

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Abderrahmane Mira de Béjaïa

Faculté des sciences exactes

Département informatique



Mémoire de fin de cycle

En vue d'obtention du diplôme de master professionnel en
informatique

Option : Administration et Sécurité des Réseaux

Thème :

Estimation du mouvement d'une caméra dans une
séquence d'images

Réalisé par :

M^{lle} AMAZOUZ Sonia

M^{lle} MEHDI Tinhinane

Devant le jury composé de :

Président : *M^r S. AISSANI*

Encadreur : *M^{me} S. Ait Kaci Azzou*

Examinatrice : *M^{me} S. Aloui*

Année universitaire 2015/2016

Remerciements

Au terme de ce travail, nous tenons à remercier Dieu le tout puissant de nous avoir donné le courage, la volonté et la patience pour achever ce travail.

Nous avons l'honneur et le plaisir de présenter notre profonde gratitude et nos sincères remerciements à notre chère encadreur Mme S.AIT KACCI AZZOU pour sa précieuse aide, sa confiance, ces orientations et le temps qu'elle nous a accordé pour notre encadrement.

Nos remerciements les plus sincères et les plus profonds sont adressés à Mr S.AISSANI et Mmme S.ALOUI pour avoir accepté d'examiner notre travail.

Nous remercions profondément tous les enseignants qui nous ont encouragés et soutenus.

Nous remercions également tous ceux qui ont contribué de prêt ou de loin à l'achèvement de notre Travail.

Sonia et Tinhinane

Dedicaces

Je dédie ce modeste travail :

*Du plus profond de mon cœur et avec tout le respect que je leurs dois
je dédie sincèrement ce travail :*

*Aux êtres les plus chères dans ma vie ; mes très chers parents à qui je dois tout le
bonheur du monde pour leur sacrifices et leur patience, et d'avoir été toujours là
pour moi. Sans vous, ce travail n'aurait jamais pu avoir lieu.*

*À mes frères et ma soeur pour leurs disponibilités, leur soutien moral et leur
encouragements incessants.*

À mon fiancé pour son soutien moral, ses encouragements et ses conseils.

À ma belle famille, à mes cousins et cousines

À toute ma famille de près et de loin.

À tous mes chers ami(e)s : pour leur amitié, leur soutien moral, et leur conseils.

À ma binôme Tinhinane et à toute sa famille.

Sonia

Dedicaces

À mes parent à qui je dois tout le bonheur du monde. Sans vous, ce travail n'aurait
jamais pu avoir lieu.

À mes frères et Soeurs pour leurs disponibilités, leur soutien moral et leurs
encouragements incessants.

À toute ma famille de prés et de loin.

À tous mes chères ami(e)s : pour leur amitié, leur soutien moral, et leurs conseils.

À ma binôme Sonia ainsi qu'à toute sa famille

Tinhinane

Table des matières

Table des Matières	iii
Table des figures	vii
Introduction Générale	1
1 Généralités	3
1.1 Introduction	3
1.2 Domaines d'application	3
1.3 Paradigme de David Marr	3
1.4 Vision par ordinateur [4]	4
1.5 Segmentation	5
1.6 Systèmes de vision	6
1.7 Classification des systèmes de vision par ordinateur	6
1.7.1 Les systèmes bidimensionnelles	6
1.7.2 Les systèmes tridimensionnelles	6
1.8 Stéréovision	7
1.9 Calibration	8
1.10 Mise en correspondance	8
1.10.1 Contrainte épipolaire	9
1.10.2 Contrainte d'ordre[4]	9
1.10.3 Contrainte d'unicité	10
1.10.4 Contrainte figurale	10
1.11 Reconstruction	11
1.12 Conclusion	11

2	Introduction au Mouvement des caméras	13
2.1	Introduction	13
2.2	Problèmes liés au mouvement [10]	13
2.3	Technique d'estimation du mouvement	14
2.4	Flot optique	14
2.5	Mise en correspondance	15
2.5.1	Méthodes de la mise en correspondance	16
2.5.2	Classification des méthodes de mise en correspondance	16
2.5.3	RANSAC	17
2.5.4	Sans mise en correspondance	17
2.6	Conclusion	18
3	Conception	19
3.1	Introduction	19
3.2	Modélisation géométrique d'un système de vision monoculaire en mouvement	19
3.3	Choix des primitives	20
3.4	Points d'intérêt	21
3.5	Equations de transformation 3D au 2D (Projection)	21
3.6	Relation entre les pentes d'un segment de droite dans les images successives [25]	22
3.6.1	Equations de mouvement d'un segment I	22
3.7	Détermination du mouvement	23
3.8	Appariement RANSAC	25
3.9	Conclusion	26
4	Résultat d'expérimentation	27
4.1	Introduction	27
4.2	Etapes d'expérimentation	27
4.2.1	Extraction des points d'intérêt	28
4.2.2	Appariement des segments	28
4.2.3	Regroupement des segments	29
4.2.4	Estimation du mouvement	30
4.2.5	Images résultat avec les points en correspondance	31
4.3	Résultats d'expérimentation sur une autre image	32
4.4	Extraction des points d'intérêt	32

4.4.1	Estimation du mouvement	33
4.4.2	Images résultat avec les points en correspondance	34
4.5	Conclusion :	35
	Conclusion générale et perspectives	35
	Bibliographie	32

Table des figures

1.1	Le paradigme de David Marr.	4
1.2	vision par ordinateur	5
1.3	Système de stéréovision.	7
1.4	mise en correspondance entre deux images.	8
1.5	géométrie épipolaire.	9
1.6	contrainte d'ordre.	9
1.7	contrainte d'unicité.	10
1.8	contrainte de la continuité figural.	10
1.9	Reconstruction par La triangulation.	11
2.1	Exemple de flot optique.	15
3.1	Modélisation géométrique de la tête de vision stéréoscopique.	19
3.2	détection de point d'intérêt.	21
4.1	Images de test.	27
4.2	Extraction des points d'intérêts.	28
4.3	les segments.	28
4.4	Regroupement des segments.	29
4.5	Estimation du mouvement.	30
4.6	Images résultat.	31
4.7	Images de test.	32
4.8	Extraction des points d'intérêt.	32
4.9	Estimation du mouvement.	33
4.10	Images résultat.	34

Introduction generale

Contexte :

Le monde a une structure 3D et il est composé d'objets, d'êtres humains qui savent parfaitement décrire et interpréter ce monde. Cependant, l'information disponible sur la rétine n'est qu'un ensemble de points (environ un million de Picture éléments ou pixels)... Les objets (téléphone, voiture...) n'existent pas sur la rétine, et pourtant on les voit : leur interprétation est le résultat du processus visuel.

La vision humaine est extrêmement complexe (neurosciences). La vision par ordinateur ne cherche pas à comprendre ou à reproduire la vision humaine, mais à construire un modèle algorithmique et théorique par lequel l'information utile relative à l'environnement peut être automatiquement extraite et analysée à partir d'une image, d'un ensemble d'images ou d'une séquence d'images.

L'acquisition d'images est réalisée par l'intermédiaire d'une caméra vidéo dont le signal est numérisé. Le positionnement des objets dans la scène nécessite la connaissance de la position et l'orientation de la caméra par rapport à un repère lié au mouvement. Le problème de la localisation de la caméra est important ; divers travaux sont menés dans ce domaine pour le résoudre.

Motivation de l'estimation d'un mouvement de caméra :

La connaissance du mouvement d'une caméra a de nombreuses applications. Elle est à la base des méthodes appelées " structure from motion ", méthodes estimant un plan des profondeurs de la scène filmée (ou structure) à partir des images du film et du mouvement de la caméra. Elle est utilisée pour la compensation du mouvement entre deux images, opération très efficace en compression pour diminuer le coût de

codage d'une vidéo [1], ou pour stabiliser une séquence d'images.

Un autre champ d'application important est la réalité augmentée. L'idée est d'ajouter à des images d'un monde réel des objets virtuels [2]. La connaissance du mouvement de la caméra est nécessaire pour insérer les objets virtuels dans la séquence avec le même point de vue que celui adopté par la caméra lors du tournage. En urbanisme, la réalité augmentée permet d'examiner des projets de construction en insérant un futur bâtiment dans une vidéo tournée sur le site d'implantation et Les domaines d'application sont nombreux : cinéma, architecture d'intérieur...etc

Problématique : Les travaux de recherche diffèrent selon le domaine d'application. On retrouve des travaux qui utilisent le mouvement pour retrouver la structure 3D. D'autres essayent de déterminer la position de la caméra ou de l'objet dans la scène. Dans ce dernier cas, la majorité des travaux détermine le mouvement en utilisant des appariements de primitives.

Dans notre cas, nous essayerons de proposer une méthode d'estimation du mouvement sans connaissance au préalable des appariements de primitives ni des modèles géométriques de la caméra utilisée.

