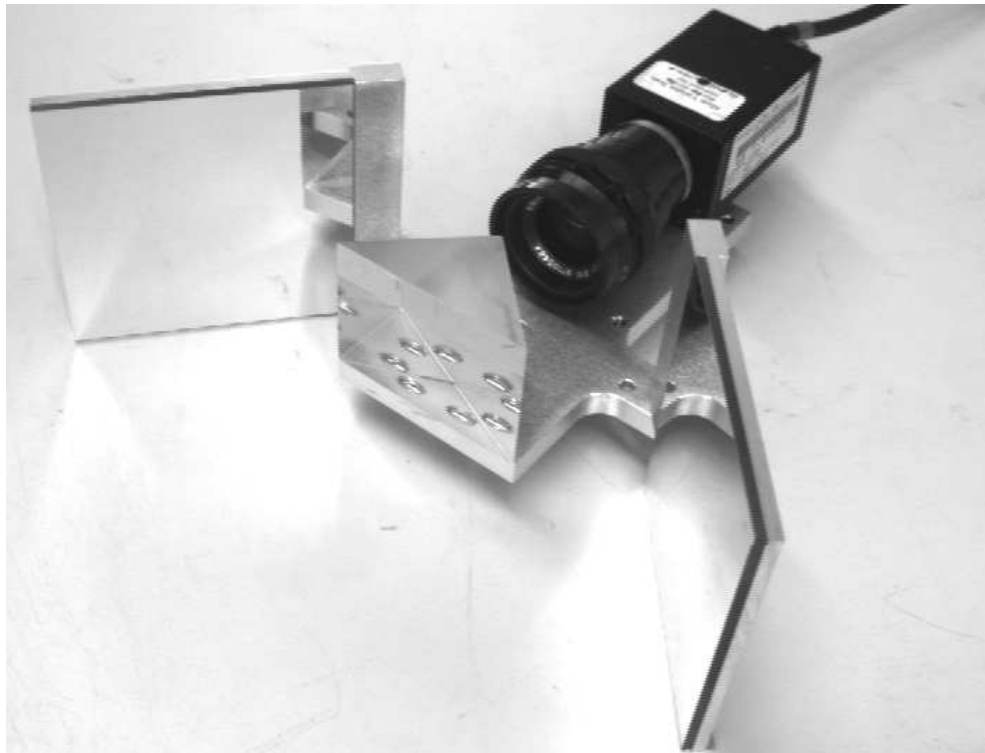


FIG. 3 – Photo du système en condition de fonctionnement

Une amélioration intéressante du système consiste à remplacer les deux premiers miroirs par un prisme avec deux faces réfléchissantes. Cela permet en effet de ne pas avoir à tailler deux miroirs en ayant la contrainte d'un positionnement à  $90^\circ$ .

On désigne par système de gauche (respectivement de droite), les deux miroirs situés à gauche (respectivement à droite) lorsqu'on se place au niveau de la caméra. On conservera cette convention pour la suite de ce rapport.

On peut de suite noter un des intérêts de ce système pour la vision 3D : il est compact et permet de conserver l'intuition du positionnement de l'objet observé ce qui n'est pas le cas avec la plupart des autres systèmes à base de miroirs. On conserve par exemple l'orientation de la scène puisque on a deux miroirs à gauche comme à droite. C'est donc un capteur dont on peut envisager plus facilement une utilisation pratique.

### 3.2 Champ de vision obtenu

Dans notre système on a donc deux miroirs mobiles. Cela nous donne la possibilité de faire varier le champ de vision commun en jouant sur la position de ces miroirs. En n'effectuant aucun calcul, on peut d'avance dire qu'on aura trois types de configuration pour le champ de vision commun : inexistant, présent à partir d'un certain point et de profondeur infinie, présent à partir d'un certain point et de profondeur finie. Ces trois configurations sont résumées dans la figure 4.

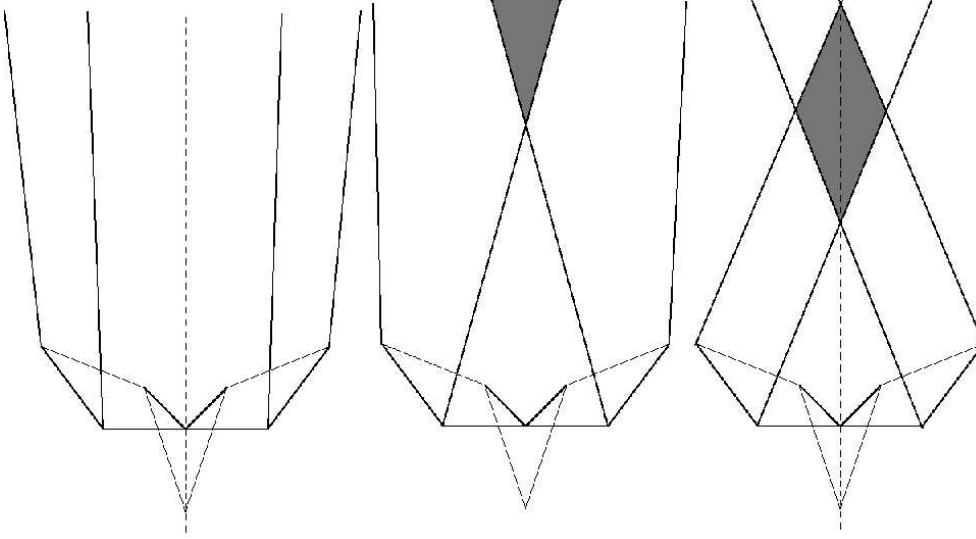


FIG. 4 – Les trois types de champ de vision commun

On va maintenant donner les principales équations permettant de définir les champ de vision des deux parties du système. On calcule pour ceci les équations des droites qui délimitent le champ de vision d'une partie du système et on effectue leur intersection avec les droites de l'autre partie du système. La configuration étant symétrique on ne présente les résultats que pour le système de gauche. On obtient ceux du système de droite en prenant l'opposé de l'ordonnée et en remplaçant  $\beta$  par  $\beta'$ . La figure 5 présente les notations et le repère employé pour ce calcul. On a confondu le centre du repère et le centre optique du système.

On pose :

$$l_2 = \frac{(a+b)\sin\omega}{\sin(\beta-\omega)}$$

Equation de  $\Delta_1$  :

$$y = b - (x - a) \tan\left(2\beta - \frac{\pi}{2}\right)$$

Equation de  $\Delta_2$  :

$$y = b + l_2 \cos\beta + (x - a - l_2 \sin\beta) \tan\left(\frac{\pi}{2} - 2\beta + \omega\right)$$

Condition pour avoir un champ de vision commun : on doit avoir intersection de  $\Delta_1$  et  $\Delta'_1$  c'est à dire  $\beta + \beta' \geq \frac{\pi}{2}$ . Dans ce cas, on a un point d'intersection de coordonnées :

$$x = a + \frac{2}{\tan\left(2\beta - \frac{\pi}{2}\right) + \tan\left(2\beta' - \frac{\pi}{2}\right)}$$

$$y = b \frac{\tan\left(2\beta' - \frac{\pi}{2}\right) - \tan\left(2\beta - \frac{\pi}{2}\right)}{\tan\left(2\beta - \frac{\pi}{2}\right) + \tan\left(2\beta' - \frac{\pi}{2}\right)}$$

