

Du traitement de l'information visuelle à la planification de trajectoires : Application à la robotique mobile d'assistance.

Mohamed Moncef Ben Khelifa, Moez Bouchouicha

Docteurs en Sciences et Technologie Industrielles,

Postdocs au Laboratoire SIS/AI (Systèmes Information et Signal / Automatique et Images)

khelifa@univ-tln.fr, + 33 4 94 14 20 77

moez@univ-tln.fr, + 33 4 94 14 24 93

Adresse professionnelle

Université de Toulon et du Var | BP 132 | F-83957 La Garde Cedex

Résumé

Les systèmes Robotiques d'assistance aux personnes handicapées, visent à apporter une aide, par exemple au pilotage des fauteuils roulants robotisés, tout en répondant aux besoins d'aide à la planification de trajectoire, et à la navigation en utilisant des techniques de vision dynamique issues de la robotique mobile. Dans cette thèse nous avons étudié les analogies qui existent entre la neuro-vision humaine et la vision artificielle dans le but de pallier aux problèmes des déficits visuo-spatiaux chez des IMC sur fauteuils roulants.

Dans un premier temps nous avons traité le problème de la détection des points d'intérêt par une implantation stable des lissages du signal avec des filtres Gaussiens suivie d'une approximation du maximum local à des positions non entières. Le détecteur ainsi développé nous a donné une précision sous-pixellique des points d'intérêt.

La répétabilité de ce détecteur, nous a permis de calculer la géométrie épipolaire caractérisée par une stabilité de la matrice fondamentale, celle-ci jouant un rôle très important dans la recherche des primitives à apparier dans une séquence d'images.

Une approche stratifiée de calibration projective, affine et Euclidienne a été utilisée dans le but d'autocalibrer une caméra CCD embarquée. La calibration affine nous a permis d'estimer une matrice stable de l'homographie du plan à l'infini, suivie d'une estimation des paramètres intrinsèques et extrinsèques en utilisant une calibration euclidienne basée sur la méthode de la conique absolue (équations de Kruppa simplifiées).

Enfin, nous validons ces algorithmes par deux expériences : l'une consiste en une planification de trajectoire suivie d'une navigation le long d'un couloir, l'autre en l'évitement d'obstacles dans un environnement interne.

Mots clés : déficits neuro-visuels, robotique d'assistance, détecteur de Harris, géométrie épipolaire, autocalibration, planification de trajectoire, navigation.

1. Qui sont les IMC ?

L'*IMC* (Infirme Moteur Cérébral) est une personne victime de lésions cérébrales précoces (anté-, péri-, ou postnatales, mais survenues avant l'âge de un an), stables (non évolutives), responsables d'une déficience motrice patente.

2. Les Grandes fonctions neurovisuelles

Parmi la complexité des symptômes neurovisuelles, il est capital de pouvoir analyser chacun d'eux, le replacer dans son cadre structurel, éviter confusions et amalgames au vu des manifestations de surface, en comprendre les mécanismes sous-jacents et en déduire les conséquences fonctionnelles. La figure 1 résume et synthétise les principaux modules cliniquement exploitables en *pathologie neurovisuelle* chez l'enfant IMC, ceci afin de comprendre les mécanismes et les spécificités : il est indispensable pour appréhender les symptômes pathologiques et leurs répercussions dans tous les secteurs cognitifs qui rendent la vie quotidienne d'un enfant polyhandicapé sur fauteuil roulant parfois difficile.

Nous distinguerons :

Les voies afférentes : *sensori-gnosiques*, voies de la vision ;

Les voies efférentes : voies de saisie visuelle, voies du regard.

3. Le fauteuil roulant M.A.R.H. (Mobile Autonome Robotisé pour Handicapés)

Il s'agit d'un concept de fauteuil roulant robotisé figure 2, et innovant de part sa constitution, son mode de fonctionnement et la prise en compte de la sécurité de la personne transportée. Il est l'aboutissement de plusieurs années de réflexion menées par un ergothérapeute, de l'Institut de Rééducation Fonctionnelle Pomponiana.Olbia de Hyères (Var), Monsieur Jean-Pierre BELHEUR [Bel. 95].

Le M.A.R.H est un robot mobile destiné au transport d'une personne handicapée qui se compose d'une base motorisée qu'on appelle l'Esclave, portant un fauteuil électrique qu'on appelle le Maître, pilotés par une commande adaptée aux handicaps de l'utilisateur [Ben et al. 99].

Cette base est équipée de cinq roues, deux motrices et trois directrices et d'une ceinture de détecteurs à infrarouges, ce qui permet à l'utilisateur de s'affranchir de toutes les difficultés habituelles de contournement d'obstacles, de franchissement de portes, de déplacement le long d'une table ou d'un plan de travail...

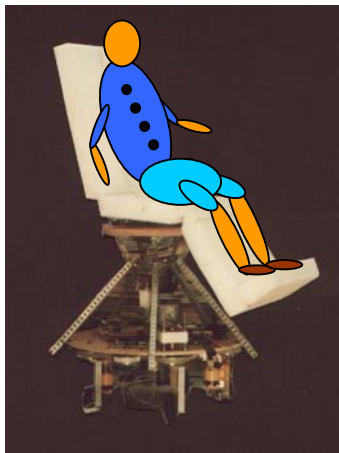


Fig. 2 : Le fauteuil roulant, M.A.R.H

Le siège est réglable en hauteur et en position, entre deux limites : horizontale et verticale. Cette fonctionnalité est rendue nécessaire par la diversité des handicaps rencontrés. De plus, le siège peut effectuer une rotation complète sur sa base grâce à un moto-réducteur électrique. Ceci permet à la personne handicapée, une diminution importante des commandes à effectuer et une plus grande précision de positionnement.

4. Conséquences

Dans notre projet, consacré à la robotique d'assistance aux enfants IMC, nous rappelons qu'il

existe un lien entre dyspraxie et vision. Ces enfants porteurs de dyspraxie constructive sont les mêmes que ceux souffrant des troubles de la vision, cette association ne saurait être fortuite, mais là encore les spécialistes en constatent la quasi-constance, sans pouvoir en tirer d'argument définitif sur les liens qu'entretiennent ces pathologies.

Il est patent que certaines praxies se construisent ou s'affinent chez l'enfant sous le contrôle de la vision. D'une façon générale, les spécialistes dans ce domaine insistent sur la nécessaire coordination d'une part entre l'œil (direction du regard, saisie fovéale, calibrage des saccades, ...) et d'autre part la vision périphérique, la position de la tête dans l'espace, la motricité, la proprioception et la kinesthésie.

Quoi qu'il en soit, on ne peut nier que la pathologie du regard induit certaines caractéristiques de la pathologie gestuelle, ce qui se traduit dans la pratique par l'utilisation du terme "dyspraxie visuo-spatiale". On souligne ainsi, non seulement le fait que la dyspraxie s'accompagne de troubles du regard et des difficultés spatiales qui s'y rattachent, mais surtout que les troubles de l'assemblage, au premier plan, sont imbriqués avec des difficultés d'analyse visuelle de la tâche ou du modèle, influencés par les aspects spatiaux de la tâche (relations topologiques des divers éléments, obliques, miroirs, orientation des différentes pièces, ...), difficultés surajoutées, attribuées aux anomalies dites "visuo-spatiales".

5. Vision par ordinateur

Nous allons maintenant envisager, un moyen de compenser le problème de dyspraxie visuo-spatiale par l'utilisation des outils de la vision par ordinateur qui visent à apporter une aide, par exemple au pilotage des fauteuils roulants robotisés, tout en répondant aux besoins d'aide à la planification de trajectoire, et à la navigation. Ces techniques de vision dynamique sont issues de la robotique mobile.

6. Détection des points d'intérêt dans une scène : Détecteur de Harris amélioré

Beaucoup d'applications de vision par ordinateur dépendent d'un calibrage ou d'un autocalibrage précis permettant de déterminer les caractéristiques intrinsèques et extrinsèques des caméras CCD. Mais ces travaux d'autocalibration ne se basent pas sur la détection des points d'intérêts (pixels des images) avec une résolution sous-pixellique (en dessous du pixel).

Dans ce paragraphe, nous présentons l'étape initiale de notre algorithme d'autocalibration de caméra CCD et d'appariement basé sur l'extraction de points d'intérêt dans une scène, à l'aide du

détecteur de Harris [Har 88], à une résolution en dessous du pixel.