Organisation du rapport :

Pour cela notre mémoire est organisé de la manière suivante :

Chapitre 1 donne un aperçu sur la vision par ordinateur.

Chapitre 2 est dédié aux techniques d'estimation du mouvement de la caméra.

Chapitre 3 est consacré à la description de la méthode développée pour l'estimation du mouvement.

Chapitre 4 présentera les résultats d'expérimentation obtenus en testant notre méthode sur des données réelles.

Nous terminerons par une conclusion résumant notre travail, et mettant en avant les perspectives futures.

Chapitre 1

Généralités

1.1 Introduction

La vision par ordinateur est une discipline qui tente de simuler la vision humaine en établissant des modèles qui possèdent des propriétés proches de la perception visuelle humaine. Dans ce chapitre, on va donner quelques définitions concernant la vision par ordinateur ainsi les systèmes de vision

1.2 Domaines d'application

Les applications de la vision par ordinateur sont nombreuses et intéressent pratiquement tous les domaines tels que l'industrie(Vision robotique, Inspection et mesures automatiques, Vision robotique...), la médecine(Cytologie, Tomographie, Echographie...), le militaire(Surveillance, Guidage automatique et poursuite d'engins...),la sécurité, etc.

1.3 Paradigme de David Marr

Vers la fin des années 70 David Marr [3] a proposé un modèle calculatoire pour le traitement et la représentation de l'information visuelle. Voici les principaux traits de ce paradigme :

- **La segmentation** : À partir d'une ou de plusieurs images un processus d'extraction de caractéristiques produit une description en matière d'attributs bidimensionnels ; ce niveau de représentation est appelé première ébauche " primal sketch ".

- **La reconstruction** : la première ébauche constitue l'entrée d'un certain nombre de processus plus ou moins indépendants qui calcule des propriétés tridimensionnelles locales relatives à la scène ; il s'agit d'une représentation centrée sur l'observateur, appelée ébauche 2.5D ; ces processus opèrent sur une séquence d'images (analyse du mouvement), sur une paire d'images (stéréoscopie) ou sur une seule image. Dans ce dernier cas, il s'agit de processus d'inférence qui utilise des connaissances géométriques (analyse des contours) et statistiques (analyse des textures) photométriques (analyse des ombrages) ou colorimétriques (analyse des reflets).

- **La reconnaissance** : L'ébauche 2.5D est la mise en correspondance avec des connaissances 3D afin de construire une description de la scène en matière d'objets et de relations entre les objets, il s'agit d'une représentation centrée sur la scène (la description ne dépend plus de la position de l'observateur).

En pratique, le paradigme de David Marr se traduit par trois étapes de traitement : segmentation, reconstruction et reconnaissance.

On peut résumer le paradigme de David Marr selon la figure 1.1

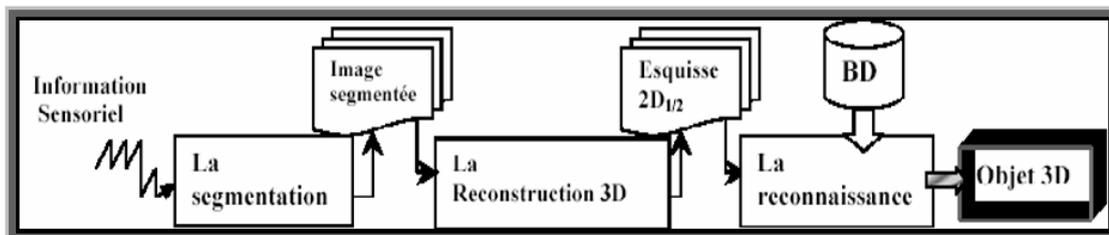


FIG. 1.1 – Le paradigme de David Marr.

1.4 Vision par ordinateur [4]

La vision par ordinateur désigne l'ensemble des méthodes et techniques qui permettent de doter l'ordinateur de facultés visuelles et de compréhension de scènes. Elles incluent des méthodes d'acquisition de traitement et d'analyse d'images et font appel de différentes théories telles que le traitement du signal, les statistiques, la physique, l'optique, l'intelligence artificielle, etc

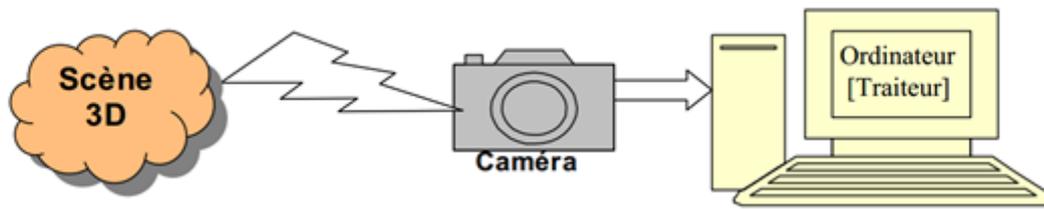


FIG. 1.2 – vision par ordinateur

- * **Un capteur vision** : pour la perception du monde extérieur par les organes de la vue (ex : caméra).
- * **Traitement** : traitement + analyse + interprétation.
 - Traitement (prétraitement) opérations de manipulation de l'image pour améliorer la qualité.
 - o la compression : réduction du volume de l'image.
 - o la restauration : correction des défauts dus à une source de dégradation.
 - o l'amélioration : modification de l'image dans le but de la rendre plus agréable à la caméra.
 - Analyse : suite d'opérations pour l'extraction d'informations contenues dans une image.
 - o phase de segmentation avec partition de l'image.
 - o techniques de description / modélisation pour obtenir la description structurelle de l'image.
 - Interprétation : passage de la description structurelle à la description sémantique en gardant certains objectifs (mesure de paramètres sur des formes, description du contenu de la scène en matière de concepts non mathématiques).

1.5 Segmentation

La segmentation [5], en général est le découpage spatial de l'image en zones homogènes. Elle joue un rôle prépondérant dans le traitement et l'analyse d'image et dans la vision par ordinateur.

Elle consiste en la détection des contours ou de région, la détection des contours est une étape préliminaire à de nombreuses applications de l'analyse d'images. Les contours constituent en effet des indices riches, au même titre que les points d'intérêts, pour toute interprétation ultérieure de l'image. ces contours dans une

image proviennent des discontinuités de la fonction de réflectance (texture, ombre), et des discontinuités de profondeur (bords de l'objet), et sont caractérisés par des discontinuités de la fonction d'intensité dans les images. Le principe de la détection de contours repose donc, sur l'étude des dérivées de la fonction d'intensité dans l'image : les extrema locaux du gradient de la fonction d'intensité et les passages par zéro du Laplacien. La difficulté réside dans la présence de bruits dans les images.

1.6 Systèmes de vision

Un système informatique de vision [6] est un ensemble de processus fortement liés et composé de matériels et de logiciels qui reçoit en entrée une séquence d'images qu'il traite en utilisant la connaissance spécifique aux différents niveaux de traitements, en sortie, il fournira une description en matière d'objets et de relations entre ces derniers.

1.7 Classification des systèmes de vision par ordinateur

On distingue deux types de systèmes [4] qui sont :

1.7.1 Les systèmes bidimensionnelles

Ces systèmes permettant une acquisition de l'image de la scène 2D et n'incluent pas la notion de profondeur, ils ne peuvent reconnaître que les objets plans, on les exploite dans la médecine (radiographique), télédétection et la reconnaissance de caractères, la robotique industrielle...

1.7.2 Les systèmes tridimensionnelles

Ces systèmes sont capables d'identifier les objets à partir d'une image d'un monde tridimensionnel et ils incluent la notion de profondeur, il existe deux types de système tridimensionnel.

Il existe deux types de vision tridimensionnelle :

1.7.2.1 Systèmes actifs [4]

Ils contrôlent l'éclairage de la scène en projetant une énergie lumineuse sur un ensemble de points de la scène afin de la sur éclairer, parmi ces systèmes :

- Lumière structurée
- Temps de vol
- Stéréo active

1.7.2.2 Systèmes passifs

le rôle des méthodes passives[4] est simplement de recevoir les signaux lumineux émis ou réfléchis par la scène. Les capteurs utilisés sont principalement les caméras. L'information directement acquise n'est pas tridimensionnelle, mais seulement bi-dimensionnelle. Une ou plusieurs images sont traitées afin d'extraire l'information tridimensionnelle souhaitée.

1.8 Stéréovision

La stéréovision[6] vise à réaliser une tâche similaire à la vision par humaine, à l'aide d'un ou plusieurs périphériques de capture d'image, reliés à un ordinateur. On utilise deux caméras, vidéo ou CCD, disposées comme les yeux humains, qui vont donner deux images d'une scène (figure 1.3). Ces images sont formées de pixels et constituent une paire stéréoscopique. On considère que l'aspect géométrique de l'espace qui nous entoure est une représentation tridimensionnelle d'entités physiques. Une image prise par une seule caméra est considérée comme une représentation bidimensionnelle de cet espace. Il y a perte d'informations durant le processus de formation d'image et En particulier la troisième dimension. La récupération de cette dernière est le but de la stéréovision.