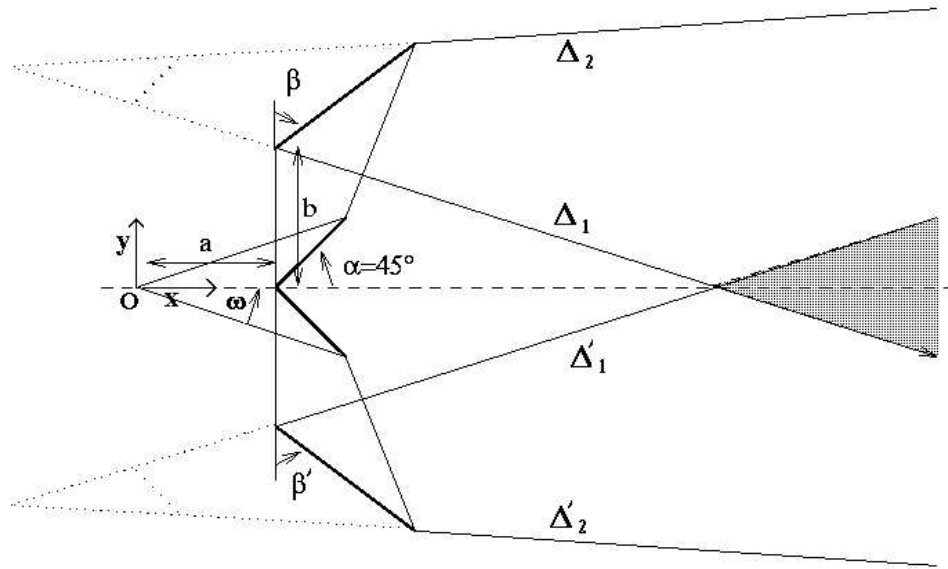


FIG. 5 – Notations pour la calcul des champs de vision

Condition pour avoir un champ de vision commun de taille finie : on doit avoir intersection de  $\Delta_2$  et  $\Delta'_2$  c'est à dire  $\frac{\pi}{2} + \omega \leq \beta + \beta'$ . Dans ce cas, on a un point d'intersection de coordonnées :

$$x = a + l_2 \sin \beta - 2 \frac{b + l_2 \cos \beta}{\tan(\frac{\pi}{2} - 2\beta + \omega) + \tan(\frac{\pi}{2} - 2\beta' + \omega)}$$

$$y = \frac{\tan(\frac{\pi}{2} - 2\beta' + \omega) - \tan(\frac{\pi}{2} - 2\beta + \omega)}{\tan(\frac{\pi}{2} - 2\beta + \omega) + \tan(\frac{\pi}{2} - 2\beta' + \omega)} (b + l_2 \cos \beta)$$

Enfin on peut calculer les positions des centres de projection virtuels des deux systèmes en faisant l'intersection de  $\Delta_1$  et  $\Delta_2$  ainsi que celle de  $\Delta'_1$  et  $\Delta'_2$ .

$$\text{Système de gauche} \begin{cases} x = a - (a + b) \sin 2\beta \\ y = b - (a + b) \cos 2\beta \end{cases}$$

$$\text{Système de droite} \begin{cases} x = a - (a + b) \sin 2\beta' \\ y = -b + (a + b) \cos 2\beta' \end{cases}$$

### 3.3 Premières applications

#### 3.3.1 Prises de vue

Après avoir un peu manipulé le capteur et l'avoir réglé pour obtenir un champ de vision commun, on peut obtenir les premières acquisitions. On s'aperçoit à ce moment-là que l'arête formée par les deux miroirs génère un flou au centre de l'image. Cela entraîne une perte supplémentaire de résolution au niveau de l'image et il faut essayer de limiter ce problème en utilisant le prisme et aussi en augmentant la distance entre cette arête et le centre de projection de la caméra. La figure 6 donne un exemple d'acquisition effectuée avec le système.



FIG. 6 – Exemple d'acquisition

### 3.3.2 Calibration simple

Afin de pouvoir reconstruire une information 3D, il faut d'abord passer par une étape de calibration. Cela consiste à évaluer les matrices de projection qui nous font passer du monde 3D au monde 2D. Dans le cadre de ce stage on a employé des méthodes dites de calibration forte, c'est à dire en utilisant une mire placée devant la caméra et dont on connaît parfaitement les coordonnées dans l'espace. Les principales étapes sont : l'acquisition des données, la mise en correspondance des points 2D avec les points 3D et l'évaluation de la matrice de projection. J'ai bénéficié pour réaliser ces étapes d'un logiciel déjà développé à l'Inria nommé Tele2. Ce programme effectue l'ensemble des étapes évoquées précédemment de manière plus ou moins automatique et en tout cas avec une grande rapidité. La méthode employée pour le calcul de la matrice fondamentale est celle de Toscani [6]. Cette méthode n'est pas la plus précise mais elle nous permettra par la suite de fournir une solution initiale aux algorithmes pour une calibration plus fine. La figure 7 donne un exemple d'acquisition utilisée pour la calibration. La mire utilisée est formée de 3 plans mais on peut se contenter de deux plans pour effectuer une calibration.

### 3.3.3 Reconstruction 3D

Une fois le système calibré, il est possible d'effectuer une reconstruction en 3D d'un objet placé en face de la caméra.

#### Méthode employée

On utilise de nouveau le logiciel Tele2 comme pour la calibration. On commence par sélectionner sur l'image de gauche un point. Le logiciel nous donne sur l'image de droite la droite épipolaire correspondante et on propose un point à mettre en correspondance en utilisant un critère de ressemblance basé sur une mesure de corrélation. On peut toutefois choisir un autre point de la droite si la proposition n'est pas satisfaisante. Une fois ces deux points mis en correspondance, on effectue une triangulation pour calculer les coordonnées 3D du point correspondant. Ces coordonnées sont



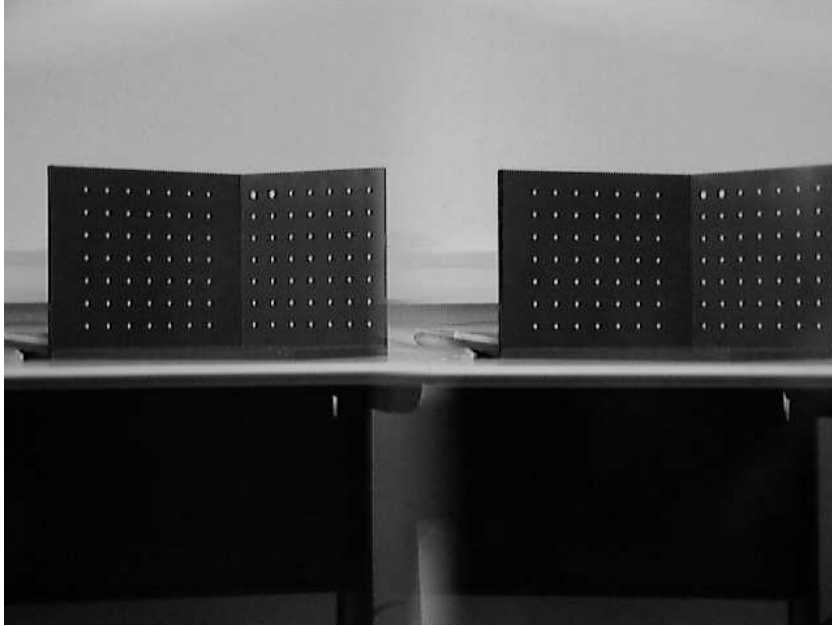


FIG. 7 – Image utilisée pour la calibration

exprimées dans le repère de la caméra de gauche. On procède ainsi pour l'ensemble des points souhaités.