6.1. Le détecteur de Harris

Harris a amélioré l'approche de Moravec [Mor 81] en calculant une matrice liée à la fonction d'auto-corrélation qui prend en compte les valeurs des dérivées premières du signal sur une fenêtre. Ceci est une amélioration par rapport à Moravec, car la discrétisation utilisée pour calculer la fonction d'auto-corrélation, due au déplacement et à la direction choisis, n'est plus nécessaire. Il obtient donc la matrice \mathbf{M} suivante :

$$\mathbf{M} = \begin{pmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) & \left(\frac{\partial I}{\partial y}\right)^2 \end{pmatrix} \quad (1)$$

Les valeurs propres de cette matrice sont les courbures principales de la fonction d'auto-corrélation. Si ces deux courbures sont grandes, ceci indique la présence d'un point d'intérêt. L'utilisation des courbures est améliorée par l'utilisation de la valeur minimale comme l'avait proposé Moravec. Toutefois, pour ne pas extraire les valeurs propres, Harris utilise une mesure reposant sur le déterminant et la trace de la matrice, et qui est supérieure à zéro dans le cas d'un coin :

$$R = \det(\mathbf{M}) - k(\text{trace}\mathbf{M})^2$$

avec $k = 0.04$ (d'après l'expérimentation de Harris).

6.2 Lissage multi-échelle (Harris précis)

Il est connu que le calcul numérique de la dérivée est un problème délicat, en effet même un bruit faible peut modifier le résultat de façon importante. Pour illustrer ce manque de robustesse, considérons les fonctions $f(x)$ et $\hat{f}(x) = f(x) + \xi \sin(\omega x)$, [Sch. 96]. Elles sont similaires, si ξ est petit. Toutefois $f'(x)$ peut différer beaucoup de $\hat{f}'(x)$ pour un ω grand ($\xi \ll \omega$). En conséquence un bruit de relativement haute fréquence dans une image peut modifier considérablement la première dérivée et donc a fortiori les dérivées d'ordre supérieur.

Préalablement à tout calcul de dérivation, il est donc nécessaire d'effectuer un lissage. Puisque la différenciation commute avec la convolution :

$\partial_i(g * f) = g * \partial_i f = \partial_i g * f$, un tel lissage peut être obtenu, soit en lissant l'image, soit en lissant l'opérateur de dérivation. Un moyen simple pour stabiliser les calculs de dérivation est donc d'utiliser les dérivées d'une fonction de lissage. Le choix le plus communément fait pour une telle fonction est la gaussienne [Wit 83, Tor 86, Rom 94, Flo 93, Lin 94]. La formule de la gaussienne $G(\mathbf{x}, \sigma)$ pour un \mathbf{x} de dimension 2 est :

$$G(\mathbf{x}, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{\mathbf{x}^2}{2\sigma^2}} \quad (2)$$

La dérivée n-ième de cette fonction par rapport aux variables x_k ($k = 1, 2$) est la suivante :

$$G_{i_k i_n}(\mathbf{x}, \sigma) = \frac{\partial^n}{\partial_{i_1} \partial_{i_2} \dots \partial_{i_n}} G(\mathbf{x}, \sigma) \quad n = 0 \wedge N$$

où i_k indique l'axe du système de référence et donc $i_k = 1, 2$ dans le cas d'un système de dimension 2.

Les convolutions avec les dérivées des gaussiennes intervenant dans le calcul sont implémentées de manière récursive ce qui permet des calculs rapides [Der 93]. Pour effectuer ces calculs récursifs la gaussienne et ses dérivées sont approximées de façon très précise.

6.3. Points d'intérêt à une résolution sous-pixellique (Harris amélioré)

La version ci-dessus nous donne seulement des points d'intérêt en des positions entières. Il est possible de détecter et localiser leurs positions sous-pixellique.

Soit le signal multi-échelle interpolé et défini par :

$$W(x) = \sum_i w(i) g_\sigma(x - i),$$

Les positions sous-pixelliques sont celles pour lesquelles la dérivée du signal interpolé W est maximale. Il est possible d'interpoler le signal entier (pour calculer $W(x)$ pour chaque réel x).

Nous traitons les maxima locaux de $W'(x)$ comme un problème de maximisation. Etant donnée une estimation grossière de la position d'un coin x_0 , nous voulons obtenir le maxima local de $W'(x)$, c'est à dire trouver un x lorsque $W''(x)$ est nul.

En utilisant la méthode de Raphson Newton nous obtenons l'algorithme itératif suivant :

$$x_{k+1} = x_k - \frac{W''(x_k)}{W'''(x_k)} \quad (3)$$

La méthode de Raphson Newton converge très rapidement, mais manque de robustesse, ainsi, nous avons limité le pas des itérations comme suit :

$$x_{k+1} = x_k - \begin{cases} 0.3 & \text{si } \frac{W''(x_k)}{W'''(x_k)} > 0.3 \\ -0.3 & \text{si } \frac{W''(x_k)}{W'''(x_k)} < -0.3 \\ \frac{W''(x_k)}{W'''(x_k)} & \text{par ailleurs} \end{cases}$$

La figure 3 compare les trois versions Harris standard, et les versions stabilisées, Harris précis et Harris amélioré. Le critère de comparaison est le taux de répétabilité pour la $i^{\text{ème}}$ image est défini par :

$$r_i(\varepsilon) = \frac{|D(\varepsilon)|}{\min(n_i, n_1)} \quad (4)$$

avec $|D(\varepsilon)|$ l'ensemble des points détectés à la même position est déterminé par un seuil de distance ε , et (n_i, n_1) le nombre de points détectés dans les images respectives.

On peut observer que la version stabilisée donne de meilleurs résultats. Dans le cas d'une rotation les résultats obtenus avec la version standard se détériorent pour les angles 30 et 120 degrés. Ceci est dû au fait que cette implémentation standard du détecteur de Harris utilise des filtres trop discrets pour approximer de façon précise et isotrope les dérivées gaussiennes.

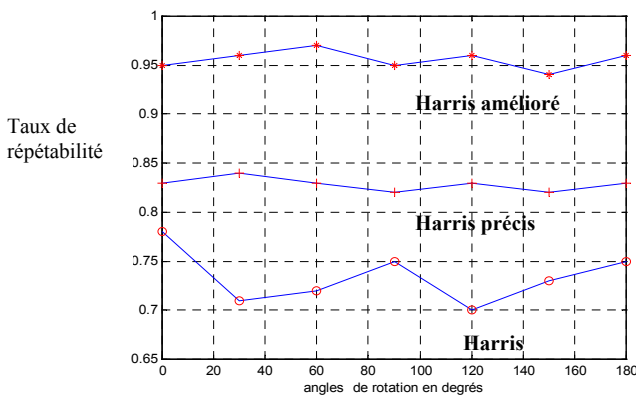


Fig. 3 : Comparaison de Harris, Harris précis, et Harris amélioré ($\varepsilon = 2$ pixels)

7. Géométrie épipolaire

La matrice fondamentale décrit une relation géométrique entre deux images prises sous deux angles de vues différents appelée géométrie

épipolaire. Dans le cas d'images obtenues avec une ou plusieurs caméras non étalonnées, la connaissance de la matrice fondamentale est considérée comme un étalonnage faible « weak calibration ». La relation avec la matrice essentielle introduite par Longuet-Higgins [Lon. 81] a été montré dans [Luo. 92]. En particulier, quand les coordonnées normalisées sont utilisées, la matrice fondamentale coïncide avec la matrice essentielle. De ce fait, cette matrice joue un rôle central pour résoudre des problèmes de vision tel que le calcul linéaire de structures projectives, l'autocalibration, la poursuite de points dans une séquence d'images et le calcul du mouvement d'une caméra.

Dans ce paragraphe nous traitons le calcul de la géométrie épipolaire, entre images successives d'une séquence sur la base d'un certain nombre de correspondances considérées connues. Cette géométrie épipolaire est représentée par une matrice 3×3 appelée *matrice fondamentale*. Cette matrice étant de rang deux et définie à un coefficient multiplicatif près, au minimum huit correspondances sont nécessaires pour permettre son estimation de façon unique, robuste et rapide.

7.1. Estimation robuste de la géométrie épipolaire

L'équation vérifiée par \mathbf{F} s'écrit pour tout couple de points images appariés $(\mathbf{m}, \mathbf{m}')$:

$$\mathbf{m}'^T \mathbf{F} \mathbf{m} = 0 \quad (5)$$

\mathbf{F} est de rang 2 et au minimum huit correspondances images sont nécessaires pour la calculer.

Avec $\mathbf{m} = (u, v, 1)^T$ et $\mathbf{m}' = (u', v', 1)^T$. Chaque couple d'appariement $(\mathbf{m}, \mathbf{m}')$ fournit donc une équation du type $\mathbf{U}^T \mathbf{f} = 0$ qui est une équation linéaire et homogène pour les neuf éléments de la matrice \mathbf{F} .