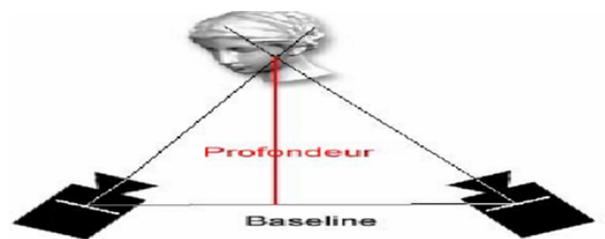


FIG. 1.3 – Système de stéréovision.

1.9 Calibration

Le modèle sténopé de la caméra permet d'exprimer le passage des points 3D de la scène aux points 2D de l'image, le calibrage permet de retrouver les paramètres intrinsèques (distance focale de la caméra, sa position) et les paramètres extrinsèques (l'angle de rotation, mouvement), quoique souvent, seule la géométrie interne est recherchée (e.g. pour une caméra qui se déplace au cours d'une application)[7].

1.10 Mise en correspondance

La mise en correspondance [8] [9] consiste à déterminer les couples de points dont les caractéristiques communes. Les caractéristiques de chaque point peuvent être calculées en tenant compte de la couleur ou de la texture au voisinage des points, une primitive peut être de points (pixel, points d'intérêts.etc.), des segments (segment de droite, arcs de cercle,etc...) ou des régions.

Le problème principal de la mise en correspondance se trouve principalement dans la mise en correspondance elle-même. En effet, l'être humain n'éprouve aucune difficulté à appairer des points, en revanche cette correspondance peut s'avérer relativement difficile à réaliser d'un point de vue purement algorithmique. Des contraintes supplémentaires sont nécessaires pour pallier à ce problème.

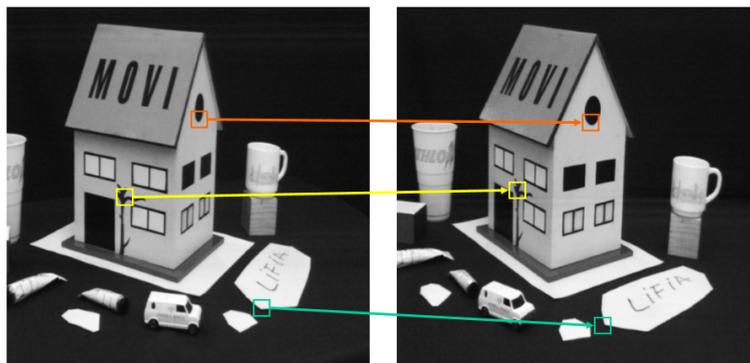


FIG. 1.4 – mise en correspondance entre deux images.

1.10.1 Contrainte épipolaire

La contrainte épipolaire est une contrainte géométrique qui réduit l'ensemble des correspondants d'un point à une droite dans l'image. La figure (1.5) résume la géométrie épipolaire.

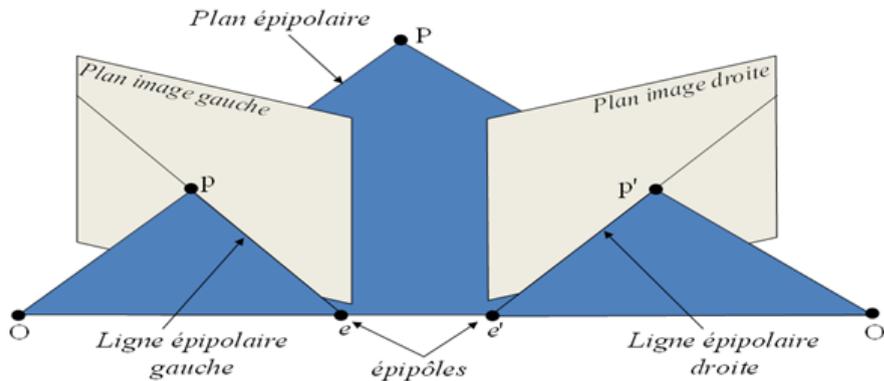


FIG. 1.5 – géométrie épipolaire.

La contrainte épipolaire permet d'obtenir une relation entre les points de l'image de gauche et ceux de l'image de droite qui ne dépendent que des paramètres des caméras et des coordonnées image et qui sont indépendantes des coordonnées du point de la scène.

1.10.2 Contrainte d'ordre[4]

Les contraintes épipolaires et de disparités permettent de réduire le nombre d'appariements possibles entre les primitives de l'image de gauche et celles de l'image de droite, la contrainte d'ordre implique que la projection des objets d'une scène conserve le même ordre dans les deux projections images.(figure 1.6).

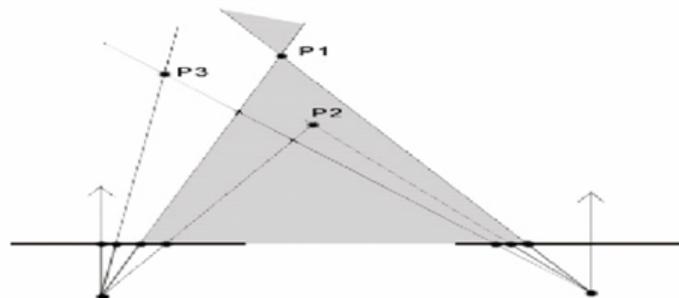


FIG. 1.6 – contrainte d'ordre.

1.10.3 Contrainte d'unicité

La contrainte d'unicité est vérifiée si tout point de l'image gauche possède au plus un correspondant dans l'image de droite. La contrainte d'unicité découle directement de la contrainte d'ordre dans le sens où elle décrit un cas limité de cette dernière. Lorsque deux points P_1 et P_2 se projettent en un même point P dans l'image de gauche et en deux points différents p'_1 , p'_2 dans l'image de droite. (Figure 1.7)

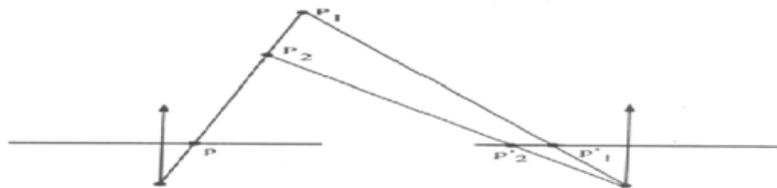


FIG. 1.7 – contrainte d'unicité.

1.10.4 Contrainte figurale

Les contraintes figurales sont définies comme telles car elles n'ont pas à l'instar des contraintes géométriques de bases théoriques.

Disparité locale constante : Lorsqu' une scène comporte des objets proéminents dont la surface est fortement inclinée par rapport aux plans des deux images, il y a un risque que la contrainte d'ordre ne soit pas respectée. Pour éviter cela on peut imposer une limite au gradient de disparité entre deux points appariés consécutifs.

Continuité figurale : Cette contrainte limite les variations de disparité le long des contours (Figure 1.8)

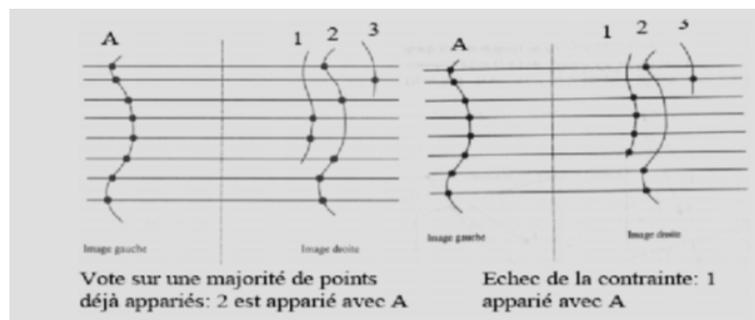


FIG. 1.8 – contrainte de la continuité figurale.

1.11 Reconstruction

La reconstruction 3D est une interprétation géométrique particulière de ce qui est observable dans un espace image caractérisé par sa géométrie épipolaire [12]. Le capteur stéréoscopique constitué de deux caméras fournit des paires d'images stéréo. La reconstruction 3D à partir d'une paire d'images suppose que l'on dispose des correspondances entre ces deux images. (Figure 1.9).

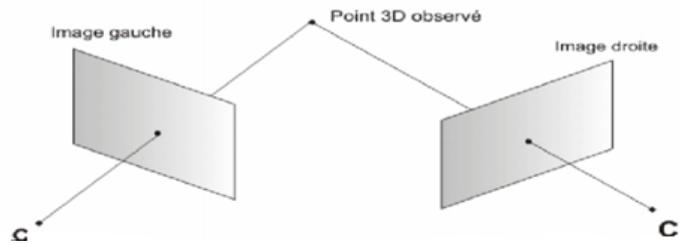


FIG. 1.9 – Reconstruction par La triangulation.