Pour des applications pratiques une méthode plus automatique est mise en oeuvre. On utilise un détecteur pour obtenir des points d'intérêt qui seront facilement repérables d'une image à l'autre. On ne demande pas confirmation du point associé dans l'image de droite et on reconstruit directement les points en 3D. Cependant, en phase de développement, l'utilisation manuelle est plus pratique pour contrôler les résultats.

### Résultats

On reprend l'objet de la figure 6 pour effectuer cette reconstruction 3D.

Les points sélectionnés sont représentés par les carrés rouges sur la figure 8. On les a choisis parmi les carrés verts qui correspondent à des points d'intérêt. Ces points ont été obtenus à l'aide d'un détecteur suivant la méthode d'Harris et Stephens[8].

La figure 9 nous donne une image de la reconstruction obtenue. Elle correspond à la reprojexion sur l'image de gauche des points 3D obtenus par triangulation.

Une telle reconstruction nous permet déjà de nombreuses applications. On peut tout d'abord calculer de nouvelles vues du modèle 3D. Il suffit de se donner une nouvelle matrice de projection et de calculer les coordonnées de chacun des projeté des points 3D. C'est ce qui a été fait pour obtenir la figure 10.

On peut aussi profiter des coordonnées 3D pour mesurer des distances ou des angles de l'objet observé. Des expériences ont été faites dans ce sens. Les résultats obtenus sont encourageants mais ils ne fournissent guère plus qu'une approximation moyenne des grandeurs physiques. Pour améliorer ces résultats il faudra tout d'abord augmenter la précision de la calibration et de l'appariement des points.



FIG. 8 – Points sélectionnés pour la reconstruction

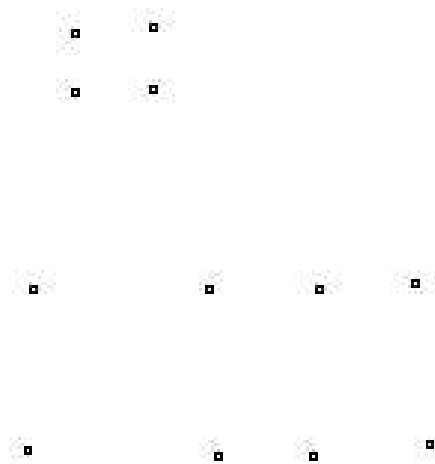


FIG. 9 – Reconstruction obtenue

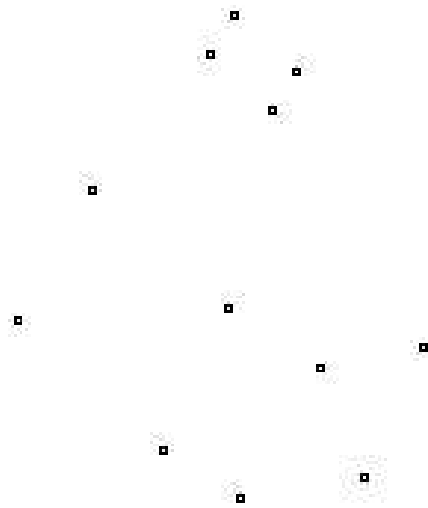


FIG. 10 – Synthèse d'une nouvelle vue à partir de la reconstruction



$$D_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & a \\ 0 & -1 & 0 & a \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Le deuxième miroir est situé à une distance  $d = b \sin \beta - a \cos \beta$  et fait un angle  $\beta$  avec l'axe des  $y$  ce qui nous donne le résultat suivant :

$$D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2 \sin^2 \beta & -\sin 2\beta & -d \sin \beta \\ 0 & -\sin 2\beta & 1 - 2 \cos^2 \beta & -d \cos \beta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On obtient alors la modélisation du système de gauche en effectuant le produit  $D_1 D_2$ .

$$D_1 D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sin 2\beta & \cos 2\beta & b \sin 2\beta - a \cos 2\beta \\ 0 & -\cos 2\beta & \sin 2\beta & a + b - b \cos 2\beta - a \sin 2\beta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Pour terminer, on pose  $\theta = 2\beta - \frac{\pi}{2}$  et on a le résultat suivant :

$$D_1 D_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & b \cos \theta + a \sin \theta \\ 0 & \sin \theta & \cos \theta & a + b + b \sin \theta - a \cos \theta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On obtient alors un déplacement dont l'angle de la rotation est  $\theta$ . On peut brièvement commenter ce résultat. Les deux miroirs placés verticalement entraînent deux réflexions qui en se combinant génèrent un déplacement (au sens rotation puis translation). L'angle de la rotation est  $\theta$  et correspond comme on s'y attend au double de l'angle formé par les deux miroirs. De plus la translation est dans un plan normal à l'axe de la rotation et on a donc un mouvement planaire.

On obtient des résultats similaires pour le système de droite.

On pose ici  $\theta' = \frac{\pi}{2} - 2\beta'$  et on a :

$$D'_1 D'_2 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta' & -\sin \theta' & a \sin \theta' - b \cos \theta' \\ 0 & \sin \theta' & \cos \theta' & a + b - b \sin \theta' - a \cos \theta' \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

En notant  $R$  et  $t$  la rotation et la translation qui permettent de passer du repère de la mire au repère de la caméra comme indiqué dans la figure 11, on peut dès lors donner les matrices de projection des systèmes de gauche et de droite :

$$M = \begin{pmatrix} K & 0 \end{pmatrix} D_1 D_2 \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

$$M' = \begin{pmatrix} K & 0 \end{pmatrix} D'_1 D'_2 \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$

### 4.3 Reconstruction 3D dans le repère caméra

Maintenant que la modélisation est effectuée, on est capable de faire une reconstruction 3D dans le repère de la caméra réelle. Jusqu'à présent on pouvait seulement effectuer cette reconstruction dans les repères des caméras virtuelles gauche et droite étant donné qu'on n'avait pas mis en évidence l'effet des miroirs. Cela suffisait pour calculer des distances ou des déplacements mais cela nous empêchait de positionner l'objet par rapport à la caméra réelle.

On note  $(X, Y, Z, T)$  les coordonnées projectives du point recherché.

On note  $(u_1, v_1)$  et  $(u_2, v_2)$  les coordonnées de ce point dans les images de gauche et de droite.