7.1.1. Critère linéaire

La première méthode pour estimer la matrice fondamentale est basée sur le fait que l'équation $\mathbf{m}'^T \mathbf{F} \mathbf{m} = 0$ est linéaire et homogène pour les neuf coefficients de \mathbf{F} :

$$\mathbf{U}^T \mathbf{f} = 0$$

où

$$\mathbf{U} = [uu', vv', u'u', uv', v'v', v'u', u, v, 1]^T$$

$$\mathbf{f} = [F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{31}, F_{32}, F_{33}]^T$$

Si l'on dispose de huit appariements $(\mathbf{m}, \mathbf{m}')$, on peut trouver une solution unique pour \mathbf{F} . En pratique, on dispose d'un nombre n bien plus grand d'appariements images $(\mathbf{m}_i, \mathbf{m}'_i)$ ($i = 1, K, n$). Alors, on peut résoudre les équations

$\mathbf{U}^T \mathbf{f} = 0$ au sens des moindres carrés, ce qui revient à minimiser le critère

$$\min_{\mathbf{F}} \left[\sum_{i=1}^n (\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i)^2 \right], \quad \text{réécrit comme}$$

$$\min_{\mathbf{f}} \|\tilde{\mathbf{U}} \mathbf{f}\|^2 \text{ sous la contrainte } \|\mathbf{f}\|^2 = 1, \text{ où :}$$

$$\tilde{\mathbf{U}} = \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{M} \\ \mathbf{U}_n^T \end{bmatrix}$$

La solution \mathbf{f} est le vecteur propre associé à la plus petite valeur propre de la matrice $\tilde{\mathbf{U}}^T \tilde{\mathbf{U}}$.

La méthode utilisée pour résoudre ce critère est non itérative (« closed form ») et le calcul est très sensible au bruit, malgré l'utilisation d'un grand nombre d'appariements, et que la contrainte $\det(\mathbf{F}) = 0$ (rang de \mathbf{F} égal 2) ne soit pas satisfaite. Par ailleurs, le critère n'étant pas normalisé, un biais est introduit sur la localisation des épipoles. Pour pallier ce problème, nous introduisons d'autres critères non linéaires.

7.1.2. Critères non linéaires

Soit $(\mathbf{m}_i, \mathbf{m}'_i)$, ($i = 1, K, n$) un ensemble d'appariements images. L'équation $\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i = 0$ est directement liée avec la distance d du point \mathbf{m}'_i par rapport à sa droite épipolaire $\mathbf{l}'_i = \mathbf{F} \mathbf{m}_i$. La première idée serait, alors d'utiliser un critère non linéaire minimisant la somme :

$$\sum_{i=1}^n d^2(\mathbf{m}'_i, \mathbf{F} \mathbf{m}_i) \text{ avec :}$$

$$d(\mathbf{m}'_i, \mathbf{F} \mathbf{m}_i) = \frac{|\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i|}{\sqrt{(\mathbf{F} \mathbf{m}_i)_1^2 + (\mathbf{F} \mathbf{m}_i)_2^2}}$$

où le terme $(\mathbf{F} \mathbf{m}_i)_i$ signifie $i^{\text{ème}}$ élément du vecteur $\mathbf{F} \mathbf{m}_i$.

Pour donner un rôle symétrique aux deux images et pour obtenir une géométrie épipolaire cohérente, il convient de minimiser le critère suivant :

$$\sum_{i=1}^n \{d^2(\mathbf{m}'_i, \mathbf{F} \mathbf{m}_i) + d^2(\mathbf{m}_i, \mathbf{F}^T \mathbf{m}'_i)\}$$

où

$$d(\mathbf{m}_i, \mathbf{F}^T \mathbf{m}'_i) = \frac{|\mathbf{m}_i{}^T \mathbf{F}^T \mathbf{m}'_i|}{\sqrt{(\mathbf{F}^T \mathbf{m}'_i)_1^2 + (\mathbf{F}^T \mathbf{m}'_i)_2^2}}$$

est la distance du point \mathbf{m}_i à sa ligne épipolaire $\mathbf{F}^T \mathbf{m}'_i$.

Compte tenu du fait que $\mathbf{m}_i{}^T \mathbf{F}^T \mathbf{m}'_i = \mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i$, on aboutit finalement à la minimisation du critère suivant (qu'on appelle critère *DIST*) qui est un critère normalisé :

Critère *DIST* :

$$\sum_{i=1}^n \left[\frac{(\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i)^2}{(\mathbf{F} \mathbf{m}_i)_1^2 + (\mathbf{F} \mathbf{m}_i)_2^2} + \frac{(\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i)^2}{(\mathbf{F}^T \mathbf{m}'_i)_1^2 + (\mathbf{F}^T \mathbf{m}'_i)_2^2} \right]$$

La minimisation de ce critère est un problème de minimisation non linéaire que l'on peut résoudre par une méthode itérative du type Levenberg-Marquardt [Pho. et al. 95]. La valeur initiale est celle fournie par le critère linéaire présenté ci-dessus.

Une autre façon d'améliorer le critère linéaire est la suivante. Les coordonnées des appariements images $(\mathbf{m}_i, \mathbf{m}'_i)$, ($i = 1, K, n$), sont généralement entachées d'un certain bruit. On peut alors multiplier chaque terme sommé $(\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i)^2$ du critère linéaire par un coefficient de telle sorte que tous les termes aient la même covariance. Si le bruit est de moyenne nulle, isotrope, uniformément réparti et indépendant selon chaque axe image, alors on montre que le nouveau critère à minimiser est un critère normalisé (qu'on appelle critère *GRAD*) [Luo. 92] :

Critère *GRAD* :

$$\sum_{i=1}^n \frac{(\mathbf{m}'_i{}^T \mathbf{F} \mathbf{m}_i)^2}{(\mathbf{F} \mathbf{m}_i)_1^2 + (\mathbf{F} \mathbf{m}_i)_2^2 + (\mathbf{F}^T \mathbf{m}'_i)_1^2 + (\mathbf{F}^T \mathbf{m}'_i)_2^2}$$

Ce critère à minimiser est semblable au précédent. Il peut également être minimisé par une méthode numérique itérative du type Levenberg-Marquardt.

Résultats et analyse de stabilité

Nous présentons ici deux types de résultats. Le premier concerne la validation avec des données réelles. Nous montrons expérimentalement qu'en utilisant uniquement le critère linéaire la solution obtenue est presque aussi correcte que la solution utilisant l'étape non linéaire, ce est loin d'être le cas

de la solution linéaire obtenue avec la méthode classique basée sur l'équation (5), [Luo 92].

Le second type de résultats concerne une analyse de stabilité et de précision des calculs avec des données réelles.

Dans toutes nos expériences avec les images réelles, les appariements de points sont obtenus automatiquement avec la méthode de « Harris amélioré » décrite dans le paragraphe ci-dessus.

Pour tester la qualité d'une matrice fondamentale \mathbf{F} étant donné un ensemble d'appariements $(\mathbf{m}_i, \mathbf{m}'_i)$, $i = 1 \dots n$, nous avons utilisé la formule suivante :

$$Q_F = \frac{\sum_{i=1}^n d(\mathbf{m}'_i, \mathbf{F}\mathbf{m}_i) + d(\mathbf{m}_i, \mathbf{F}^T \mathbf{m}'_i)}{2n} \quad (6)$$

avec Q_F la qualité de la matrice fondamentale \mathbf{F} et d la distance euclidienne d'un point à la droite épipolaire en pixels. Notons qu'ici les points images sont donnés par leurs coordonnées pixels. En d'autres termes, Q_F est la distance moyenne des points à leurs droites épipolaires.

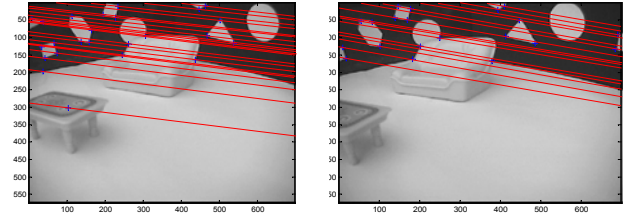
Nous avons testé cet algorithme avec une caméra CCD à focale fixe, $f = 4 \text{ mm}$, embarquée sur un robot miniature « Khepera », le tout est relié, via une liaison série RS232 à un PC (Pentium II, 266 MHz). Une carte graphique Meteor II est utilisée pour l'acquisition des images de résolution 576×768 . Nous avons préparé une scène « maison miniature » pour nos tests comportant deux parties :

- Une scène d'intérieur du « Séjour » (figure 4) : le mouvement entre les deux images est une rotation de 26° . Nous remarquons que nous avons une meilleure répétabilité des points détectés (ici 23 points). La qualité de la géométrie épipolaire est caractérisée par le facteur Q_F qui est au dessous du pixel dans les deux cas : linéaire et non linéaire, ceci limitera l'espace de recherche pour la mise en correspondance entre les points détectés dans chaque vue.
- Une scène d'intérieur du « Couloir Séjour / Cuisine » : le mouvement ici est une translation pure le long de l'axe optique de la caméra ; c'est à dire que la deuxième image est obtenue en avançant la caméra vers l'avant. Ici 96 points sont détectés. Nous pouvons remarquer que l'épipole se trouve presque au centre de l'image (figure 5).