La reconstruction dépend du type de calibrage dont on dispose. On peut distinguer les cas suivants :

- Le capteur stéréoscopique est calibré et on dispose de paramètres intrinsèques de chaque caméras ainsi que de la transformation rigide entre les deux caméra (paramètre extrinsèques) : dans ce cas, on obtient une reconstruction euclidienne dans le repère de calibrage.
- Les paramètres internes de chaque caméra sont connus mais la transformation rigide entre les deux caméras est inconnue. Dans ce cas, on obtient une reconstruction euclidienne dans le repère de l'une ou l'autre des deux caméras.
- Si aucun calibrage n'est disponible, il faut estimer la matrice fondamentale (calibrage). A partir de la matrice fondamentale, on peut obtenir une reconstruction projective tridimensionnelle.

1.12 Conclusion

Nous avons vu dans ce chapitre, un aperçu sur les principes de bases de la vision par ordinateur, l'un des objectifs de cette dernière, est une reconstruction de la structure tridimensionnelle (3D) de l'espace à partir des images acquises par les capteurs (un ou plusieurs).

Pour pouvoir reconstruire, il faut deux images (systèmes stéréoscopiques) c'est pourquoi au lieu d'utiliser deux caméras, nous allons utiliser une caméra en mouvement pour simuler la stéréoscopie, le chapitre 2 est dédié au mouvement de la caméra.

Introduction au Mouvement des caméras

2.1 Introduction

L'analyse du mouvement occupe une place essentielle en vision par ordinateur. Les zones de mouvement dans une séquence d'images correspondent généralement à des événements sur lesquels on souhaite focaliser l'attention d'un système de vision. Des méthodes de détection fiables et automatiques sont indispensables pour de nombreuses applications telles que la surveillance, la compression ou l'indexation vidéo. Globalement, il existe un besoin en méthodes simples et robustes en vision par ordinateur.

La détection de mouvement est une tâche préliminaire à beaucoup de traitements de plus haut niveau en vision par ordinateur tels que la compression ou le suivi.

2.2 Problèmes liés au mouvement [10]

Toutes les méthodes de détection du mouvement ont deux principales contraintes : l'éclairage de la scène et le mouvement du capteur. En effet, la situation est différente selon que le capteur est fixe, qu'il bouge lentement ou qu'il soit mobile avec un mouvement compensé. De même, un éclairage constant, des variations d'éclairage à basse fréquence ou des changements soudains d'intensité lumineuse sera géré différemment par un système de détection du mouvement. A ces difficultés s'ajoutent le bruit du capteur (bruit d'acquisition et de numérisation) et la gestion des zones homogènes (lorsque la différence de luminance entre deux instants est inférieure à un certain seuil).

l'estimation du mouvement est utilisé :

Détection du mouvement :

- Est-ce que quelque chose a bougé dans la scène ?

Estimation du mouvement :

- Localisation, vitesse et direction.

Suivi du mouvement (tracking) :

- Correspondance entre plusieurs images pour un même mouvement.

Reconnaissance du mouvement :

- Reconnaître le scénario correspondant au mouvement.

Estimation de la structure à partir du mouvement :

- Structure from motion.

Interpolation de vues :

- Synthèse de nouvelle image dans la séquence.

2.3 Technique d'estimation du mouvement

On a deux principales grandes familles de techniques [11] d'estimation du mouvement qui sont :

- les méthodes différentielles, basées sur le calcul de la vitesse locale apparente (flot optique).
- les techniques de mise en correspondance basées sur l'extraction et l'appariement de primitives telles que les ponts d'intérêts, segments...

2.4 Flot optique

Le terme de flot optique [12] a été inventé par le psychologue James Jérôme Gibson en 1950 ; dans une étude sur la vision humaine. Le flot optique entre deux images successives est le mouvement apparent des pixels d'une image à l'autre. L'utilisation de cette méthode consiste à extraire l'information de la vitesse à partir d'une séquence d'images avec l'hypothèse que l'intensité des pixels des objets est conservée au cours du déplacement. Sous cette hypothèse, on peut établir une relation entre la vitesse apparente v (déplacement dans l'image d'indices visuels tels que des régions délimitées par des contours supposés représenter la projection du mouvement

3D des objets de la scène et /ou du mouvement de la caméra) et les variations spatio-temporelles de l'intensité. En notant [13] $I(x, y, t)$ l'intensité d'un point en position (x, y) de l'image à l'instant t et (U_x, U_y) la vitesse de déplacement de ce point en cet instant, on écrit que la dérivée par rapport au temps du champ d'intensité est nulle, ce qui amène à l'équation fondamentale (Horn et Schunck 1981).

$$\frac{\sigma I}{\sigma x} \cdot U_x + \frac{\sigma I}{\sigma y} \cdot U_y + \frac{\sigma I}{\sigma t} = 0$$

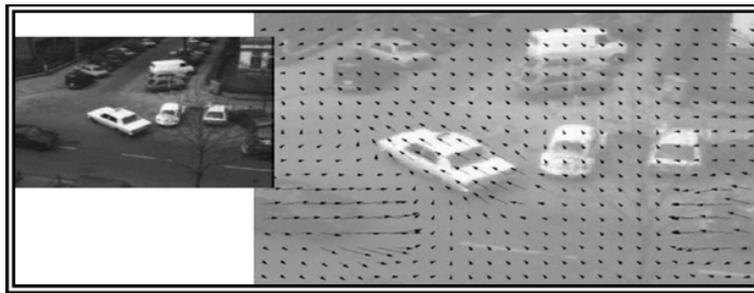


FIG. 2.1 – Exemple de flot optique.

Avantages :

- Engendre des informations riches.
- Elle est robuste à l'ajout de bruit gaussien sur les composantes du flot.

Inconvénients :

- Assez coûteux en temps de calcul.
- Sensible à la fiabilité des résultats.
- Sensible à l'ajout de bruit impulsionnel.
- L'estimation du flot optique sur l'image entière présente trop de valeurs aberrantes pour que la méthode fournisse de bons résultats.

2.5 Mise en correspondance

La mise en correspondance de points entre deux images est le plus souvent basée sur le principe suivant : un point de la première image représente le même point physique qu'un autre point de la seconde image (on dit aussi que les deux points se correspondent) si les deux points se " ressemblent ". Cette ressemblance doit prendre en compte les voisinages des points à cause du bruit dans les images. Les méthodes de mise en correspondance réalisent l'appariement entre caractéristiques extraites

de différentes images [14]; C'est le cœur du processus de mise en correspondance d'images. Pour cela, ces méthodes doivent être efficaces et robustes indépendamment de l'invariance des caractéristiques.

2.5.1 Méthodes de la mise en correspondance

Les méthodes de mise en correspondance cherchent à trouver des informations similaires dans deux images espacées temporellement (ou spatialement dans le cas de la stéréovision). Les informations utilisées peuvent être : des primitives (points, lignes, cercles, ...), des modèles (déformables ou articulés par exemple) extraits de l'image, ou des informations statistiques sur les niveaux d'intensité des pixels (histogramme, entropie ...). Au-delà de la simple mise en correspondance, l'objectif final de ces méthodes est le calcul de paramètres d'un modèle de mouvement paramétrique ou le suivi d'objet. Les méthodes utilisées sont généralement des méthodes d'optimisation ou d'ajustement des paramètres par rapport à une erreur mesurée (asservissement).

2.5.2 Classification des méthodes de mise en correspondance

La mise en correspondance d'images est l'opération qui réalise l'appariement spatial entre les éléments d'images basé sur leur intensité. Les méthodes peuvent être divisées en trois principales approches selon l'acquisition[15] :

1. Les images à mettre en correspondance sont prises par un (ou plusieurs) capteur(s) identiques avec des orientations différentes et la stéréovision est un exemple.
2. Dans la deuxième approche, il s'agit de mettre en correspondance des images prises par le capteur pendant une période de temps ; un système de surveillance ou l'analyse de mouvement sont de cette catégorie.
3. La dernière approche couvre l'analyse multi modale d'images prises par de différents types de capteurs. Ainsi, les formats de l'information extraite sont différents.

Avantages[12] :

- minimisation directe de l'erreur de reconstruction.
- fiabilité.
- simplicité du principe, et parallélisation possible.
- méthode connue, maîtrisée, optimisée...

Inconvénients :

- complexité pour un espace de recherche important.
- pauvreté de la modélisation pour une complexité raisonnable.