On doit donc avoir :

$$\begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} \simeq (K \ 0) D_1 D_2 \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

$$\begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \simeq (K \ 0) D'_1 D'_2 \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

On en déduit alors que  $(X, Y, Z, T)$  est un élément non nul du noyau de :

$$\begin{pmatrix} a_u & (u_0 - u_1) \sin \theta & (u_0 - u_1) \cos \theta & \delta_2 (u_0 - u_1) \\ 0 & (v_0 - v_1) \sin \theta + a_v \cos \theta & (v_0 - v_1) \cos \theta - a_v \sin \theta & a_v \delta_1 + (v_0 - v_1) \delta_2 \\ a_u & (u_0 - u_2) \sin \theta' & (u_0 - u_2) \cos \theta' & \delta'_2 (u_0 - u_1) \\ 0 & (v_0 - v_2) \sin \theta' + a_v \cos \theta' & (v_0 - v_2) \cos \theta' - a_v \sin \theta' & a_v \delta'_1 + (v_0 - v_1) \delta'_2 \end{pmatrix}$$

avec

$$\begin{aligned} \delta_1 &= b \cos \theta + a \sin \theta \\ \delta_2 &= a + b + b \sin \theta - a \cos \theta \\ \delta'_1 &= a \sin \theta' - b \cos \theta' \\ \delta'_2 &= a + b - b \sin \theta' - a \cos \theta' \end{aligned}$$

On a donc le moyen, par une équation linéaire, d'obtenir une reconstruction 3D des points qu'on a appariés dans les deux images. Cependant cette équation ne servirait dans une implémentation pratique qu'à fournir une solution initiale pour le calcul de la solution. On préfère en effet des algorithmes qui nous permettent de quantifier l'erreur commise en effectuant les calculs. On pourra donc utiliser par exemple un algorithme de type Levenberg-Marquardt pour minimiser l'erreur de reprojection.

On a donc besoin pour ce calcul des paramètres de la modélisation. Leur extraction fait l'objet de la partie suivante de ce rapport.

### 4.4 Calibration du système

Le but de la modélisation effectuée précédemment a été de mettre en évidence les paramètres du système. On va maintenant essayer d'extraire ces paramètres par une calibration. Cela nous permettra d'effectuer la reconstruction 3D dans le repère caméra présentée dans le paragraphe précédent.

#### 4.4.1 Calibration avec une paire d'images

La première question à se poser avant de se lancer dans la calibration est de savoir quels sont les paramètres qui pourront être extraits. En effet on a l'intuition qu'on perd un peu l'idée du positionnement de la scène observée par rapport à la caméra. On peut ainsi imaginer obtenir des scènes

similaires en tournant le système et en changeant l'orientation des miroirs. C'est malheureusement ce qu'on va constater en s'intéressant à la calibration avec une unique paire d'images.

La calibration plus fine du système revient à essayer de factoriser deux matrices de projection mise sous la forme canonique 2.4.1 en deux matrices mises sous la forme obtenue lors de la modélisation 4.2. Cette factorisation doit être unique pour permettre l'extraction de l'ensemble des paramètres.

Considérons le mouvement qui nous fait passer du monde 3D vu par le système gauche à celui vu par le système droit. Il est obtenu par la matrice suivante :

$$T = D_1 D_2 (D'_1 D'_2)^{-1}$$

c'est à dire :

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\theta - \theta') & -\sin(\theta - \theta') & (a + b) \sin(\theta - \theta') + 2b \cos \theta \\ 0 & \sin(\theta - \theta') & \cos(\theta - \theta') & (a + b)(1 - \cos(\theta - \theta')) + 2b \sin \theta \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Cela nous permet de mettre en évidence que, s'il est facile d'extraire  $\theta - \theta'$ , il est impossible d'extraire les quatre paramètres  $\theta$ ,  $\theta'$ ,  $a$  et  $b$  puisqu'il y a une infinité de solutions. On peut s'en convaincre en se donnant une valeur pour  $\theta$  et en considérant le système d'équation formé par les éléments  $T_{24}$  et  $T_{34}$  de la matrice et avec pour inconnues  $a$  et  $b$ . On trouve alors pour chaque valeur de  $\theta$  un couple solution  $(a, b)$ .  $\theta'$  s'obtient indépendamment puisqu'on connaît  $\theta - \theta'$ .

On peut tout de même affiner la calibration du système avec une seule paire d'images en imposant la forme de  $T$ , c'est à dire un mouvement planaire.

Ceci a malheureusement une conséquence importante sur le système. L'impossibilité d'extraire les paramètres physiques entraîne que toute reconstruction ne pourra se faire que dans le repère d'une des caméras virtuelles. On est en effet dans l'incapacité d'extraire  $R$  et  $t$  et on ne connaît donc pas la transformation entre le repère mire et le repère caméra. Concrètement cela signifie qu'on ne pourra pas effectuer de mesure de positionnement d'un objet par rapport à la caméra. On pourra toujours évaluer les déplacements d'un objet d'une position à une autre mais sans jamais savoir s'il est en face de la caméra ou s'il est à sa gauche et que les miroirs sont orientés dans ce sens.

#### 4.4.2 Calibration avec deux paires d'images

On a vu précédemment qu'il était impossible d'extraire tous les paramètres physiques du système avec une seule paire d'images. Même si on peut se contenter de la calibration obtenue pour de nombreuses applications, on a cependant envie de compléter ce résultat. Pour ce faire on va utiliser deux paires d'images. La seule différence entre les deux acquisitions est qu'on aura fait varier un des angles de réglage des grands miroirs. Considérons qu'on a modifié  $\theta'$  uniquement. On désigne par  $\theta''$  la deuxième valeur.

On applique tout d'abord la calibration précédente sur chacune des paires d'images. On obtient deux matrices de déplacement  $T$  et  $T'$  entre les repères gauche et droit.  $T$  correspond à l'angle  $\theta'$  (respectivement  $T'$  à  $\theta''$ ). On exploite alors le fait que  $\theta$  est le même pour les deux systèmes et on déduit les relations suivantes avec la dernière colonne des deux matrices  $T$  et  $T'$  :

$$2b \cos(\theta) = \frac{T'_{24} \sin(\theta - \theta') - T_{24} \sin(\theta - \theta'')}{\sin(\theta - \theta') - \sin(\theta - \theta'')}$$

$$2b \sin(\theta) = \frac{T'_{34}(1 - \cos(\theta - \theta')) - T_{34}(1 - \cos(\theta - \theta''))}{\cos(\theta - \theta') - \cos(\theta - \theta'')}$$

De plus on a la possibilité d'identifier dans les matrices les valeurs de  $\cos(\theta - \theta')$ ,  $\cos(\theta - \theta'')$ ,  $\sin(\theta - \theta')$  et  $\sin(\theta - \theta'')$ . On peut alors en déduire les valeurs de  $b$  et  $\theta$ . Comme on connaît les valeurs de  $\theta - \theta'$  et  $\theta - \theta''$ , on a donc aussi les valeurs de  $\theta'$  et  $\theta''$ . Enfin  $a$  s'obtient facilement en reprenant une des équations précédentes.

On est donc en mesure de calibrer complètement le système mais il faut posséder deux paires d'images pour lesquels les paramètres sont tous identiques sauf un des angles des grands miroirs.

### 4.4.3 Implémentation pratique

L'idée générale de la calibration est d'améliorer une calibration effectuée de manière classique et à partir de laquelle on aura extrait des valeurs initiales pour les paramètres physiques du système.

Comme étape préliminaire, on procède donc à la calibration des deux caméras séparément. Cela revient à effectuer l'acquisition d'une image de la mire, une mise en correspondance de points 2D avec des points 3D et une évaluation des matrices de projection. On extrait ensuite de ces matrices de projection les paramètres intrinsèques et extrinsèques de la caméra.