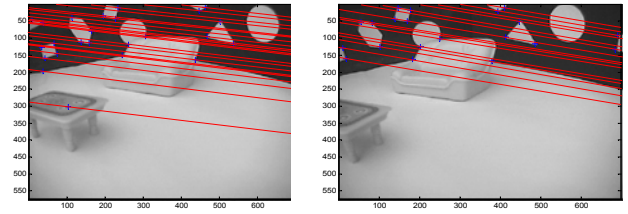
Ces résultats obtenus avec des données réelles pour différents mouvements est différentes scènes prouvent la stabilité du calcul linéaire de \mathbf{F} .

Ce dernier conditionne la convergence du calcul non linéaire.

Dans la méthode classique, le calcul linéaire de \mathbf{F} est très instable ; la convergence du calcul non linéaire n'est alors plus garantie.

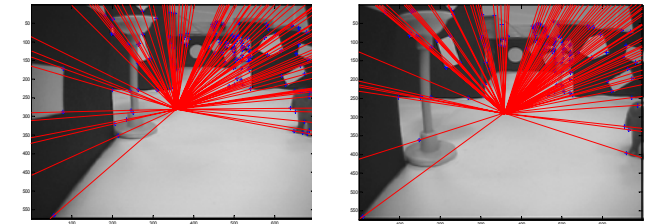


Calcul linéaire de la matrice fondamentale,
 $Q_F = 0.51$ pixels

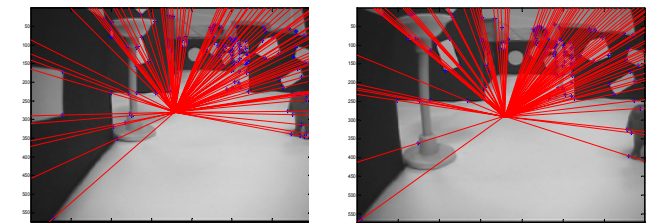


Calcul non linéaire de la matrice fondamentale,
 $Q_F = 0.22$ pixels

Fig. 4 : Résultat du calcul de la géométrie épipolaire (matrice fondamentale) pour une paire d'images d'intérieur « Séjour » : la caméra a effectuée une rotation de 26° vers la gauche.



Calcul linéaire de la matrice fondamentale,
 $Q_F = 0.39$ pixels



Calcul non linéaire de la matrice fondamentale,
 $Q_F = 0.19$ pixels

Fig. 5 : Résultat du calcul de la géométrie épipolaire (matrice fondamentale) pour une paire

d'images d'intérieur « Couloir Séjour / cuisine » : la caméra a effectuée une translation de 5 cm vers l'avant.

$$\lambda \mathbf{m}'_{\infty} = [\mathbf{K} \quad \mathbf{0}] \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{M}_{\infty} \\ \mathbf{0} \end{bmatrix} = \mathbf{K} \mathbf{R} \mathbf{M}_{\infty}$$

$$\Rightarrow \lambda \mathbf{m}'_{\infty} = \mathbf{K} \mathbf{R} \mathbf{M}_{\infty} = \mathbf{K} \mathbf{R} \mathbf{K}^{-1} \mathbf{m}_{\infty} = \mathbf{H}_{\infty} \mathbf{m}_{\infty} \quad (7)$$

8. Autocalibration de la caméra CCD

Nous présentons dans ce paragraphe le rôle très important que la connaissance de l'équation du plan infini peut jouer pour le calibrage d'une caméra mobile. Ces réflexions que nous présentons sont inspirées des travaux de Zeller et Faugeras et de Luong et Faugeras [Zel. 96], [Luo. 94].

L'approche qui suit est basée sur la conique absolu et le plan infini. La propriété principale que nous allons exploiter est la suivante : une translation laisse le plan à l'infini invariant. Nous allons démontrer qu'ayant les correspondances images, dans deux vues différentes prises par une caméra en mouvement planaire, de quelques points de la scène se trouvent très loin par rapport à la caméra, il est possible de calculer l'homographie du plan de l'infini. Ayant l'homographie du plan de l'infini de deux vues différentes, il est possible de calculer entre deux vues successives les paramètres intrinsèques et extrinsèques de cette caméra.

Nous pouvons alors utiliser une approche stratifiée qui consiste à déterminer d'abord la structure projective, puis affine, et ensuite, la structure euclidienne de la scène. La structure affine est obtenue en déterminant le plan à l'infini à partir de différents mouvements. Les paramètres intrinsèques sont alors calculés à partir des équations de Kruppa, mettant en jeu une ou plusieurs homographies à l'infini entre des vues consécutives.

8.1. Estimation de la matrice \mathbf{H}_{∞}

Afin de calculer les paramètres intrinsèques de la caméra, nous devons calculer la matrice \mathbf{H}_{∞} qui représente l'homographie du plan à l'infini. Soit donc la caméra en mouvement. Considérons le repère qui lui est associé à l'instant τ . Nous allons utiliser ce repère comme repère de référence. Exprimons la projection des points scène à l'infini dans le repère de référence. Si \mathbf{K} est la matrice des paramètres intrinsèques, alors :

$$\lambda \mathbf{m}_{\infty} = [\mathbf{K} \quad \mathbf{0}] \begin{bmatrix} \mathbf{M}_{\infty} \\ \mathbf{0} \end{bmatrix} = \mathbf{K} \mathbf{M}_{\infty}$$

avec λ un facteur d'échelle, qu'on appelle aussi la profondeur.

La caméra se déplace, et à l'instant $\tau + \Delta\tau$ elle prend une autre image de la scène. Si $[\mathbf{R} \quad \mathbf{t}]$ est son déplacement entre les instants τ et $\tau + \Delta\tau$, alors on a :

D'après l'équation (7), la matrice \mathbf{H}_{∞} lie les projections images de tout point scène appartenant au plan à l'infini. Pour la déterminer il faut au minimum quatre correspondances images $(\mathbf{m}_i, \mathbf{m}'_i)$ ($i = 1, \dots, n$) provenant de points scène appartenant au plan à l'infini. Pour cela, il suffit de considérer au moins 4 points de l'horizon mis en correspondance dans une séquence d'images (au moins deux vues).

Si plusieurs correspondances images sont disponibles (avec $n \geq 4$, n étant le nombre d'appariements), il est possible de minimiser le critère non linéaire du type :

$$\min_{\mathbf{H}_{\infty}} f(\mathbf{H}_{\infty}) = \sum_{i=1}^n [(\mathbf{m}'_i - \mathbf{H}_{\infty} \mathbf{m}_i)^2 + (\mathbf{m}_i - \mathbf{H}_{\infty}^T \mathbf{m}'_i)^2] \quad (8)$$

qui peut être minimisé en utilisant la méthode de Levenberg-Marquardt.

8.2. Estimation des paramètres intrinsèques

Nous présentons dans ce paragraphe une méthode originale permettant de calculer les paramètres intrinsèques et le déplacement relatif d'une caméra. Cette méthode s'appuie sur des outils de géométrie projective. Luong, Faugeras et Maybank, [Luo. 92], sont à la base de cette théorie dite « conique absolue et équations de Kruppa ».

De l'équation (7), nous pouvons remarquer que $\mathbf{H}_{\infty} = \mathbf{K} \mathbf{R} \mathbf{K}^{-1}$. Donc cette homographie ne contient que les composantes rotationnelles du mouvement de la caméra, elle est indépendante de la translation, comme on pouvait s'y attendre, en effet, il est connu que toute translation laisse les points à l'infini invariants.

