

## ESTIMATION DE CARTES DE PROFONDEUR A PARTIR D'UNE CAMERA MOBILE

Reçu le 13/01/2001 – Accepté le 02/06/2003

### Résumé

Dans cet article, nous présentons un système d'estimation de cartes de profondeur à partir d'une caméra mobile. Ce système est basé sur la modélisation, par un modèle dynamique, de la structure changeante de la scène perçue par une caméra CCD dont le mouvement relatif par rapport à la scène est inconnu. L'état de ce système est la profondeur et la mesure est le flot optique. Nous employons ce modèle de systèmes dynamiques dans la construction d'un filtre de Kalman qui estime de façon optimale la profondeur en tout point de l'image.

Le traitement est effectué pour chaque paire d'images successives en trois étapes :

- Estimation du champ du flot optique et de sa variance,
- Estimation du mouvement 3D de la caméra par la méthode des moindres carrés,
- Récupération d'une carte dense de profondeur de l'environnement en utilisant un filtre de Kalman par pixel d'image.

Les performances de ce système sont illustrées sur des séquences d'images synthétiques et réelles.

**Mots clés:** Carte de profondeur, Flot optique, Système dynamique, Mouvement 3D, Filtre de Kalman.

### Abstract

In this paper, we present a depth map estimation system from a moving camera. This system is based on the modelisation, by a dynamical system, of the changing structure of the scene perceived in the camera, moving in an environment with unknown motion. The state of this system is the depth and the measurement is the optical flow. We use this dynamical systems model in the construction of a Kalman filter which optimally estimates the depth of each image point.

The treatment is carried out for each pair of successive images in three stages :

- Estimate of the optical flow field and its variance,
- Estimate of the 3D motion of the camera by the square least method,
- Recovery a dense depth map of the environment using one Kalman filter per image pixel.

The performances of this system are illustrated on synthetic and real image sequences.

**Keywords:** Depth map, Optical flow, Dynamical system, 3D motion, Kalman filter.

A. BOUBIDI  
H. TEBBIKH

L.A.I.G.

Université Guelma  
BP 401, 24000 Guelma  
Algérie

### ملخص

في هذه المقالة، نعرض نظاماً لتقدير بطاقات العمق عن طريق كاميرا متحركة. هذا النظام يعتمد على نمذجة، بواسطة نظام ديناميكي، البنية المتغيرة للمشاهد الملتقط بكاميرا تتحرك بالنسبة للمشاهد بحركة غير معروفة. حالة هذا النظام هي العمق والقياس هو التغير الضوئي. نستعمل هذا النموذج للأنظمة الديناميكية في تكوين مصفاة "كالمان" التي تقدر وبصورة مثلى العمق في كل نقطة من الصورة. تتم المعالجة من أجل كل ثنائية لصورتين متتاليتين على ثلاثة مراحل :

- تقدير التغير الضوئي وتباينه،
  - تقدير الحركة ثلاثية الأبعاد للكاميرا الحاصلة بين الصورتين بواسطة طريقة أقل التربيغات،
  - استرجاع بطاقة مكثفة لعمق المحيط باستعمال مصفاة "كالمان" عند كل نقطة من الصورة.
- نجاحة هذا النظام بينت على سلاسل متتابعة من الصور الإصطناعية والحقيقية.

**الكلمات المفتاحية:** بطاقة عمق، التغير الضوئي، نظام ديناميكي، حركة ثلاثية الأبعاد، مصفاة كالمان.

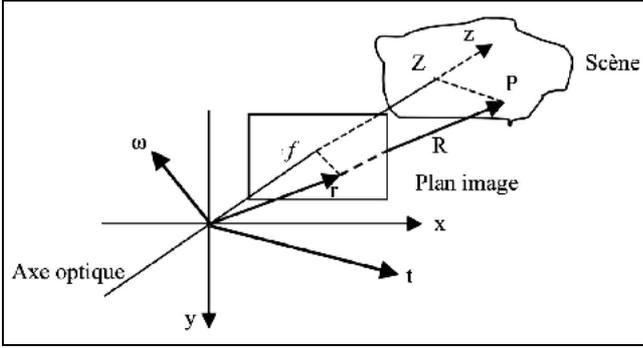
Le problème d'estimation de la structure a été largement étudié ces deux dernières décennies. En particulier, deux voies importantes ont été largement examinées : la stéréovision dynamique [1, 2] dans laquelle le système d'acquisition mobile est composé de plusieurs caméras rigidement liées, et la profondeur à partir du mouvement [3, 4], dans laquelle une caméra unique observe la scène rigide.

Si la première voie a continué à motiver les chercheurs ces dernières années [5], la seconde a été abandonnée au profit de la vision active [6]. Cependant, nous estimons qu'il est toujours intéressant d'exploiter l'image naturelle, du moins pour l'élargissement du champ de la connaissance... C'est dans cet esprit que nous avons développé, dans ce travail, un système de vision dynamique à base d'une caméra CCD standard. L'objectif est de modéliser automatiquement des scènes réelles, sous forme de cartes de profondeur, à partir de séquences d'images acquises par la caméra déplacée devant la scène dont le mouvement relatif est inconnu. Un tel processus nécessite la manipulation de l'incertitude, ce qui exige l'introduction d'un modèle de bruit tel que le filtrage de Kalman qui a montré son efficacité à la reconstruction 3D de primitives [7] aussi bien que de cartes denses de profondeur [8, 4].