2.5.3 RANSAC

l'algorithme RANSAC ("Random Sample Consensus") qui est proposé par Fischler et Bolles en 1981 [16] est le plus utilisé peut éliminer les données aberrantes. Des travaux récents confirment son actualité [17-18] de même que ses évolutions depuis 2000 [19-20]. Torr et Zisserman [19] ont amélioré cet algorithme en choisissant la solution qui maximise la vraisemblance au lieu de considérer simplement le nombre de points ne correspondant pas au modèle.

Son principe est le suivant :

- un sous-ensemble de données est construit de manière aléatoire ;
- le modèle correspondant à ces données est estimé,
- et le nombre de points qui ne suivent pas ce modèle est évalué.

Cette procédure est répétée en nombre fixé de fois. Le modèle donnant le moins de points aberrants (c'est-à-dire ne suivant pas le modèle) est sélectionné.

L'avantage de RANSAC est sa capacité à calculer de manière robuste les paramètres du modèle, c'est-à-dire qu'il peut estimer les paramètres avec un degré élevé de précision, même si une quantité importante de valeurs aberrantes (outliers) est présente dans les données.

L'inconvénient de RANSAC est qu'il n'y a pas de limite supérieure sur le temps de calcul de ces paramètres. Lorsqu'une limite est utilisée (un nombre maximal d'itérations), la solution obtenue peut ne pas être la solution optimale. L'autre inconvénient de RANSAC est qu'il suppose de fixer des seuils spécifiques au problème traité.

2.5.4 Sans mise en correspondance

La recherche du mouvement sans appariement photométrique a été proposée par Dellaert et al. [21]. Néanmoins, il s'agit d'une approche purement combinatoire du problème : il est explicitement supposé que le même nombre de points détecté dans les images sont correctement détectés (i.e. qu'ils correspondent tous à un point 3D). Le problème est résolu par maximisation de la vraisemblance géométrique des appariements possibles, dans une méthode de recuit simulé permettant d'éviter les

minima locaux.

Si cette approche est intéressante pour contenir la combinatoire importante du problème, les hypothèses semblent trop contraignantes pour qu'elles puissent être utilisées dans un cadre réaliste.

Domke et Aloimonos [22] définissent la vraisemblance des mouvements de caméra en s'appuyant sur un modèle probabiliste mêlant ressemblance photométrique et cohérence géométrique (distance aux lignes épipolaires). Aucun appariement a priori n'est nécessaire. Néanmoins, la maximisation de la vraisemblance est comme dans [21] un problème combinatoire très difficile (même dans leur contexte où les caméras sont calibrées au préalable, réduisant la représentation du mouvement à la matrice essentielle).

La solution présentée par les auteurs permet de trouver une solution lorsque l'on dispose d'une prédiction du mouvement attendu, mais qui est beaucoup trop coûteuse dans un autre cadre.

2.6 Conclusion

Dans ce chapitre nous avons présenté des méthodes pour l'analyse et l'estimation du mouvement qui se résume en deux méthodes :

- La première est basée sur le calcul de flux optique : cette technique nécessite une estimation des dérivées première et seconde de la fonction de l'intensité, qui est très sensible au bruit.
- La seconde est basée sur l'appariement d'un certain nombre de primitives : quel que soit le système monoculaire ou stéréoscopique, nous pouvons constater que la correspondance des primitives pourra être calculée à partir de l'algorithme de RANSAC.

Dans le prochain chapitre, nous proposons une méthode qui permet l'estimation du mouvement de la caméra sans appariement de primitive au préalable.

Chapitre 3

Conception

3.1 Introduction

Un système de vision stéréoscopique utilise des caméras mobiles effectuant des mouvements de convergence, d'inclinaison et de translation, pour produire un ensemble d'images prises à des endroits différents et sous des angles de vue différents, ce qui permet d'obtenir une carte de profondeur plus complète de l'environnement. Dans ce chapitre nous allons présenter la démarche à suivre pour l'estimation de mouvement monoculaire, après avoir fait un choix de primitive et un algorithme.

3.2 Modélisation géométrique d'un système de vision monoculaire en mouvement

La figure (Fig.3.1) suivante illustre le model de système de vision stéréoscopique équipé de la caméra en mouvement de rotation. Où :

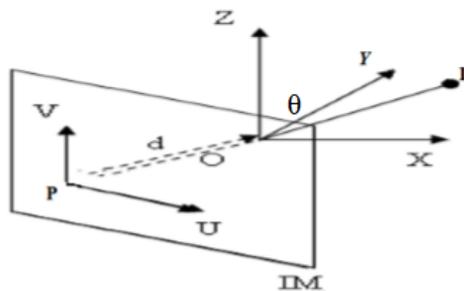


FIG. 3.1 – Modélisation géométrique de la tête de vision stéréoscopique.

- IM est le plan de projection de la caméra,
- L : l'entille de la caméra (centres optiques),
- P : point d'impact de l'axe optique sur le plan IM
- (P U V) : repère cartésien orthonormé dit théorique lié au plan IM, construit ainsi :
 - P U est parallèle aux lignes de l'image,
 - P V est parallèle aux colonnes de l'image IM (sur une caméra CCD les lignes sont perpendiculaires aux colonnes),
 - l'unité, que ce soit sur O U est soit le pixel ou le millimètre (mm). Le passage d'une unité à l'autre (pixel -mm par exemple) obéit à la transformation suivante : $U=(ex).u$, $V= (ez).v$ Sachant que (ex, ez) représentent les facteurs de changement d'échelle et (U ,V) (resp. (u, v)) Les coordonnées d'un point image dans le système où l'unité est le pixel (resp. le mm)
- (O x y z) : repère cartésien orthonormé -dit théorique- lié à la caméra construit de la manière suivante :
 - O est situé sur l'axe optique de la caméra à une distance d de P.
 - O Z est parallèle à (P V),
 - l'unité utilisée sur ce repère est le mm,
 - O x sont colinéaires.
 - θ :angle de rotation de la caméra autour de l'axe Oz
 - f : distance focale.

3.3 Choix des primitives

Le choix s'est porté sur l'appariement des segments de droite car :

- Les segments fournissent une représentation simplifiée de l'image. Aussi, ils répondent à certains critères de choix tels que : la robustesse, la discriminance, la précision et la densité [23].
- Les segments de contours réduisent la complexité combinatoire : par construction, il y a moins de segments que de points, donc, moins d'appariements potentiels.

- Le monde est fait de surfaces continues qu'on peut approximer par des segments.
- Appariement des segments, c'est intégrer la contrainte de continuité des formes.

3.4 Points d'intérêt

Les points d'intérêt[24] sont également des primitives pertinentes que l'on peut extraire des images ; ils sont largement utilisés pour l'appariement d'image, Nous avons opté pour l'utilisation des points d'intérêt comme extereement de segment. La méthode de Harris qui est largement utilisée en matière d'extraction de points d'intérêt. L'idée de base de ce détecteur est d'utiliser la fonction d'auto-corrélation pour déterminer les positions où le signal change dans deux directions simultanément. on construit une matrice M liée à cette fonction d'auto-corrélation en chaque point de l'image.



FIG. 3.2 – détection de point d'intérêt.

3.5 Equations de transformation 3D au 2D (Projection)

En utilisant le modele géométrique de la caméra, tout point $M(x,y,z)$ de l'espace tridimensionnel se projette en un point $M(u,v)$ sur IM de la manière suivante :[25]

$$U = (ex) \cdot f \cdot \frac{-\cos(\theta) \cdot x + \sin(\theta) \cdot y}{\sin(\theta) \cdot x + \cos(\theta) \cdot y + d - f} \dots\dots\dots 3.1$$

$$V = (ez) \cdot f \cdot \frac{-Z}{\sin(\theta) \cdot x + \cos(\theta) \cdot y + d - f} \dots\dots\dots 3.2$$

Après un certain nombre de rotations des caméras, les coordonnées de projection du point objet $M(x,y,z)$ sur l'images IM, Que deviennent alors ces coordonnées après k rotations d'angle α autour de O Z pour la caméra ? Supposons que la caméra effectue des rotations d'un angle (α) autour de O Z à partir de sa position initiale θ . Après cette rotation, les coordonnées deviennent comme suit :

$$U^k = (ex) \cdot f \cdot \frac{-\cos(\theta + \alpha) \cdot x + \sin(\theta + \alpha) \cdot y}{\sin(\theta + \alpha) \cdot x + \cos(\theta + \alpha) \cdot y + d - f} \dots\dots\dots 3.3$$

$$V^k = (ez) \cdot f \cdot \frac{-Z}{\sin(\theta + \alpha) \cdot x + \cos(\theta + \alpha) \cdot y + d - f} \dots\dots\dots 3.4$$