Soit la matrice de rotation $\mathbf{R} = \mathbf{K}^{-1} \mathbf{H}_{\infty} \mathbf{K}$, et sachant que $\mathbf{R} \mathbf{R}^T = \mathbf{I}$, on a :

$$(\mathbf{K}^{-1} \mathbf{H}_{\infty} \mathbf{K}) (\mathbf{K}^{-1} \mathbf{H}_{\infty} \mathbf{K})^T = \mathbf{I}$$

$$\Rightarrow \mathbf{H}_{\infty} \mathbf{K} \mathbf{K}^T \mathbf{H}_{\infty}^T = \mathbf{K} \mathbf{K}^T$$

$$\Rightarrow \mathbf{H}_{\infty} \Delta \mathbf{H}_{\infty}^T = \Delta$$

où Δ est appelée matrice de Kruppa, qui définit la conique duale de ω , l'image de la conique Ω de l'espace scène se trouvant au plan de l'infini,

$$\text{or} \quad \Delta = \mathbf{B}^* = \det(\mathbf{B}) (\mathbf{B}^{-1})^T,$$

avec $\mathbf{B} = (\mathbf{K}^{-1})^T \mathbf{K}^{-1}$. Et d'après les solutions des équations de Kruppa, nous obtenons :

$$\Rightarrow \lambda \Delta = \lambda \mathbf{K} \mathbf{K}^T = \lambda \begin{bmatrix} u_0^2 + \frac{\alpha_u^2}{\sin^2 \theta} & u_0 v_0 - \frac{\alpha_u \alpha_v \cot \theta}{\sin \theta} & u_0 \\ u_0 v_0 - \frac{\alpha_u \alpha_v \cot \theta}{\sin \theta} & v_0^2 + \frac{\alpha_v^2}{\sin^2 \theta} & v_0 \\ u_0 & v_0 & 1 \end{bmatrix}$$

Par conséquent, si \mathbf{H}_∞ est connu, Δ peut être déterminé à l'aide de l'équation $\mathbf{H}_\infty \Delta \mathbf{H}_\infty^T = \Delta$ et par là, la matrice des paramètres intrinsèques \mathbf{K} de la caméra peut être entièrement récupérée. Deux vues suffisent pour déterminer l'homographie \mathbf{H}_∞ et donc les éléments de la matrice Δ . Toutefois pour des raisons de stabilité numérique, il est préférable d'utiliser deux déplacements (et donc trois vues), un premier générant une disparité horizontale et un second, une disparité verticale.

8.3. Détermination des longueurs en calibration affine

Une fois que la matrice homographie de l'infini est déterminée, le bi-rapports des longueurs de tout segments de la scène peut être calculé directement sur les images. En effet, soient $\mathbf{M}_1, \mathbf{M}_2$ et \mathbf{M}_3 trois points objet alignés, comme présentés sur la figure 6, à partir de leurs points images $(\mathbf{m}_1, \mathbf{m}'_1)$, $(\mathbf{m}_2, \mathbf{m}'_2)$ et $(\mathbf{m}_3, \mathbf{m}'_3)$, nous pourrions calculer les points images $(\mathbf{m}, \mathbf{m}')$ de l'intersection de cette ligne avec le plan de l'infini en utilisant \mathbf{H}_∞ . Nous pourrions calculer sur chaque image le bi-rapport de ces quatre points. Comme un invariant projectif, ce bi-rapport est donc égal au rapport de $\mathbf{M}_1, \mathbf{M}_2$ et \mathbf{M}_3 ; plus précisément :

$$\frac{\overline{\mathbf{M}_1 \mathbf{M}_3}}{\overline{\mathbf{M}_2 \mathbf{M}_3}} = \frac{\overline{\mathbf{m}_1 \mathbf{m}_3}}{\overline{\mathbf{m}_2 \mathbf{m}_3}} \quad \frac{\overline{\mathbf{M}_1 \mathbf{M}_3}}{\overline{\mathbf{M}_1 \mathbf{M}_\infty}} = \frac{\overline{\mathbf{m}_1 \mathbf{m}_3}}{\overline{\mathbf{m}_1 \mathbf{m}}}$$

8.4. Estimation des paramètres extrinsèques

Une fois la matrice des paramètres intrinsèques est connue, la matrice de rotation est calculée depuis l'équation :

$$\mathbf{R} = \mathbf{K}^{-1} \mathbf{H}_\infty \mathbf{K} \quad (9)$$

et pour calculer le vecteur de translation \mathbf{t} il suffit que le point scène \mathbf{M} n'appartient pas au plan de l'infini, donc :

$$\mathbf{m}' \neq \mathbf{H} \mathbf{m}$$

Ceci est vérifié par les équations suivantes :

$$\mathbf{m} = \begin{bmatrix} \mathbf{K} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{M} \\ 1 \end{bmatrix} \Rightarrow \mathbf{m} = \mathbf{K} \mathbf{M}$$

$$\mathbf{m}' = \begin{bmatrix} \mathbf{K} \mathbf{R} & \mathbf{K} \mathbf{t} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{M} \\ 1 \end{bmatrix} = \mathbf{K} \mathbf{R} \mathbf{M} + \mathbf{K} \mathbf{t}$$

$$\Rightarrow \mathbf{m}' = \mathbf{K} \mathbf{R} \mathbf{K}^{-1} \mathbf{m} + \mathbf{K} \mathbf{t} \Rightarrow \mathbf{H}_\infty \mathbf{m} + \mathbf{K} \mathbf{t}$$

Ayant un certain nombre de correspondances $(\mathbf{m}_i, \mathbf{m}'_i)$ ($i=1, K, n$) entre deux images successives de la séquence et ayant calculé l'homographie de l'infini et les paramètres intrinsèques de la caméra, le vecteur translation \mathbf{t} peut être récupéré (à un facteur d'échelle près) par résolution aux moindres carrés de l'équation :

$$\sum_{i=1}^n \|\mathbf{H}_\infty \mathbf{m}_i + \mathbf{K} \mathbf{t} - \mathbf{m}'_i\| \quad (10)$$

Expérimentations et résultats

Pour valider ces résultats théoriques, nous avons utilisé une caméra CCD (modèle VPC-735-PAL) de focale $f = 4mm$, figure 7, embarquée sur un robot miniature « Khepera », figure 8. La résolution de l'image est de (564×760) .



Fig. 7 : Caméra CCD, VPC-735-PAL

Fig. 8: Robot KHEPERA

Nous appliquons cet algorithme d'auto-calibration sur une séquence de 25 images prises dans la « Cuisine » de la « Maisonnée miniature » construite à cet effet. figure 9.

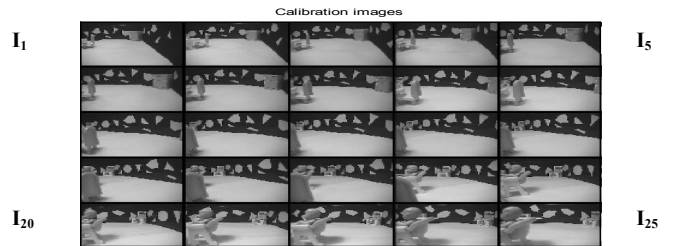


Fig. 9: Séquence « Cuisine », 25 prises de vue

Ces images présentent une répartition de niveaux de gris assez variable. Ceci est dû aux différentes formes et matières des objets observés dans la scène. La figure 10 représente la répartition de niveau de gris de la ligne numéro 154 de l'image I_{20} .

Cette propriété est assez importante pour l'auto-calibration. En effet, plus la répartition des niveaux de gris varie plus la disparité entre les primitives devient significative, figure 11. Ce qui permet d'éviter certaine ambiguïté dans la phase de mise en correspondance, due aux faux appariements.

Les points d'intérêt, coins et intersections des jonctions, sont détectés automatiquement sur toute la séquence grâce au détecteur de « Harris amélioré ». Des primitives d'appariement sont formées grâce à la variation significative de la répartition des niveaux de gris dans les images.

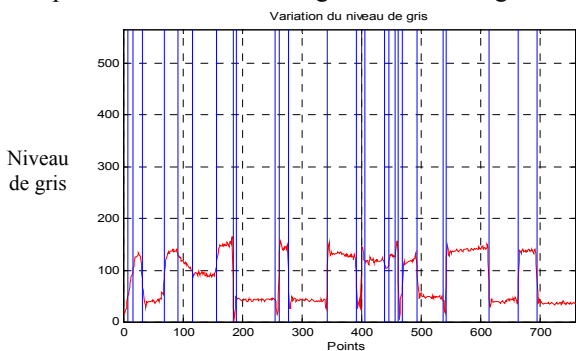


Fig.10 : Variation du niveau de gris sur la ligne 154 de l'image I_{20}

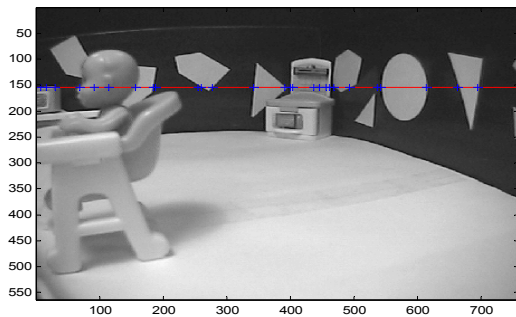


Fig. 11 : 27 points détectés sur la ligne numéro 154 de l'image I_{20}

Dans cette partie, l'appariement a été effectué en respectant les contraintes suivantes :

La contrainte épipolaire : Cette contrainte nous permet de réduire l'espace de recherche à une droite au lieu de l'image entière. La géométrie épipolaire est caractérisé par le calcul de la matrice fondamentale F qui permettant d'associer à un point de la première image (respectivement un point de la deuxième image), une droite appartenant à la deuxième image (respectivement une droite appartenant à la première image). Pour trouver le

point correspondant, il suffit de mener la recherche sur la droite épipolaire,

La contrainte d'unicité : Cette contrainte stipule qu'un point de la première image à appairier, ne peut avoir qu'un seul point homologue sur la deuxième image.