On va ensuite calibrer les matrices de projection pour les mettre sous la forme suivante :

$$M = \begin{pmatrix} K & 0 \end{pmatrix} \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix}$$
$$M' = \begin{pmatrix} K & 0 \end{pmatrix} \begin{pmatrix} R' & t' \\ 0 & 1 \end{pmatrix}$$

On utilise un algorithme de type Levenberg-Marquardt avec lequel on minimise l'erreur de reprojection en jouant sur les 16 paramètres du système (4 paramètres intrinsèques et deux fois six paramètres extrinsèques). Les valeurs initiales nous sont fournies par l'étape précédente. On force ainsi l'égalisation des paramètres intrinsèques. On a de plus la possibilité d'évaluer le déplacement entre les repères des caméras gauche et droite :

$$T = \begin{pmatrix} R & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} R' & t' \\ 0 & 1 \end{pmatrix}^{-1}$$

Deux possibilités s'offrent à nous à cette étape de la méthode.

Si on ne dispose que d'une seule paire d'images on ne pourra pas extraire tous les paramètres et il ne nous reste plus qu'à utiliser à nouveau un algorithme de type Levenberg-Marquardt pour forcer  $T$  à être de la forme suivante :

$$T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) & u \\ 0 & \sin(\alpha) & \cos(\alpha) & v \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

On a donc calibré notre système avec treize paramètres : quatre paramètres extrinsèques, six paramètres pour le déplacement dans le repère mire et trois paramètres pour le déplacement relatif entre les deux caméras.

Si on dispose par contre de deux paires d'images on peut effectuer la calibration complète décrite précédemment. On évalue les deux déplacements relatifs  $T$  et  $T'$ . On extrait les paramètres du système et on les injecte dans un algorithme qui minimise l'erreur de reprojection sur les quatre images en même temps. On a quinze paramètres qui évoluent : quatre paramètres intrinsèques, six paramètres pour le déplacement vers le repère mire et cinq paramètres du système.

Cette seconde méthode de calibration est un peu désagréable puisqu'on doit agir sur le système physiquement pour le calibrer entièrement. Cependant c'est la seule manière pour obtenir une calibration complète. On a sinon une calibration qui nous place dans le repère d'une des caméras et on ne peut pas faire de mesure de positionnement par rapport à la caméra. Par contre on pourra toujours mesurer des déplacements d'objets ou faire de la reconstruction dans un repère virtuel avec la première méthode de calibration.

### 4.4.4 Résultats

La réalisation pratique d'un logiciel m'a été demandée. Il s'appuie sur les résultats de Tele2 pour effectuer le calcul des matrices de projection du système et pour en extraire les paramètres extrinsèques et intrinsèques. Ces valeurs nous permettent d'initialiser comme décrit précédemment nos algorithmes de minimisation.

On effectue des essais avec trois caméras différentes pour voir comment le type d'objectif peut influencer la qualité de la calibration. Ces trois caméras sont :



- SONY EVI G21
- IMASYS avec objectif 8 mm
- IMASYS avec objectif KINOPTIK 12,5 mm

On effectue une calibration complète avec deux paires d'images. Le critère de qualité est l'erreur de reprojection. On la mesure pour la calibration de chacune des paires d'images (erreur 1 et erreur 2) et pour la calibration complète (erreur finale). Les résultats obtenus sont les suivants :

<i>Caméra</i>	<i>erreur 1</i>	<i>erreur 2</i>	<i>erreur finale</i>
SONY EVI G21	0,1 pixel	0,1 pixel	0,35 pixel
IMASYS avec objectif 8 mm	0,25 pixel	0,25 pixel	0,35 pixel
IMASYS avec objectif KINOPTIK 12,5 mm	0,05 pixel	0,1 pixel	0,16 pixel

Ces résultats sont obtenus en prenant un angle  $\theta' - \theta''$  aussi grand que possible. Cela permet de rendre la calibration beaucoup plus précise. En effet si cet angle est trop faible le système devient numériquement assez instable et ne peut faire mieux que dans le cas où on a une seule paire d'images. C'est une contrainte assez importante au moment de la calibration et cela rend cette étape assez peu commode.

On remarque aussi que la qualité de la calibration dépend du type d'objectif utilisé. Ceci concerne notamment la distorsion qui apparaît lorsqu'on s'éloigne du centre de l'image. En effet en calibrant le système, on effectue des calculs uniquement sur la partie gauche de l'image puis uniquement sur la partie droite de l'image alors qu'en temps normal on étudie des cibles relativement centrées. On peut reproduire ce phénomène en calibrant une caméra seule et en faisant évoluer la position de la mire. On s'aperçoit alors que le centre optique a tendance à se déplacer en fonction de cette position de la mire. Ce déplacement est d'autant plus important que l'objectif est de moins bonne qualité. On imagine aisément les effets sur notre système puisque la mire est toute à gauche ou toute à droite. Les résultats qui suivent permettent d'avoir l'ordre de grandeur de la variation des paramètres intrinsèques. Ils ont été obtenus avec l'objectif KINOPTIK et la caméra Imasys.

<i>Type de calibration</i>	$a_u$	$a_v$	$u_0$	$v_0$
Image de gauche	-1331,2	1346,9	271,5	437,9
Image de droite	-1520,7	1537,3	275,0	382,9
Paire d'images	-1469,5	1485,1	266,2	391,6

Les résultats obtenus avec ces caméras sont dans l'ensemble satisfaisants et ils laissent envisager la possibilité de nombreuses applications relativement précises.

## 5 Utilisation du capteur en mouvement

### 5.1 Problématique

Le but de l'étude menée dans la fin de ce rapport est d'imaginer la méthode à mettre en oeuvre pour créer un capteur 3D utilisant son mouvement autour d'un objet pour créer une reconstruction précise de celui-ci. On peut s'imaginer un tel appareil comme un appareil photo 3D que l'on déplacerait autour d'un objet pour en récupérer un modèle dans l'espace.

Le principal problème auquel on est confronté pour toute reconstruction est celui de la mise en correspondance des points. Lorsqu'on utilise une séquence d'images la mise en correspondance est facile d'une image à l'autre. Cependant on aimerait utiliser les images gauche et droite pour calculer cette reconstruction de manière précise. Or pour faire une mise en correspondance de l'image de gauche vers l'image de droite, la calibration ne nous donne qu'une droite épipolaire le long de laquelle on doit chercher un appariement ce qui est trop coûteux. Le principal dé est alors de faire à moindre frais la mise en correspondance stéréo.

On dispose pour résoudre ce problème d'un système considéré comme parfaitement calibré. On pourra donc pour nos besoins utiliser en toute confiance le déplacement entre les caméras de gauche et de droite, ou bien la géométrie épipolaire. C'est donc avec deux points mis en correspondance dans une paire d'images qu'on voudrait effectuer la reconstruction 3D. On profite ainsi du fait que notre système génère des paires d'images parfaitement synchronisées. Il n'y a donc aucun risque de détériorer la qualité de la reconstruction par un quelconque décalage temporel.

### 5.2 Méthode développée

On envisage une solution relativement modulaire, chaque élément peut donc être traité séparément et modifié sans remettre en cause l'ensemble du travail. Quelques uns des éléments de cette méthode proviennent de [10] qu'on pourra consulter pour plus de détails.

#### 5.2.1 Détection de points

Notre but n'est pas de faire de la corrélation dense et d'apparier tous les points d'une image de gauche avec celle de droite. Il est au contraire de reconstruire un nombre limité de points mais apportant une grande quantité d'information. On va donc devoir extraire ce qu'on appelle des points d'intérêts. Les choix possibles sont nombreux et variés. Deux éléments sont à prendre en compte pour évaluer la qualité d'un détecteur de points d'intérêts : la précision de la localisation et la capacité à retrouver dans plusieurs images le même point détecté.