3.6 Relation entre les pentes d'un segment de droite dans les images successives [25]

Soit $v_0 = a_{i,0} \cdot u_0 + b_{i,0}$ l'équation du segment dans une image où $a_{i,0}$ est la pente du segment et (u_0, v_0) les coordonnées 2D d'un point de segment de droite. Sachant que les coordonnées d'un point de l'image sur IM peuvent s'écrire :

$$U_0 = (e_x) \cdot f \cdot \frac{-X_{i,0}}{Y_{i,0} + d - f} \dots\dots\dots 3.5$$

$$V_0 = (e_z) \cdot f \cdot \frac{-Z_{i,0}}{Y_{i,0} + d - f} \dots\dots\dots 3.6$$

Où :

$$\begin{cases} X_{i,0} = \cos(\theta + \alpha) \cdot x - \sin(\theta + \alpha) \cdot y \\ Y_{i,0} = \sin(\theta + \alpha) \cdot x + \cos(\theta + \alpha) \cdot y \\ Z_{i,0} = Z \end{cases}$$

Alors après un certains nombres de transformation, on obtient :

$$\frac{a^k - a_{i,0} \cos(\alpha)}{\sin(\alpha)} = \frac{b^0}{ex \cdot f} = \delta$$

3.6.1 Equations de mouvement d'un segment I

- - Après rotation d'un angle α :

$$\frac{a_{i,1} - a_{i,0} \cos(\alpha)}{\sin(\alpha)} = \frac{b_{i,0}}{e_{x,f}} \dots\dots\dots 3.7$$

- Après rotation d'un angle β :

$$\frac{a_{i,1} - a_{i,0} \cos(\beta)}{\sin(\beta)} = \frac{b_{i,0}}{e_{x,f}} \dots\dots\dots 3.8$$

3.7 Détermination du mouvement

pour déteminer le mouvement, on a suivi la méthode ci-après :

A partir du rapport entre les équations (3.7, 3.8), on peut déduire que : pour le segment i

$$\frac{a_{i,1} - a_{i,0} \cos(\alpha)}{a_{i,2} - a_{i,0} \cos(\beta)} = \frac{\sin(\alpha)}{\sin(\beta)} \dots\dots\dots 3.9$$

Pour le segment j, on obtient :

$$\frac{a_{j,1} - a_{j,0} \cos(\alpha)}{a_{j,2} - a_{j,0} \cos(\beta)} = \frac{\sin(\alpha)}{\sin(\beta)} \dots\dots\dots 3.10$$

A partir de (3.9, 3.10), on aura :

$$\frac{a_{i,1} - a_{i,0} \cos(\alpha)}{a_{i,2} - a_{i,0} \cos(\beta)} = \frac{a_{j,1} - a_{j,0} \cos(\alpha)}{a_{j,2} - a_{j,0} \cos(\beta)} \dots\dots\dots 3.11$$

Cette égalité devient une équation à deux inconnus de la forme suivante :

$$c_1 * (\cos \beta) + c_2 * (\cos \alpha) = c_3 \dots\dots\dots A$$

Tels que :

$$\begin{cases} c_1 = a_{j,1} * a_{i,1} - a_{i,1} * a_{j,0} \\ c_2 = a_{i,2} * a_{j,0} - a_{j,2} * a_{i,0} \\ c_3 = a_{j,1} * a_{i,2} - a_{i,1} * a_{j,2} \end{cases}$$

Pour résoudre cette équation, il faut avoir un système d'équation à deux inconnues pour cela, il suffit de prendre deux autres segments k et l pour obtenir la deuxième équation manquante qui sera de la forme :

$$c_{p,1} * (\cos \beta) + c_{p,2} * (\cos \alpha) = c_{p,3} \dots\dots\dots B$$

Tels que :

$$\begin{cases} c_{p,1} = a_{l,1} * a_{k,0} - a_{k,1} * a_{l,0} \\ c_{p,2} = a_{k,2} * a_{l,0} - a_{l,2} * a_{k,0} \\ c_{p,3} = a_{l,1} * a_{k,2} - a_{k,1} * a_{l,2} \end{cases}$$

La résolution du système linéaire à deux équations (A et B) pour deux inconnus $\cos \alpha$ et $\cos \beta$ est comme suit :

$$\begin{cases} c_1 * \cos \beta + c_2 * \cos \alpha = c_3 \\ c_{p,1} * \cos \beta + c_{p,2} * \cos \alpha = c_{p,3} \end{cases}$$

$$Det = [(c_1 * c_{p,2}) - (c_{p,1} * c_2)] \neq 0$$

$$\cos \alpha = [(c_3 * c_{p,2}) - (c_{p,3} * c_2)] \setminus Det$$

$$\cos \beta = [(c_1 * c_{p,3}) - (c_{p,1} * c_3)] \setminus Det$$

Il faut que les solutions trouvées respectent les deux conditions suivantes :

$$\begin{cases} |\beta| > |\alpha| \rightarrow \cos(\alpha) > \cos(\beta) \\ et \\ |\cos(\beta)| \leq 1 \text{ et } |\cos(\alpha)| \leq 1 \end{cases}$$

La résolution de ce système nécessite 4 segments dans 3 séquences d'images. La valeur de la solution est le mouvement effectué par la caméra pour l'obtention de l'image suivante est donc une valeur qui reste constante si les segments considérés correspondent aux bons appariements.

L'algorithme suivi :**Début :**

Extraction de points d'intérêt respectivement sur (IM1 ; IM2 ; IM3)

Elaboration des segments de IM1 ;

Calculer la pente de chaque segment de IM1

Elaboration des segments de IM2 ;

Calculer la pente de chaque segment de IM2 ;

Elaboration des segments de IM3 ;

Calculer la pente de chaque segment de IM3 ;

Rassembler en groupe de quatre les segments de IM1 dans T1 ;

Rassembler en groupe de quatre les segments de IM2 dans T2 ;

Rassembler en groupe de quatre les segments de IM3 dans T3 ;

- Pour tout groupe de segments de IM1 faire
 - Pour tout groupe de segments de IM2 faire
 - Pour tout groupe de segments de IM3 faire
 - **Calculer** $[\cos \alpha; \cos \beta]$;
- Fin pour
 - Fin pour
 - Fin pour

Calculer le score pour chaque solution trouvée.

Sélectionner les groupes de segments qui engendrent le plus grand score.

Etablir la correspondance entre les segments de droite dans les différentes images ayant engendré le plus grand score.

Fin

3.8 Appariement RANSAC

Pour vérifier si les mises en correspondances obtenues sont correctes, on a programmé en plus l'algorithme Ransac (Radom échantillon consensus) qui est un algorithme qui permet de trouver le meilleur modèle pour s'adapter aux données tout en n'étant pas trop affecté par des valeur aberrantes pour la bonne mise en correspondance

Algorithme :

Idée générale : (pas d'application particulière)

- Sélectionner plusieurs fois, au hasard, le nombre de données requises pour évaluer les paramètres du modèle.
- Évaluer la performance de cette paramétrisation avec l'ensemble des données.
- Avec les meilleurs paramètres, on raffine le modèle.

3.9 Conclusion

Dans ce chapitre nous avons proposé une méthode d'estimation du mouvement ne nécessitant pas la connaissance de l'appariement des primitives.

Elle ne dépend pas des paramètres intrinsèques (distance focale) de la caméra. Les résultats d'experimentation feraient l'objet du chapitre suivant.

Chapitre 4

Résultat d'expérimentation

4.1 Introduction

Dans ce chapitre, nous présentons les résultats d'expérimentation obtenus sur des images réelles, après avoir programmé l'algorithme d'estimation du mouvement en utilisant matlab.

4.2 Etapes d'expérimentation

Les images de test :



FIG. 4.1 – Images de test.

4.2.1 Extraction des points d'intérêt

Après avoir appliqué l'algorithme de Harris[26] on a pu détecter les points d'intérêts sur les images suivantes :

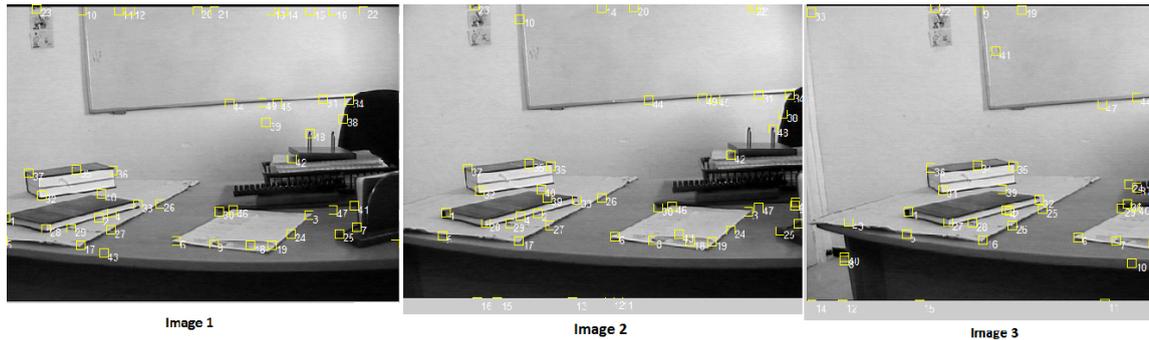


FIG. 4.2 – Extraction des points d'intérêts.