La contrainte d'ordre : L'étude de cette contrainte a été effectuée au paragraphe 7. Soient $(\mathbf{m}_1, \mathbf{m}_2)$ un couple de points de la première image I_1 , et $(\mathbf{m}'_1, \mathbf{m}'_2)$ un couple de points de la deuxième image I_2 , avec $(\mathbf{m}_1, \mathbf{m}_2) \in l$ et $(\mathbf{m}'_1, \mathbf{m}'_2) \in l'$. Si \mathbf{m}_1 correspond à \mathbf{m}'_1 et \mathbf{m}_2 correspond à \mathbf{m}'_2 , avec $u_1 < u_2$ où u_1 et u_2 abscisses des points \mathbf{m}_1 et \mathbf{m}_2 , alors nous avons : $v_1 < v_2$ où v_1 et v_2 abscisses des points \mathbf{m}'_1 et \mathbf{m}'_2 . La figure 12 illustre la contrainte d'ordre entre les images I_1 et I_2 de la séquence « Cuisine », figure 9.

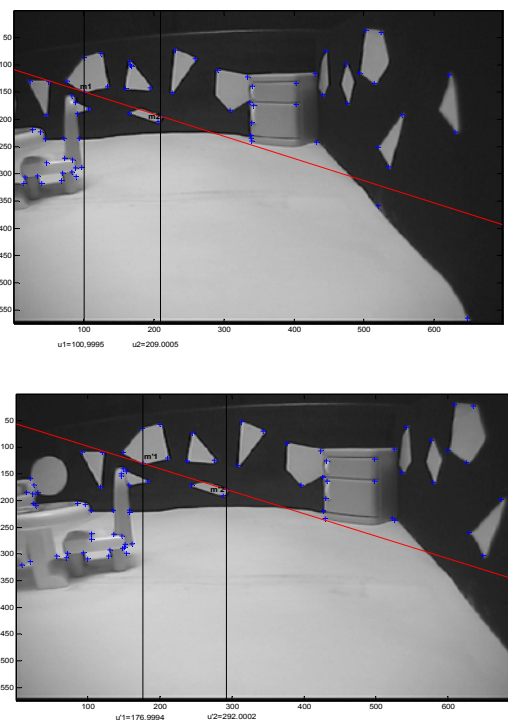


Fig. 12 : Illustration de la contrainte d'ordre

Nous avons appliqué l'algorithme d'auto-calibration ci-dessus, dans un premier temps sur 3, 4, 5, 6 images, puis sur toutes la séquence. Les résultats sont présentés sur le tableau 1. Pour chaque configuration, deux colonnes de mesures sont données. La première colonne montre les solutions de la Calibration-Affine-Euclidienne (CAE), est la deuxième colonne (σ) montre

l'estimation de l'écart-type qui représente l'incertitude de l'estimation de la première colonne. Nous remarquons que ces estimations sont très précises et très stables. Nous avons retenu dans notre étude le modèle complet de la caméra qui tient compte des distorsions radiales et tangentielles. Les incertitudes décroissent avec le nombre d'images. La valeur du rapport d'échelle de la caméra dans chacun de ces tests

$$\mu = \frac{\alpha_u \sin \Theta}{\alpha_v} \approx 1.$$

On en déduit que les pixels

sont rectangulaires donc ($\Theta = 90^\circ$), le rapport d'échelle équivaut le rapport des dimensions d'un pixel.

La dernière ligne de ce tableau indiquée par (RMS : Root of Mean Squared distances), est la racine carrée des moyennes des distances au carré entre les points détectés et les points projetés, ces mesures sont très satisfaisantes et ceci grâce à la précision de l'algorithme de détection de points d'intérêt « Harris amélioré »

Le tableau 2 présente l'estimation des paramètres extrinsèques. La matrice de rotation \mathbf{R} a été calculée grâce à l'équation (9), et le vecteur de translation (en pixel) grâce à l'équation (10).

9. Navigation

Durant la phase de l'auto-calibration nous avons calculé la matrice fondamentale puis la matrice de l'homographie à l'infini, en utilisant les méthodes présentées ci-dessus. Ces estimations nous aideront à mieux calculer les rapports de trois points alignés, par exemple : le milieu d'un couloir ou le milieu de deux obstacles.

Prenons l'exemple du milieu d'un couloir. En effet, si nous représentons les projections des deux lignes du bas du couloir S_1 et S_2 par s_1 et s_2 dans la première image et s'_1 et s'_2 dans la seconde image, figure 13, et en choisissant un point \mathbf{a} sur s_1 et un point \mathbf{b} sur s_2 , les projections dans la première image des points \mathbf{A} et \mathbf{B} appartiennent à S_1 et S_2 respectivement, qui correspondent aux points \mathbf{a}' et \mathbf{b}' dans la deuxième image. Ces points se trouvent aux intersections respectivement de la ligne épipolaire e_a de \mathbf{a} avec s'_1 et la ligne épipolaire e_b de \mathbf{b} avec s'_2 .

A partir des couples de points $(\mathbf{a}, \mathbf{a}')$ et $(\mathbf{b}, \mathbf{b}')$, nous pouvons calculer les projections $(\mathbf{m}, \mathbf{m}')$ du point milieu \mathbf{M} de \mathbf{A} et \mathbf{B} . Si S_1 et S_2 sont localement parallèles, \mathbf{M} restera toujours au milieu du couloir, et tout autre projection d'un nouveau

point sur cette ligne demeurera au milieu dans les deux images.

Nous avons utilisé la même technique de calcul, pour retrouver un point milieu de navigation de la ligne la plus large entre les obstacles, figure 14.

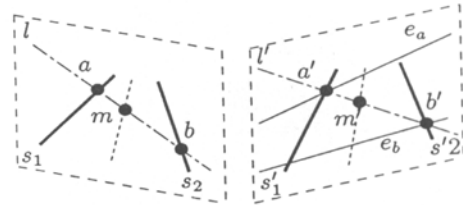


Fig. 13 : Calcul d'un point milieu dans un couloir

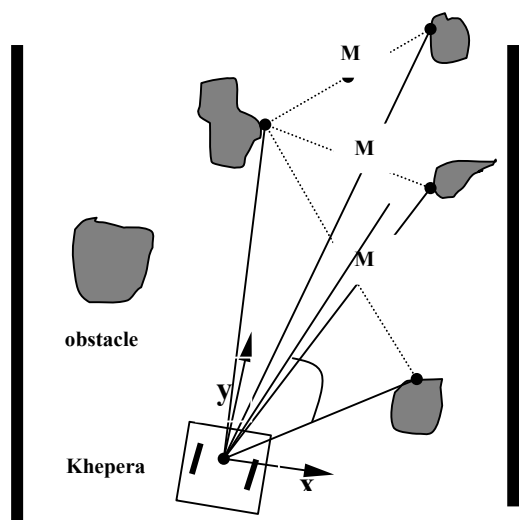


Fig. 14 : Points de navigation entre les obstacles

Pour nos simulations nous avons utilisé le robot miniature « Khepera » avec une caméra CCD embarquée. Ce mobile possède deux micro-contrôleurs pour contrôler sa position et sa vitesse à chaque mouvement. Le centre optique de la caméra correspond au centre des roues du robot.

Chaque roue est couplée à l'arbre d'un moteur courant-continu. Un encodeur d'incrémentatation est placé sur l'axe du moteur et donne 24 impulsions par tour. Ceci fournit une résolution de 600 impulsions par tour de roue et correspond à 12 impulsions par millimètre de la trajectoire du robot. Le microprocesseur principal contrôle directement l'alimentation du moteur et peut lire les impulsions de l'encodeur d'incrémentatation. Un sous programme d'interruption détecte chaque impulsion et met à jour le compteur de position de la roue.

Le robot se déplace dans une séquence d'images calibrées, où il subit plusieurs mouvements en fonction des informations visuo-spatiales reçues. Dans la « maisonnette miniature » préparée à cet effet, les séquences d'images sont prises dans des

régions ouvertes, des régions fermées, et en présence d'obstacles de différentes tailles.

Les figures 15, et 16, illustrent quelques trajectoires enregistrées en temps réelle par le générateur de trajectoire du robot. Il ne peut se référer qu'aux informations issues des images calibrées, et des estimations odométriques.