On utilise souvent dans cette optique le détecteur de coins d'Harris et Stephens [8]. Un coin correspond à un brusque changement d'intensité du signal lumineux dans les deux directions. L'avantage d'un tel détecteur est que l'information exploitée est facilement perceptible pour l'observateur humain, sensible au changement de contraste et donc aux coins dans l'image.

La figure 12 fournit un exemple d'application d'un détecteur de coins de type Harris et Stephens à une image. Les points détectés sont marqués en vert.

#### 5.2.2 Suivi de points

On va tout d'abord considérer séparément les deux séquences d'images. Dans chacune des images d'une séquence on détecte des points d'intérêts et on va effectuer l'appariement de ces points d'une image à l'autre.

Cet appariement est beaucoup plus facile à effectuer que l'appariement stéréo. On a en effet un déplacement des points qui est très faible. On peut donc effectuer l'appariement en utilisant une mesure de corrélation avec tous les points situés dans une fenêtre assez réduite autour de la position initiale du point d'intérêt. On peut visualiser un exemple de résultat d'un programme de suivi de points sur la figure 13.

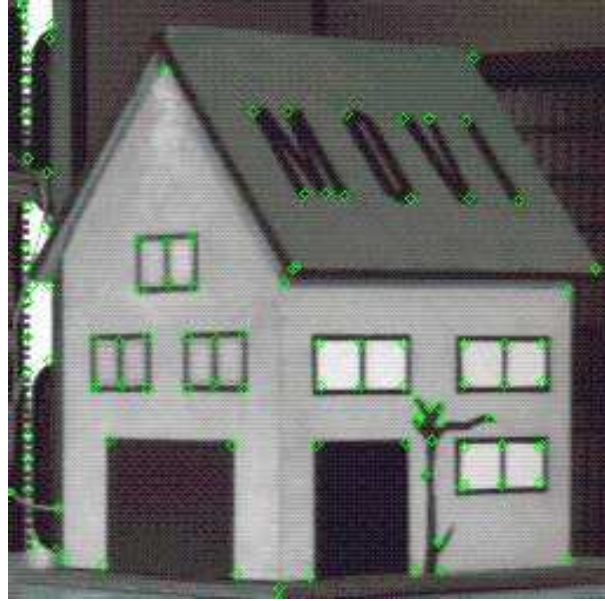


FIG. 12 – Utilisation d'un détecteur de coins sur une image

Afin de se faciliter la tâche il est possible d'utiliser des images obtenues avec un intervalle de temps très court. Le chemin parcouru est alors très court et le suivi se fait facilement. Par contre on aura besoin pour la suite de déplacements relativement importants. On ne retient donc pas les résultats intermédiaires et on conservera juste en quelque sorte l'appariement final de tout une séquence d'images. On a donc des appariements dans deux séquences monoculaires et ces deux séquences sont reliées par une calibration effectuée au préalable. On résume cette situation par le schéma 14. On y trouve aussi la numérotation des images qui sera utilisée dans la suite du rapport.

### 5.2.3 Extraction du mouvement d'une caméra

On va à partir de maintenant faire l'hypothèse que l'objet observé par la caméra est resté fixe pendant le déplacement de celle-ci lors de l'acquisition de la séquence d'images. Les points d'intérêts qu'on a suivi dans l'étape précédente sont donc l'illustration du seul mouvement de la caméra. Il est donc naturel d'essayer d'extraire par leur intermédiaire le mouvement de la caméra.

Tout d'abord on va effectuer une calibration faible du système. Cela revient à évaluer la matrice fondamentale  $F_{12}$  (c'est à dire correspondant au système formé par les images (1) et (2)) ainsi que la matrice  $F_{34}$ . Cette évaluation est rendue possible par la mise en correspondance qui a été effectuée lors du suivi de points. Dès lors il ne nous reste plus qu'à utiliser une des nombreuses méthodes d'évaluation de la matrice fondamentale. On pourra consulter notamment [9] pour avoir un aperçu globale des méthodes possibles ainsi que leurs avantages et leurs inconvénients respectifs. Il faut toutefois garder à l'esprit que cette évaluation n'est pas toujours très précise. De nombreux cas d'incertitude sont en effet possibles selon le type de mouvement de la caméra.

Comme on connaît les paramètres intrinsèques de la caméra, on est capable d'extraire à partir de  $F$  la rotation du mouvement de la caméra. On obtient aussi la translation mais seulement à un facteur près. En effet, connaissant  $K$ , on peut calculer la matrice essentielle  $E$  du système. A partir de celle-ci on peut extraire ces informations sur le mouvement en appliquant par exemple la factorisation proposée dans [7].

Il nous reste deux facteurs à identifier pour connaître parfaitement le mouvement de chacune des caméras. Or on a, à tout instant, la relation suivante qui est vérifiée :



FIG. 13 – Résultat du suivi de points d'intérêt

$$D_s D_{12} = D_{34} D_s$$

On a noté  $D_s$  le mouvement qui nous fait passer du repère de la caméra de gauche à celui de la caméra de droite. Il est bien connu après l'étape de calibration.

Cette équation matricielle nous fournit trois équations pour identifier les deux scalaires recherchés. On est donc capable d'extraire complètement le mouvement de chacune des caméras de ce système stéréo rigide.

#### 5.2.4 Alternative : utilisation de l'approximation par une homographie

L'estimation proposée précédemment de la matrice fondamentale peut s'avérer hasardeuse dans certains cas particuliers. Une autre idée peut être exploitée pour pallier cette difficulté. On va chercher une homographie 2D amenant les points de (1) aux points de (2). Si on note  $p_1$  un point de 1 et son correspondant  $p_2$  dans (2), on doit donc trouver  $H$ , matrice de taille  $3 \times 3$ , vérifiant :

$$p_2 \simeq H p_1$$

L'existence d'une telle matrice est acquise dans le cas d'une rotation autour du centre de projection de la caméra. On peut cependant espérer obtenir une bonne approximation dans les autres cas et notamment dans les cas où l'évaluation de  $F$  serait trop hasardeuse. L'intérêt de cette approximation va apparaître dans une des solutions proposées dans la partie suivante pour la mise en correspondance stéréo.

#### 5.2.5 Mise en correspondance stéréo

On va maintenant proposer trois voies possibles pour effectuer la mise en correspondance stéréo sans recherche le long de la droite épipolaire. On utilisera pour ce faire les informations obtenues précédemment sur le mouvement de chaque caméra. Il faut noter que ces trois méthodes permettent de suggérer des appariements possibles ou des zones de recherche. Elles sont donc à compléter par l'utilisation d'un critère local de comparaison de deux points comme par exemple une mesure de corrélation. Cependant on aura effectué une grande avancée puisqu'on utilisera ce critère sur une très faible portion de l'image et non le long de toute la droite épipolaire comme c'était le cas auparavant.

#### Passage par la reconstruction 3D

On se place dans le cas où on a évalué les mouvements  $D_{12}$  et  $D_{34}$  pour les deux séquences d'images.