### 1- PRESENTATION DE LA SCENE

Nous adaptons le système de coordonnées représenté dans la figure 1 dont l'origine coïncide avec le point focal de la caméra et l'axe  $z$  avec l'axe optique. Les points du monde réel sont désignés par les lettres

majuscules  $R=[X, Y, Z]^T$ , tandis que les lettres minuscules désignent les points dans l'image  $r=[x, y, z]^T$ .



**Figure 1:** Système de coordonnées et projection perspective.

On note que la distance  $f$  entre l'origine et le plan image est la longueur focale du système de la caméra qu'on suppose égale à 1 dans le but de simplifier notre étude.

Le rapport entre les points du monde réel et les points dans l'image est donné par:

$$\frac{R}{Z} = \frac{r}{f} \Rightarrow \frac{1}{Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{f} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (1)$$

Intéressons nous à la description du mouvement. Evidemment, il y a trois cas possibles, à savoir, une caméra mobile, un objet mobile ou les deux à la fois. Cependant, pour la description de notre système, il n'y a que le mouvement 3D relatif de la caméra par rapport à la scène qui est évident et qui peut être représenté par un vecteur de translation  $t=[t_x, t_y, t_z]^T$  et un vecteur de rotation  $\omega=[\omega_x, \omega_y, \omega_z]^T$ .

## 2- UN MODELE DE SYSTEME DYNAMIQUE POUR LA SCENE

### 2.1- Un modèle d'état

Soit  $P$  un point de l'objet perçu par une caméra en déplacement par un mouvement relatif par rapport à la scène dont les paramètres sont  $t$  et  $\omega$ . Si  $R$  est son vecteur de position, nous avons:

$$\dot{R} = -t - \omega AR \quad (2)$$

La troisième composante du vecteur de l'équation différentielle (2) a un intérêt particulier. Elle décrit le changement temporel de la profondeur  $Z$  de l'environnement. Cette dernière représente ainsi un attribut particulier du monde réel. Il serait donc extrêmement utile de formuler un système dynamique qui décrit le comportement temporel de cette quantité.

Plus formellement, nous trouvons que la troisième composante de l'équation (2) est:

$$\dot{Z} = -t_z - (\omega_x Y - \omega_y X) \quad (3)$$

En utilisant l'équation de la projection perspective (1), nous obtenons:

$$\dot{Z} = -t_z - (\omega_x y - \omega_y x) Z \quad (4)$$

qui est une équation d'état simple pour la profondeur  $Z$ .

Pour le moment, nous supposons que les paramètres du mouvement sont connus et nous expliquerons, dans le paragraphe 7, comment ils peuvent être obtenus.

### 2.2- Un modèle de mesure: le flot optique

Pour compléter la modélisation du système dynamique, nous devons formuler l'équation de mesure. Cette mesure est un vecteur de déplacement  $[u, v]^T$  dit "flot optique", estimé pour chaque point de l'image  $(x, y)$  en utilisant deux images comme entrées. Nous supposons que ce procédé d'estimation est disponible.

La question qui se pose est: qu'elle est la relation entre les quantités mesurables  $[u, v]^T$  et l'état  $Z$ ?

Sachant que:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \quad (5)$$

et à partir des équations (01) et (02), nous obtenons:

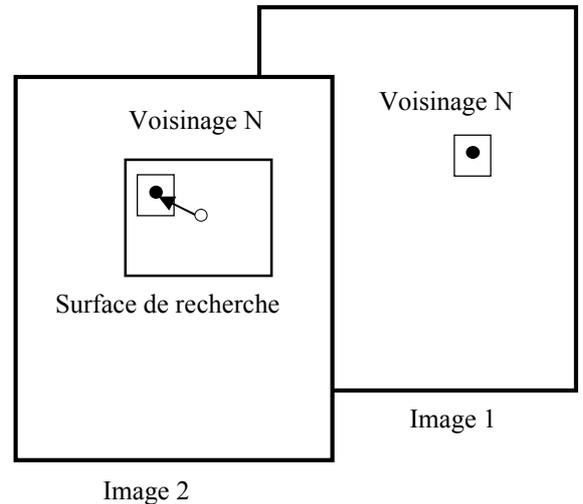
$$u = \frac{-t_x + xt_z}{Z} + \omega_x xy - \omega_y (x^2 + 1) + \omega_z y \quad (6.a)$$

$$v = \frac{-t_y + yt_z}{Z} + \omega_x (y^2 + 1) - \omega_y xy - \omega_z x \quad (6.b)$$

Ce système d'équations est la sortie de notre système dynamique.