4.2.2 Appariement des segments

Chaque deux points d'intérêts forment un segment, et pour appliquer notre algorithme on a formé les segments suivants (Figure 4.3) à partir des points d'intérêts extraits précédemment.

Liste des segments			
	image1	image2	image3
segment1	0.2520	0.2459	0.2417
segment2	-0.2979	-0.3043	-0.2979
segment3	0.2627	0.2564	0.2500
segment4	-0.2692	-0.2745	-0.2745
segment5	0.1000	0.0952	0.0898
segment6	0.6338	0.6197	0.6232

FIG. 4.3 – les segments.

4.2.3 Regroupement des segments

Pour estimer le mouvement, il faut 4 segments, d'où le regroupement précédents en groupe de 4 (voir figure 4.4) :



FIG. 4.4 – Regroupement des segments.

4.2.4 Estimation du mouvement

Après l'application de l'algorithme proposé au chapitre précédent, nous obtenons les résultats suivants pour $\cos \alpha$ et $\cos \beta$ le test a été limité à l'objet livre. Nous remarquons que pour un taux d'erreur de 0.1 on a obtenu un résultat d'angle moyen de $\alpha = 37.81$ et $\beta = 84.37$

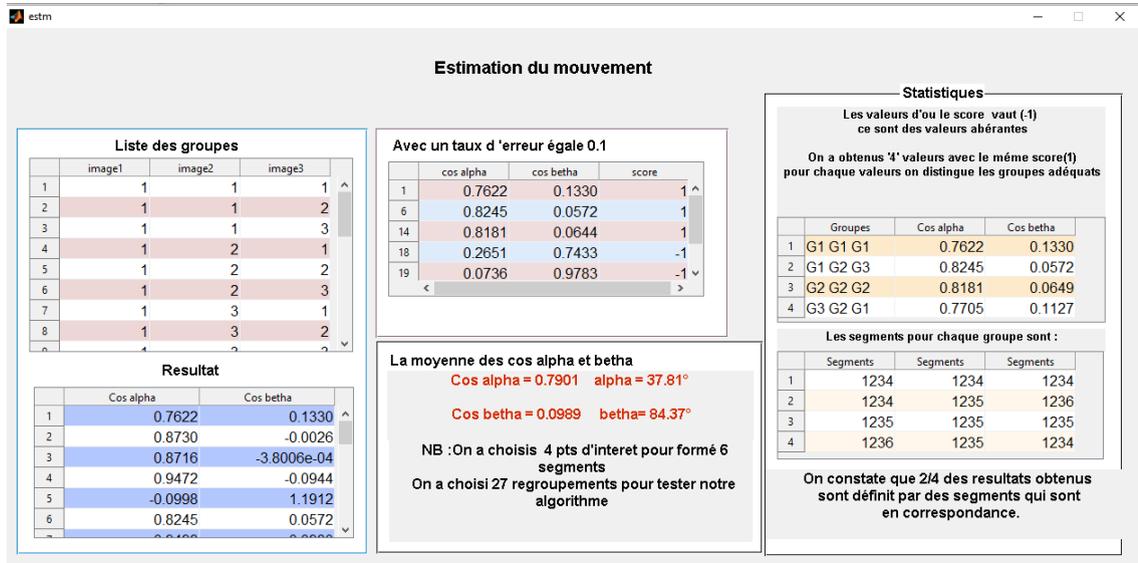


FIG. 4.5 – Estimation du mouvement.

4.2.5 Images résultat avec les points en correspondance

Les groupes de segments ayant généré le même $\cos \alpha$ et $\cos \beta$ sont en correspondance, c'est ce que montre la figure suivante (les points d'intérêt en rouge montrent les points en correspondance).

a/captu.png

FIG. 4.6 – Images résultat.

4.3 Résultats d'expérimentation sur une autre image

Le test à été effectué sur d'autres images à savoir :



FIG. 4.7 – Images de test.

4.4 Extraction des points d'intérêt

Avec le même algorithme appliqué sur les images précédente, on a détecté les points d'intérêt sur les images suivantes :

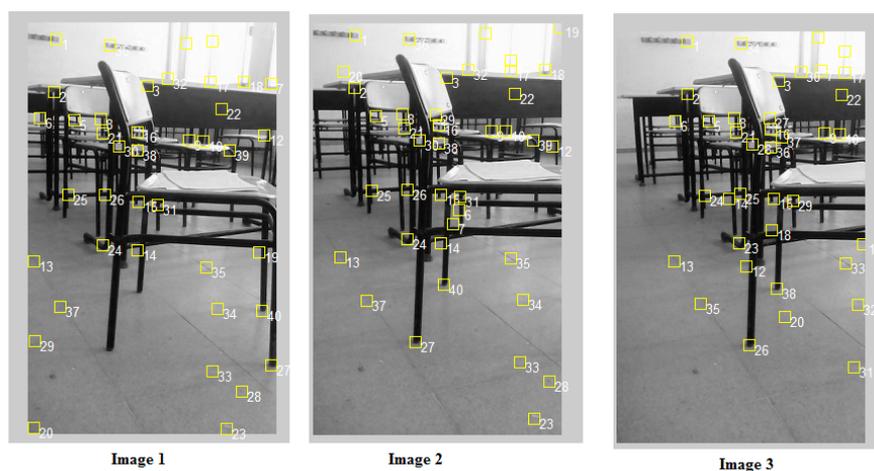


FIG. 4.8 – Extraction des points d'intérêt.

4.4.1 Estimation du mouvement

Après l'application de l'algorithme proposé au chapitre précédent nous obtenons les résultats suivants pour $\cos \alpha$ et $\cos \beta$.

Nous remarquons que pour un taux d'erreur de 0.1 on a obtenu un résultat d'angle moyen de $\alpha = 27.66$ et $\beta = 66.42$

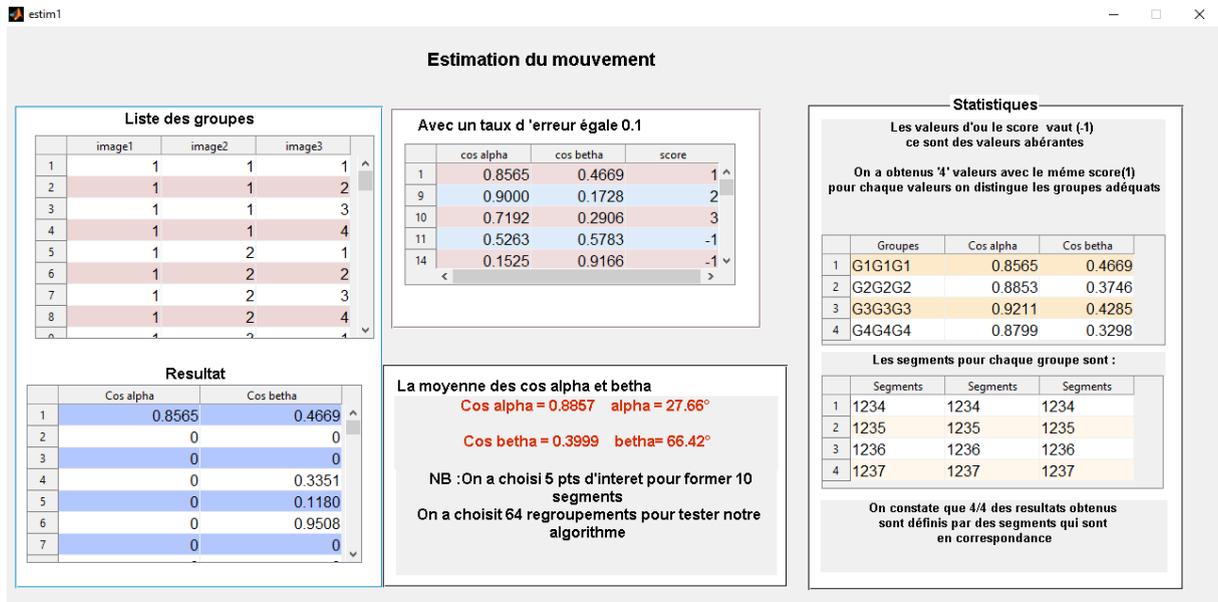


FIG. 4.9 – Estimation du mouvement.

4.4.2 Images résultat avec les points en correspondance

Les groupes de segments ayant généré le même $\cos \alpha$ et $\cos \beta$ sont en correspondance, c'est ce que montre la figure suivante (les points d'intrêt en rouge montre les points en correspondance).