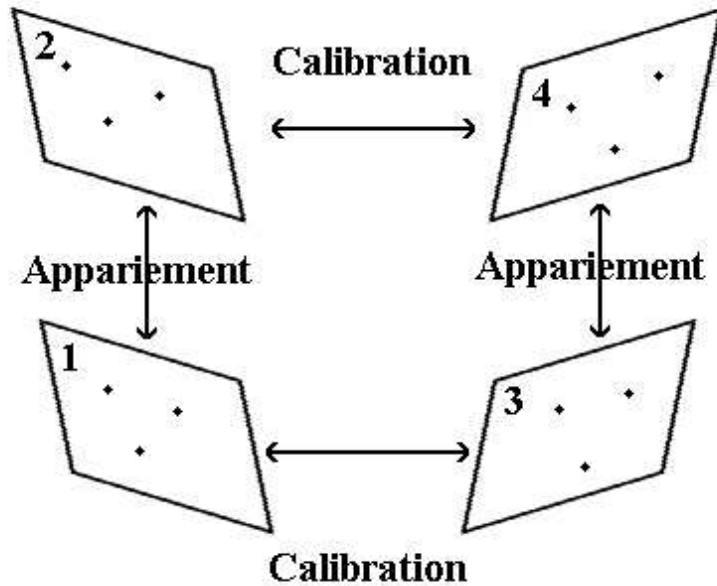


FIG. 14 – Situation après le suivi de points

On a donc pour un point qu'on aura suivi dans une séquence les deux équations suivantes :

$$\begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} \simeq P \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix} \text{ et } \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \simeq PD_{12} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

On peut ainsi déterminer  $(X, Y, Z, T)$  puisqu'on connaît la matrice de projection  $P$  grâce à la calibration et que  $(u_1, v_1)$ ,  $(u_2, v_2)$  et  $D_{12}$  ont été déterminés lors des étapes précédentes.

Dans chacune des deux caméras, on est donc capable d'effectuer une reconstruction 3D d'un certain nombre de points. De plus on connaît le mouvement pour passer d'un repère caméra à l'autre. Cela nous permet de superposer ces deux nuages de points et d'effectuer un certain nombre d'appariements. Ces appariements peuvent s'effectuer tout simplement en fonction de la distance entre deux points 3D. Un seuil doit être défini et on tente d'apparier tous les points vérifiant cette condition. Pour valider un appariement on vérifie tout de même que la mesure de corrélation entre les projections dans les séquences d'images est bonne.

### Reprojection dans la deuxième séquence

Cette deuxième méthode n'est en fait qu'une évolution de la première.

On commence par évaluer les coordonnées 3D d'un point dans une séquence. Puis on le projette dans les images de l'autre séquences. On connaît en effet le déplacement entre les deux repères et les matrices de projection par la calibration.

Ce point nous permet de définir autour de ses projections dans les images une petite fenêtre de recherche dans laquelle doit se trouver le point correspondant dans l'autre séquence. Il est tout à fait possible que les projections arrivent en dehors de l'image. Dans ce cas on a reconstruit un point qui n'est pas vu par l'autre caméra. On effectue une mesure de corrélation avec tous les points présents dans la fenêtre qu'on a pu définir. Si l'un d'entre eux s'avère concluant on obtient un appariement.

### Utilisation de la géométrie épipolaire

On va ici utiliser la géométrie épipolaire pour effectuer des appariements. On connaît déjà  $F_{13}$  par la calibration et on a donc pour chaque point de (1) une droite de (3) sur laquelle doit se trouver son image. L'idée de cette méthode est d'évaluer  $F_{23}$  pour avoir une deuxième droite épipolaire. Le point correspondant se trouvera donc alors à l'intersection des deux droites.

Pour évaluer  $F_{23}$  deux possibilités s'offrent à nous.

Si on a évalué le déplacement  $D_{12}$  de façon complète, on est donc capable de calculer  $D_{23}$  puisque la calibration nous donne le déplacement  $D_{13}$ . On a alors la possibilité de calculer  $F_{23}$  avec la formule suivante :

$$F_{23} \simeq K^{-T}[t_{23}]_x R_{23} K^{-1}$$

où  $[t]_x$  représente l'application produit vectoriel avec  $t$  et  $K$  est la matrice des paramètres intrinsèques obtenue par la calibration.

Si par contre on s'est contenté d'évaluer une homographie entre les images (1) et (2), on doit raisonner différemment.

On a obtenu une matrice  $H$  vérifiant :

$$p_1 = H^{-1}p_2$$

Or on a :

$$p_3^T F_{13} p_1 = 0$$

d'où :

$$p_3^T (F_{13} H^{-1}) p_2 = 0$$

ce qui nous donne :

$$F_{23} \simeq F_{13} H^{-1}$$

On a donc obtenu une évaluation de  $F_{23}$  grâce à l'approximation du déplacement par  $H$  et à la calibration précédemment effectuée.

L'intersection des deux droites nous donne un point dans l'image (3). Cependant il ne faut pas s'imaginer que ce point sera le bon à coup sûr. On se donne à nouveau une petite fenêtre dans laquelle on recherche parmi les points d'intérêt celui qui correspond le mieux au sens de la mesure de corrélation.

### 5.2.6 Reconstruction finale

La dernière étape de notre méthode consiste à affiner la reconstruction et les déplacements qu'on aura évalués. Cela se fait en tirant partie des appariements qui ont pu être effectués dans l'étape précédente.

Tout d'abord on effectue la reconstruction 3D des points appariés dans les deux séquences. Ces points sont présents dans les quatre images. Afin d'améliorer le positionnement on prend en compte ces quatre contraintes dans un algorithme minimisant l'erreur de reprojection. On obtient ainsi un point 3D dont le positionnement devrait être relativement précis.

Une fois ces points 3D calculés, on va s'appuyer sur eux pour évaluer plus précisément le déplacement de chaque caméra. En effet on dispose pour tous les points de la séquence de gauche des deux équations suivantes :

$$\begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} \simeq P \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix} \text{ et } \begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} \simeq PD_{12} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Cela nous donne la possibilité d'évaluer à nouveau et de manière plus précise le déplacement  $D_{12}$ . De la même manière on calcule de nouveau le déplacement  $D_{34}$ .

Enfin on utilise les nouvelles valeurs de  $D_{12}$  et  $D_{34}$  pour évaluer à nouveau les coordonnées 3D des points des deux séquences qui n'ont pas été appariés. On espère ainsi améliorer leur positionnement ce qui nous permettra peut-être de les mettre en correspondance à la prochaine étape de l'algorithme.

Comme l'illustre le schéma 15, la situation est la suivante à la fin de cette dernière étape. Certains points (en rouge) ont été appariés entre les deux images permettant le début de la fusion des deux reconstructions 3D. Ils permettent aussi d'améliorer le positionnement des points qui n'ont pas trouvé de correspondant (en vert ou en bleu). De tels points peuvent être obtenus pour de nombreuses raisons : manque de précision dans leur reconstruction, occultation dans la deuxième séquence ou tout simplement absence du champ de vue de la deuxième caméra.