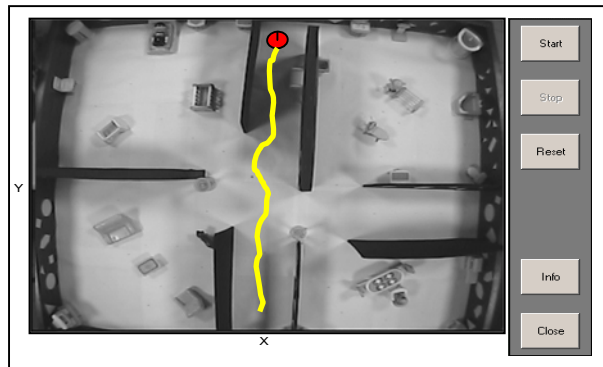


Fig. 15 : Navigation le long du couloir

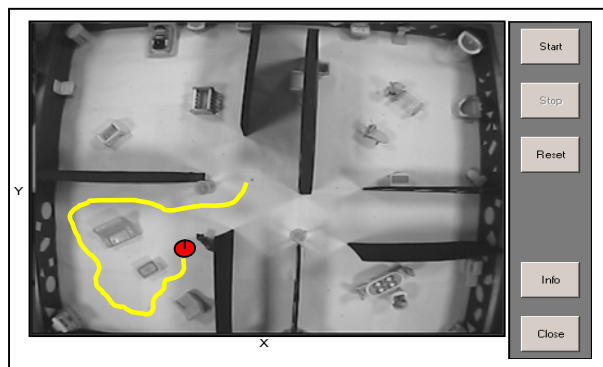


Fig. 16 : Evitement d'obstacles dans la séquence « Séjour »

10. Conclusion

La réalisation d'un fauteuil roulant robotisé pour polyhandicapé est une opération difficile, car elle met en œuvre des technologies variées (mécaniques, électroniques, informatiques,...) et des compétences multitechnologiques (robotique, intelligence artificielle, réseaux de neurones,...). L'étude présentée dans ce manuscrit a porté exclusivement sur la partie vision artificielle pour un mobile autonome robotisé.

Des pathologies neuro-visuelles affectent les différents sous-ensembles, constitutifs, des représentations spatiales chez l'enfant I.M.C. C'est à partir de la prise en compte des particularités de chaque composante de la spatialisation, que peut s'envisager une assistance visuelle efficace à la conduite d'un fauteuil roulant.

Le projet M.A.R.H. a été décrit dans le paragraphe 3. Les options technologiques retenues et les solutions à adopter pour parvenir aux spécifications

souhaitées s'inscrivent dans une conception coopérative entre l'homme et le fauteuil robotisé.

Cette coopération passe par un système de vision artificielle.

Dans le paragraphe 6, nous avons également montré qu'une implantation stable des lissages du signal avec des filtres Gaussiens permet d'améliorer de façon importante le résultat du détecteur de Harris standard. Puis une précision sous-pixelique des points d'intérêts a été réalisée à travers une approximation du maximum local à des positions non entières. Ceci nous a donné une autre forme du détecteur « Harris amélioré » qui est le plus répétable.

Les points d'intérêt mis en correspondance grâce au détecteur « Harris amélioré », nous ont permis de calculer la matrice fondamentale en utilisant les équations linéaires garantissant une convergence rapide de l'étape non linéaire utilisée à la suite.

Les tests réalisés avec des données réelles, dans le paragraphe 7 ont validé cette méthode conduisant à une stabilité acceptable de la géométrie épipolaire. Notons, qu'en l'absence du calibrage de la caméra, la matrice fondamentale joue un rôle clé en vision par ordinateur. Elle facilite la mise en correspondance et permet de calculer le mouvement entre deux positions de la caméra, ainsi qu'à estimer ses paramètres intrinsèques.

Dans le paragraphe 8, nous avons établi, comment effectuer l'auto-calibration d'une caméra CCD tout en utilisant une approche stratifiée de calibration projective et affine en estimant la matrice d'homographie du plan infini à partir de différents mouvements, et euclidienne en estimant les paramètres intrinsèques grâce à la conique absolue. Nous avons validé cette approche en utilisant des scènes d'intérieur réelles et complexes, et des mouvements planaires de la caméra embarquée. Grâce à une meilleure localisation sous-pixelique des points d'intérêt, ainsi qu'à l'efficacité de leur mise en correspondance dans une séquence d'images tout en respectant les contraintes épipolaires, d'ordre et d'unicité, cette méthode présente des résultats satisfaisants, nous permettant de localiser un obstacle par rapport à la caméra, ou mesurer la distance entre deux points de la scène.

Dans le dernier paragraphe, nous avons proposé une approche basée sur la vision pour la navigation d'un robot de table dans un environnement « miniature » d'intérieur.

Quelques simulations présentent des trajectoires enregistrées en temps réel par le générateur de trajectoire du robot Khepera se référant aux informations issues des images calibrées, et des mesures odométriques. Les résultats sont satisfaisants puisque le robot n'a pas subi de collisions.

10. Bibliographie

- [Bel. 95] **J.P. Belheur.** Domotique, Communication, Déambulation, face à la Personne Déficiante, *Rapport Technique, IRF Pomponiana Olbia de Hyères (Var)*, 1995.
- [Ben. et al. 99] **M. Ben Khelifa et al.** Un Mobile d'Autonomie Robotisé pour Polyhandicapés le M.A.R.H., *Journées de L'AFA, Toulouse, France*, 4-5 Juin 1999.
- [Der. 93] **R. Deriche.** Recursively implimenting the gaussian and its derivatives. *Rapport technique, INRIA*, 1993.
- [Flo. 93] **L. M. J. Florack.** The syntactical structure of scalare images. *Thèse de doctorat, Universiteit Utrecht*, 1993.
- [Lin. 94] **T. Lindeberg.** Scale-space Theory in Computer Vision. *Kluwer Academic Publishers*, 1994.
- [Lon. 81] **C. Longuet-Higgins.** A computer algorithm for reconstructing a scene from two projections. *Nature*, 293 : 133-135, 1981.
- [Luo 92] **Q. -T. Luong.** Matrice fondamentale et autocalibration en vision par ordinateur. *PhD thesis, Université de Paris Sud, Orsay*, Decembre 1992.
- [Luo. 94] **Q. -T. Luong et O. D. Faugeras.** An optimization framework for for efficient self-calibration and motion determination. *In Proceedings of the 12th International*
- [Mor 81] **H. Morevac.** Rover visual obstacle avoidance. Dans *Proceedings of the 7th International Joint Conference on Artificial Intelligence, pages 785-790*, 1981.
- [Pho. et al. 95] **T. Q. Phong, R. Horaud, A. Yassine, P. D. Tao.** Object Pose from 2D to 3D Point and LineCorrespondences, *International Journal of Computer Vision*, 1995.
- [Rom. 94] **B. M. ter Haar Romeny.** Geometry-Driven Diffusion in Computer Vision. *Kluwer Academic Publishers*, 1994.
- [Sch. 96] **C. Schmidt.** Appariement d'images par invariants locaux des niveaux de gris, *Thèse de Doctorat, IMAG, Grenoble*, 1996.
- [Tor. 86] **V. Torre et T. A. Poggio.** On image analysis by methods of moments. *Journal of the Optical Society of America*, 70: 920-930, 1980
- [Wit. 83] **A. P. Witkin.** Scale space filtering. *Dans proceedings of the the 8th International Joint Conference on Artificial Intelligence, pages 1019-1023*, 1983.
- [Zel. 96] **C. Zeller et O. Faugeras.** Camera self-calibration from video sequences : The Kruppa equations revisited. *Rapport de recherche 2793, INRIA*, February 1996.