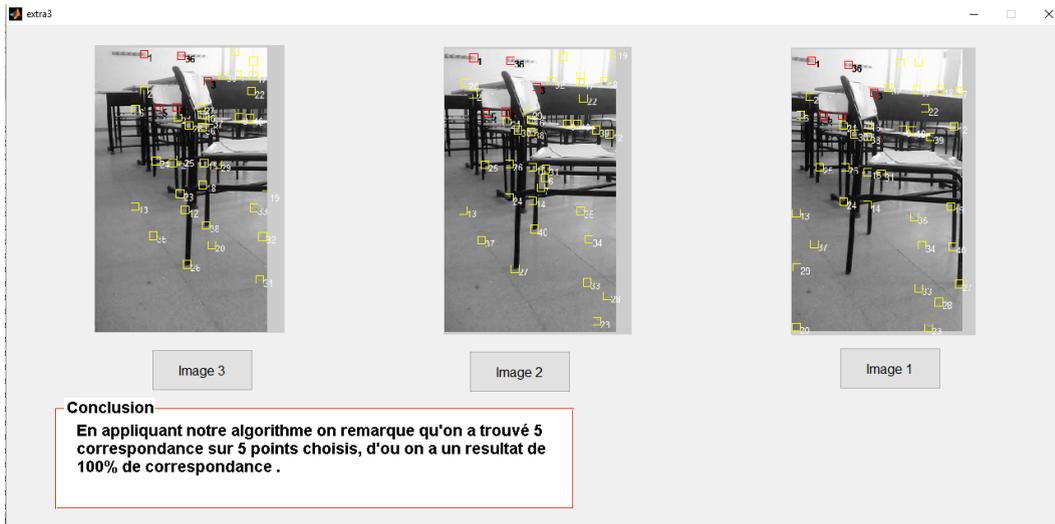


FIG. 4.10 – Images résultat.

4.5 Conclusion :

L'algorithme donne d'assez bon résultat d'appariement et d'estimation du mouvement, le teste à été effectuer sur de petits objets de l'image, Il serait inéressant de le généraliser à toute l'image afin de détrminer le taux de reconnaissance et le temps de réalisation.

Conclusion générale et perspectives

Les progrès en vision par ordinateur, branche de l'intelligence artificielle, permet aux machines d'extraire de l'information à partir d'images. Diveres approches sont apparues, permettant des applications variées. telle que la reconnaissance d'objets, la construction de modèles virtuels, le positionnement de robots mobiles, l'estimation du mouvement de la caméra qui est l'un des domaines du traitement d'images qui joue un rôle considérable dans l'analyse des séquences d'images d'ou le nombre considérable de ses applications. Dans cette étude, nous nous sommes intéressés à l'estimation du mouvement dans une séquence d'images.

Un aperçu en premier temps sur les généralités de la vision par ordinateur et des systèmes de cette dernière.

Une étude bibliographique des différentes techniques d'estimation du mouvement a été abordée, en introduisant d'abord les techniques différentielle(flots optiques), en suite les techniques de mise en correspondance et de ses contraintes fondamentales. Nous avons présenté l'algorithme à suivre pour atteindre l'objectif sur lequel est basé notre recherche et la démarche que nous avons suivie.

Nous avons donné un aperçu sur des résultats obtenus et présenté quelques interfaces qui illustrent notre travail.

Les résultats obtenus sont assez satisfaisants, le mouvement de la caméra est estimé avec le calcul de l'angle de rotation de la caméra entre les images.

En guise de perspective, nous allons affiner nos résultats pour avoir des calcul de l'angle bien précis et de meilleurs appariements en utilisant Ransac pour éliminer les données aberrantes.

Bibliographie

- [1] A.Zakhor and F.Lari, Edge based 3-d camera motion estimation with application to video coding. *IEEE Transactions on Image Processing* 2, 4 (1993), 481-498
- [2] T.Needham "Visual Complex Analysis", Oxford University Press, 1997.
- [3] D. Marr, Freeman and Company San-Francisco, 'Vision : A computational investigation into human representation and processing of visual', 1982.
- [4] R.Abaoui, Elaboration et appariement des trajectoires dans un système de vision stéréoscopique en mouvement de convergence, mémoire d'ingénierie université frehat abbasse sétif, avril 2008.
- [5] A.Miticheand et H.Sekkati, Optical Flow 3D Segmentation and Interpretation : A Variational Method with Active Curve Evolution and Level Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006. 28(11) : p. 1818-1829.
- [6] R.HORAUD et O.MONGA, Vision par ordinateur outils fondamentaux ouvrage rédigé avec le concours du Ministère de l'enseignement supérieur et de la recherche (DIST), Deuxième édition, Editions Hermès,1995
- [7] S.Martin Auto-calibrage de caméras et reconstruction 3-D 'à partir d'images Peter Sturm INRIA Rhone-Alpes 655 Avenue de l'Europe, 38330 Montbonnot , 2012
- [8] L.G. Brown. A survey of image registration techniques. *ACM Computing Surveys*, 24(4) :325-376, 1992.

- [9] B.Zitova, J.Flusser , P.Sroubek , Image Registration : A Survey and Recent Advances, ICIP 2005 [9]
- [10] A.BOUCHER -IFI. 'Mouvement (computer vision - A modern Approach, chap. 17)'.2011
- [11] M.Cagnazzo," Estimation du mouvement " ,TELECOM Paris-Tec, 14 mai 2012.
- [12] C.Jonchery.'Estimation d'un mouvement de cam'era et problèmes connexes '. Thèse PHD soutenue le 6 novembre 2006, de l'école Normale Supérieure de CACHAN.
- [13] R.Abaoui, Elaboration et appariement des trajectoires dans un système de vision stéréoscopique en mouvement de convergence, mémoire d'ingénierie université frehat abbasse sétif, avril 2008.
- [14] S.CHAMBON, 'Mise en correspondance stéréoscopique d'images couleur en absence d'occlusions'. Thèse de doctorat soutenue le 14 décembre 2005, l'école doctorale Informatique et Télécommunications, Université de TOULOUSE III.
- [15] N.SUVONVORN, 'Mise en correspondance d'images pour l'analyse du mouvement et la stéréovision '. Thèse de doctorat soutenue le 18 décembre 2006, Ecole doctorale " Sciences et techniques de l'information des Télécommunications et des systèmes ", Université de Paris Sud.
- [16] M.A.Fischler and R.C.Bolles, RandomSample Consensus : A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of the ACM, 1981, 24, p. 381-395.
- [17] Y.Furukawa,, et al, Robust structure and motion from outlines of smooth curved surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006,28(2) : p. 302-315.
- [18] J.Wills, S.Agarwal, and S.Belongie, A Feature-based Approach for Dense Segmentation and Estimation of Large Disparity Motion. International Journal

of Computer Vision, 2006,68(2) : p. 125-143.

- [19] Torr, P.H.S. and A. Zisserman, MLESAC : A New Robust Estimator with Application to Estimating Image Geometry. Computer Vision and Image Understanding, 2000, 78(1), p; 138-156.
- [20] Torr, P.H.S. and C. Davidson, IMPSAC : synthesis of importance sampling and random sample consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(3), p. 354-364.
- [21] J. Domke and Y. Aloimonos. A Probabilistic Notion of Correspondence and the Epipolar Constraint. In3DPVT, 2006.
- [22] F. Dellaert, S.M. Seitz, C. Thorpe, and S. Thrun. EM, MCMC, and chain flipping for structure from motion with unknown correspondence. Machine Learning, 50(1-2), 2003.
- [23] N. AYACHE. 'Vision stéréoscopique et Perception multi sensorielle'. Application à la robotique mobile 1989.
- [24] C.Achard sur la base des cours de J.Devars et M.Milgram "cours traitement d'image", 2002-2003
- [25] S. Ait Kaci Azzou, M.C Baba Hnini, " Tracking Line Segment without Knowledge of Camera Motion", International Journal of Computer Applications (0975 - 8887) Volume 73- No.6, July 2013
- [26] operationpixel@free.fr, Copyright © 2010-2014

Résumé

L'analyse des séquences temporelles d'images a fait l'objet de nombreux travaux. Elle permet d'accéder à des informations importantes sur les objets contenus dans les scènes observées que ce soit à des fins de détection, d'estimation, de suivi ou de reconnaissance. Notre intérêt s'est porté sur l'estimation du mouvement des caméras. Pour cela, nous avons proposé un algorithme qui permet de calculer le mouvement et de résoudre le problème d'appariement simultanément ; l'algorithme a été testé sur des données réelles.

Abstract

Image sequence analysis has been a major topic. It allows accessing important information on the objects of interest for detection, estimation, tracking and recognition tasks.

Our interest is inclined the estimation of the cameras's movement. This is why we proposed an algorithm which permits to calculate the movement resolute the problem of appariement simultaneously ; the algorithm had been tasred with real data.