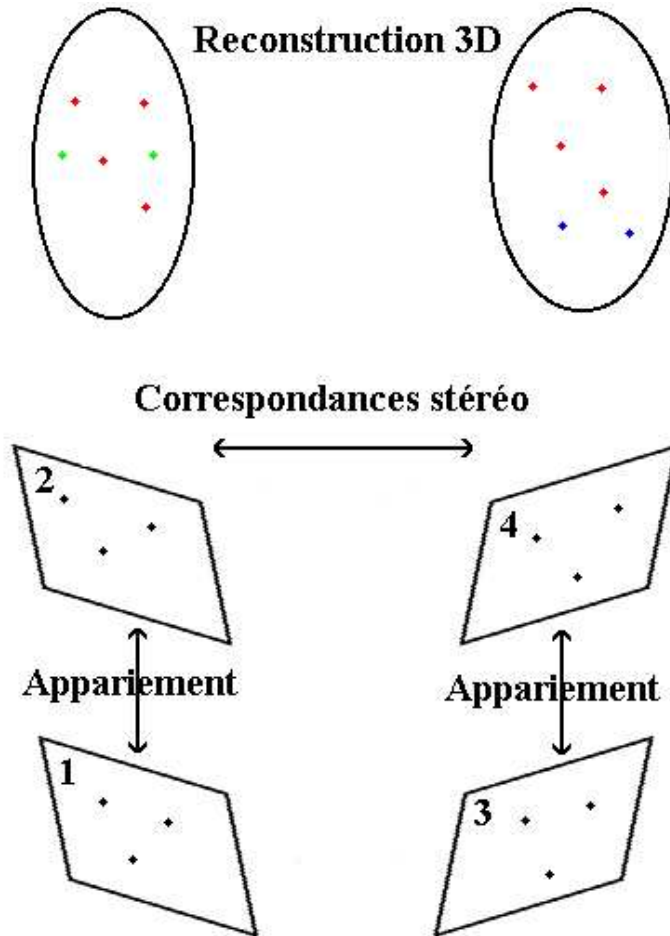


FIG. 15 – Situation à la fin de l'algorithme

### 5.3 Perspectives

On a donc vu une méthode complète permettant d'utiliser le mouvement du capteur pour se passer de la coûteuse mise en correspondance stéréo par corrélation. On a juste utilisé une courte séquence d'images et il faut itérer le procédé à nouveau en améliorant les reconstructions déjà effectuées et en calculant de nouveaux appariements et de nouveaux points 3D.

De nombreuses perspectives d'évolution sont possibles pour améliorer le capteur et son utilisation. On a par exemple supposé un objet immobile pour ne pas avoir à prendre en compte d'autres mouvements que celui de la caméra. Il est aussi possible d'envisager une utilisation avec un objet mobile mais il faudra envisager des méthodes pour segmenter l'image afin de différencier les parties fixes et les parties mobiles de la scène.

De plus on a présenté ici une méthode sur une courte séquence en s'intéressant uniquement aux problèmes du domaine de la vision par ordinateur. Or on doit aussi envisager les structures de données à mettre en oeuvre pour stocker au mieux la reconstruction 3D. On aura besoin en effet dans une utilisation concrète de connaître les points reconstruits, de savoir dans quelles images on les a trouvés et il en est de même pour les points reconstruits uniquement à l'aide des séquences monoculaires. Les performances du système dépendront beaucoup de la réponse à cette question.

Il reste donc encore beaucoup à faire avant d'avoir une implémentation complète d'un tel capteur mais cette méthode nous fournit la démarche détaillée dans laquelle on doit l'inscrire. Cette application est parfaitement adaptée aux caractéristiques de notre système : rigidité du système stéréo et synchronisation parfaite des flux vidéo.



## 6 Conclusion

Le but de mon travail était d'évaluer la qualité du système à quatre miroirs comme capteur 3D et de nombreux éléments de réponse ont été fournis dans ce cadre.

Tout d'abord ce système peut très bien s'utiliser comme système virtuel de deux caméras et il est très commode à employer pour peu qu'on adapte un peu quelques outils logiciels au fait que les deux vues sont présentes dans la même image. On apprécie notamment la conception physique du système qui rend possible le réglage du champ de vision selon les besoins de l'utilisateur.

De plus on a apporté à ce système une modélisation projective qui n'avait pas été faite. Elle a permis deux avancées importantes : calibrer complètement et plus précisément le système et reconstruire un modèle 3D dans le repère de la caméra réelle. Ce dernier résultat peut s'avérer très important dans une utilisation pratique du capteur. On pense par exemple au positionnement d'un engin utilisant ce capteur pour ses déplacements. Cette modélisation est d'autant plus importante qu'une évolution possible du système est de motoriser le déplacement des grands miroirs. Il est donc très important de connaître précisément le modèle du système pour prendre en compte un changement d'orientation.

Enfin on a développé une méthode pour utiliser ce capteur lors d'un mouvement autour d'une cible fixe. Malheureusement le stage n'a pas permis de réaliser l'implémentation complète de cette méthode mais les résultats théoriques sont encourageants. On est capable de faire de la mise en correspondance stéréo sans avoir recours à des mesures de corrélation sur une grande partie de l'image. On a pu contourner cette énorme difficulté. Encore une fois les perspectives sont grandes puisqu'on peut envisager de créer ainsi un véritable appareil photo 3D.

Le système à quatre miroirs n'est donc pas encore totalement validé dans le sens où il reste quelques incertitudes sur la précision des reconstructions obtenues et sur la viabilité du modèle développé. Cependant les résultats obtenus lors de ce stage sont très encourageants et poussent à continuer l'utilisation de ce système dans le futur.

## 7 Remerciements

Ce stage au sein de l'équipe MOVI s'est déroulé dans une très bonne ambiance. L'accueil et le soutien apportés aux stagiaires fut très agréable.

Je tiens à remercier plus particulièrement Radu Horaud pour l'encadrement de mon stage et pour la constante disponibilité dont il a fait preuve.

Je remercie aussi de manière générale les personnels de l'unité de recherche Inria Rhône-Alpes et notamment les moyens robotiques dont les compétences m'ont été fort utiles.

Enfin j'apporte une mention particulière aux autres stagiaires de l'INRIA Rhône-Alpes qui ont ajouté à ce stage la détente nécessaire pour passer trois mois très agréables.

## Références

- [1] H. Mathieu et F. Devernay. Système de miroirs pour la stéréoscopie. *Rapport technique 0172, INRIA Sophia Antipolis, 1995*
- [2] J. Gluckman et S. Nayar. Planar catadioptric stereo : geometry and calibration. *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*
- [3] J. Gluckman et S. Nayar. Rectified catadioptric stereo sensors. *2000*
- [4] S.J. Maybank. Theory of reconstruction from image motion. *Springer-Verlag Berlin, 1993*
- [5] M. Ollis, H. Herman et S Singh. Analysis and design of panoramic stereo vision using equi-angular pixel cameras
- [6] O.D. Faugeras et G. Toscani. The calibration problem for stereo. *Proceedings of the 1986 Conference on Computer Vision and Pattern Recognition*
- [7] O. Faugeras. Three-dimensional computer vision - A geometric viewpoint. *1993*
- [8] C. Harris et M. Stephens. A combined corner and edge detector. *Alvey vision conference 1988*
- [9] O. Faugeras et Q. T. Luong. The geometry of multiple images. *2001*
- [10] F. Dornaika et R. Chung. Stereo correspondence from motion correspondence. *Proceedings of the 1999 Conference on Computer Vision and Pattern Recognition*
- [11] D. Demirdjian. Le mouvement projectif. Théorie et applications pour l'autocalibrage et la segmentation du mouvement.