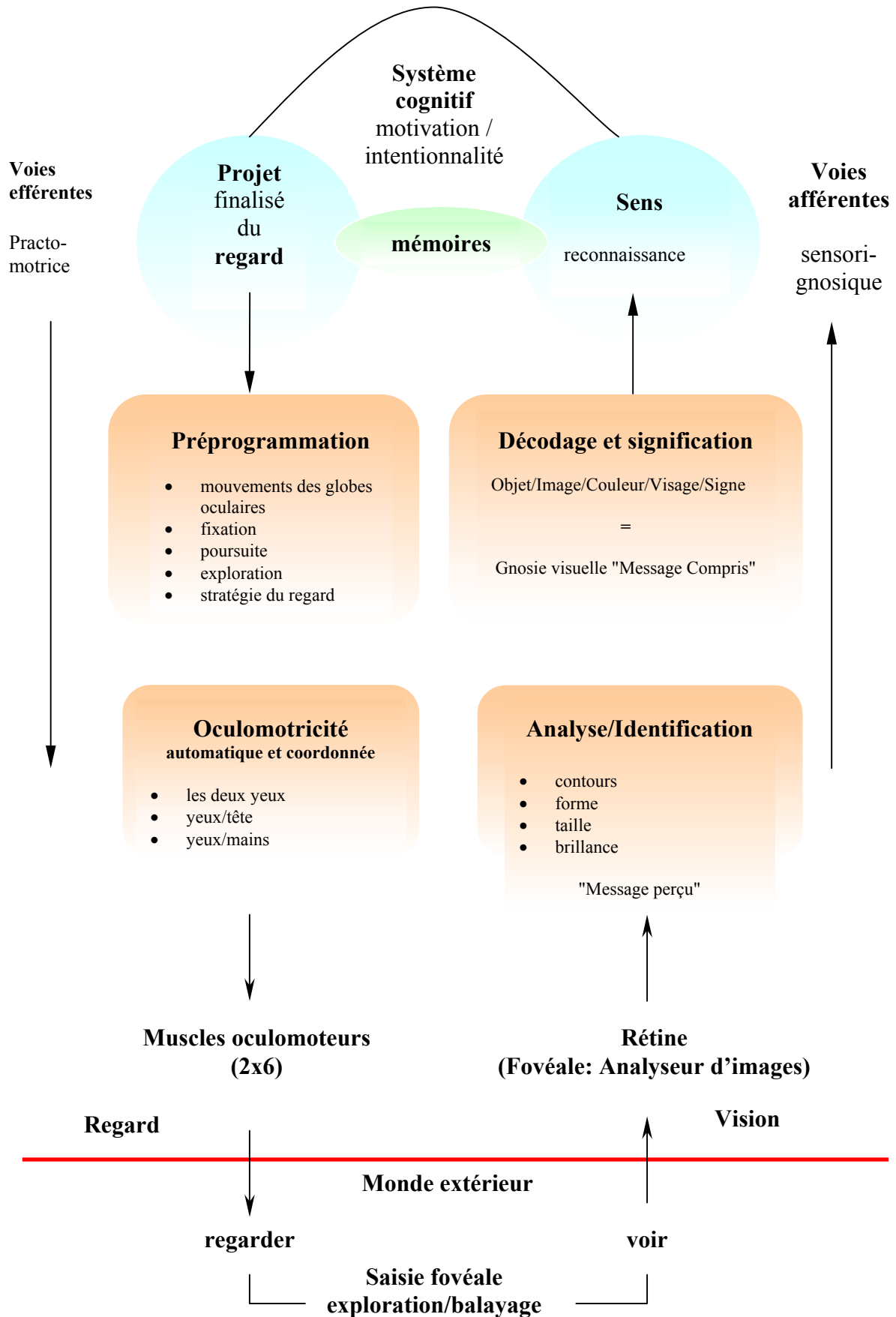


Fig. 1: Les voies visuelles : organisation fonctionnelle

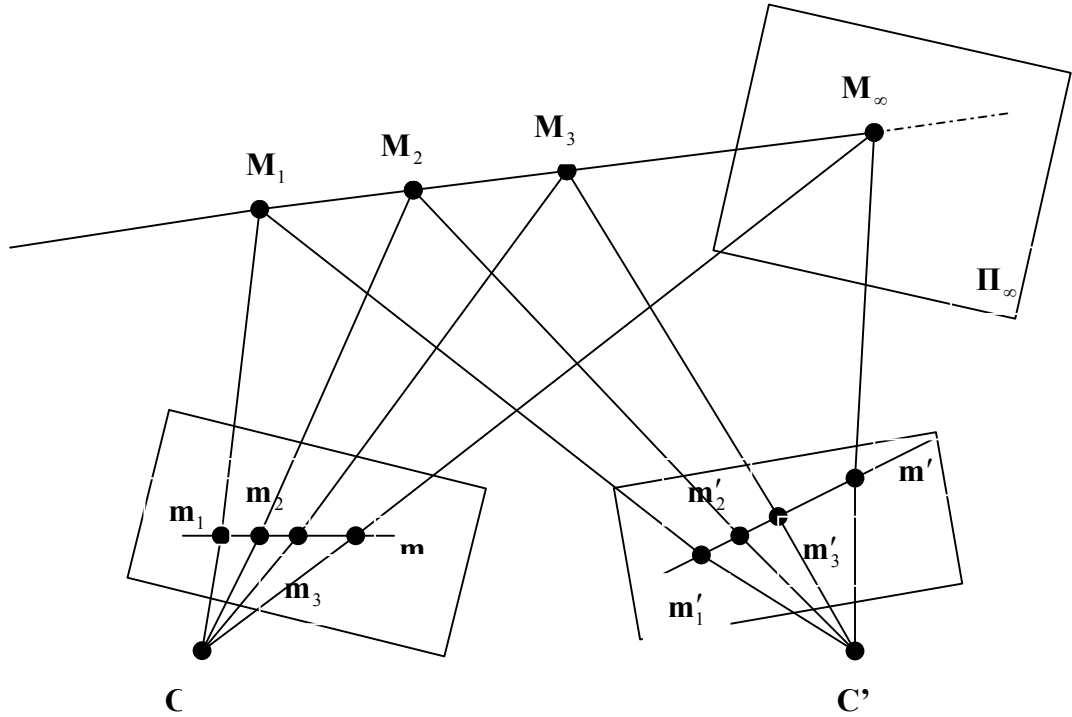


Fig. 6 : Détermination des rapports des longueurs en calibration affine

paramètres	3 images		4 images		5 images		6 images	
	CAE	σ	CAE	σ	CAE	σ	CAE	σ
α_u	528.1265	1.7430	528.8004	1.5003	529.1287	0.9888	531.2222	1.0010
α_v	527.7512	1.7890	529.5440	1.5210	528.3101	0.9774	530.9047	1.0031
u_0	379.0698	1.1101	380.2018	0.9800	378.0003	1.0247	381.0009	0.9470
v_0	281.3322	0.9026	280.4001	1.0000	278.4861	0.9089	282.7102	0.8957
k_1	-0.1270	0.0066	-0.1269	0.0061	-0.1273	0.0055	-0.1264	0.0057
k_2	0.1444	0.0332	0.1473	0.0301	0.1480	0.0204	0.1465	0.0299
J_1	-0.0055	0.0018	-0.0052	0.0021	-0.0054	0.0017	-0.0055	0.0009
J_2	0.0162	0.0011	0.0157	0.0012	0.0153	0.0010	0.0167	0.0019
RMS	0.2470	0.1020	0.2351	0.1058	0.2687	0.0980	0.2120	0.0249

Tableau 1 : Variation des résultats d'auto-calibration sur la séquence « Cuisine »

paramètres	10 images		15 images		20 images		25 images	
	CAE	σ	CAE	σ	CAE	σ	CAE	σ
α_u	530.6900	0.9102	532.5514	0.9326	531.0258	0.9411	530.6450	0.9030
α_v	529.3331	0.9509	531.9874	0.9512	531.2504	0.9000	529.1780	0.9294
u_0	380.1028	0.9058	381.9148	0.9710	382.1004	1.0001	380.8497	0.8774
v_0	281.0025	0.8024	282.3007	0.9070	281.6374	0.9090	282.0147	0.8566
k_1	-0.1252	0.0062	-0.1253	0.0058	-0.1265	0.0060	-0.1250	0.0063
k_2	0.1460	0.0323	0.1453	0.0217	0.1482	0.0158	0.1471	0.0130
J_1	-0.0050	0.0011	-0.0049	0.0007	-0.0057	0.0010	-0.0052	0.0010
J_2	0.0161	0.0016	0.0166	0.0015	0.0162	0.0023	0.0159	0.0037
RMS	0.2954	0.0201	0.3001	0.0188	0.2981	0.0271	0.3028	0.0111

Tableau 1 : Variation des résultats d'auto-calibration sur la séquence « Cuisine »

Séquences	r_x	r_y	r_z	t_x	t_y	t_z
Séq 1	-0.151176	0.178336	-0.011045	-214.9364	138.5799	1101.2453
Séq 2	-0.074716	-0.156385	0.005383	-215.7424	133.6969	1095.1061
Séq 3	-0.127951	0.123847	-0.007127	-216.3852	139.3841	1109.3270
Séq 4	-0.049708	-0.173295	0.003058	-215.3924	132.8863	1086.1321
Séq 5	-0.799335	-0.308664	0.173297	1084.9993	725.1801	3259.0296
Séq 6	-0.790040	-0.312366	0.168953	1078.9668	720.8526	3241.4882
Séq 7	0.381695	0.927837	-0.166272	1082.1936	723.4696	3251.9350
Séq 8	-0.785343	-0.301911	0.165062	1088.1453	727.3391	3268.3676

Tableau 2 : Estimation des paramètres extrinsèques