## 3- DETAILS DE LA MESURE DU FLOT OPTIQUE

Un champ du flot optique est un champ de vecteurs qui, pour chaque point  $p=(x, y)$  de l'image actuelle, qui devient  $p'=(x', y')$  dans la prochaine image, associe un vecteur  $[u, v]^T$  décrivant sa vitesse de déplacement entre les deux images. L'idée principale [8], exploitée dans notre travail, pour l'estimation du flot, est basée sur la méthode de mise en correspondance, dite SSD (Sum of Squared Differences). La figure 2 illustre cette idée.



**Figure 2:** Calcul du flot optique par la méthode SSD.

Nous supposons deux contraintes :

- Les déplacements inter-images n'excèdent pas un certain nombre de pixels dans chaque image.
- La luminosité dans une surface entourant le point demeure approximativement inchangée par le mouvement.

La première supposition mène au concept de surface de recherche qui est un ensemble de pixels dans l'image 2 que nous pouvons considérer comme correspondances possibles de  $p$ .

La deuxième supposition introduit la notion de la somme des carrés des différences. Pour savoir à quel point candidat  $p'$  dans l'image 2 correspond le mieux au point  $p$ , nous emploierons la différence entre un voisinage autour de  $p$  et le voisinage correspondant autour de  $p'$ . La mesure est donnée par la somme des carrés des différences calculée en tout point candidat de la surface de recherche.

Plus formellement, si nous notons  $E_1(x,y)$  les valeurs de luminosité aux endroits  $p=(x,y)$  dans l'image 1 et  $E_2(x,y)$  les valeurs de luminosité dans l'image 2, nous cherchons pour chaque  $p=(x,y)$  dans l'image 1 un  $p'=(x',y')$  dans l'image 2, tel que :

$$SSD(x',y') = \min_{x',y' \in S} \sum_{\Delta x, \Delta y \in N} (E_1(x+\Delta x, y+\Delta y) - E_2(x'+\Delta x, y'+\Delta y))^2 \quad (7)$$

où  $N$  désigne le voisinage du point et  $S$  la surface de recherche.

Le vecteur du flot optique au point  $p$  est alors:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{T} \begin{bmatrix} x'-x \\ y'-y \end{bmatrix} \quad (8)$$

où  $T$  est le temps entre deux images successives.

Le principe décrit précédemment détermine uniquement la vitesse de déplacement du point. Ceci est certainement insuffisant pour l'estimation de la profondeur qui nécessite la manipulation de l'incertitude. Il faut associer à chaque vecteur mouvement  $[u, v]^T$  un vecteur de variance  $[\delta_u, \delta_v]^T$  qui peut être obtenu en modélisant la surface  $SSD$  locale. La variance est relative aux paramètres de cette surface.

#### 4- APPROXIMATION EN TEMPS DISCRET

Pour pouvoir utiliser le modèle "dérivé" du système dynamique ainsi conçu dans l'implantation du filtre de Kalman, nous aurons besoin de son équivalent en temps discret. Admettons la convention  $Z(KT)=Z_K$ , où  $T$  est le temps entre deux images successives. Nous approximations les dérivées par les premières différences de telle sorte que:

$$\dot{Z} \approx \frac{Z_{K+1} - Z_K}{T} \quad (9)$$

L'équation d'état devient alors :

$$Z_{K+1} = -Tt_z + [1 - T(\omega_x y - \omega_y x)] Z_K \quad (10)$$

et la sortie est simplement :

$$u_K = \frac{-t_x + xt_z}{Z_K} + \omega_x xy - \omega_y (x^2 + 1) + \omega_z y \quad (11.a)$$

$$v_K = \frac{-t_y + yt_z}{Z_K} + \omega_x (y^2 + 1) - \omega_y xy - \omega_z x \quad (11.b)$$

#### 5- ESTIMATION DE LA PROFONDEUR PAR LE FILTRE DE KALMAN

Nous employons le modèle des systèmes dynamiques dans la construction d'un filtre de Kalman qui estime de façon optimale la profondeur d'un point de l'image. En utilisant un filtre de Kalman par pixel d'image, nous pouvons récupérer une carte de profondeur dense de l'environnement.

Le filtre se compose de deux étapes: l'une, de mise à jour, l'autre, de prédiction.

##### a) Composantes de mise à jour

La mise à jour du filtre fonctionne en corrigeant simultanément l'estimation courante de la profondeur  $Z_K^-$

et sa variance  $\delta_{Z,K}^{2-}$ . La mise à jour est indiquée par l'indice supérieur en remplaçant le "-" par le "+". Les équations de mise à jour de la profondeur et de sa variance sont respectivement:

$$Z_K^+ = Z_K^- + K_K \begin{bmatrix} u_K \\ v_K \end{bmatrix} - \begin{bmatrix} u_K^- \\ v_K^- \end{bmatrix} \quad (12.a)$$

$$\delta_{Z,K}^{2+} = (1 - K_K C_K) \delta_{Z,K}^{2-} \quad (12.b)$$

avec :

$$K_K = \delta_{Z,K}^{2-} C_K^T [C_K C_K^T \delta_{Z,K}^{2-} + R_K]^{-1} \quad (12.c)$$

$$C_K = \begin{bmatrix} (t_x - xt_z) / Z_K^- \\ (t_y - yt_z) / Z_K^- \end{bmatrix} \quad (12.d)$$

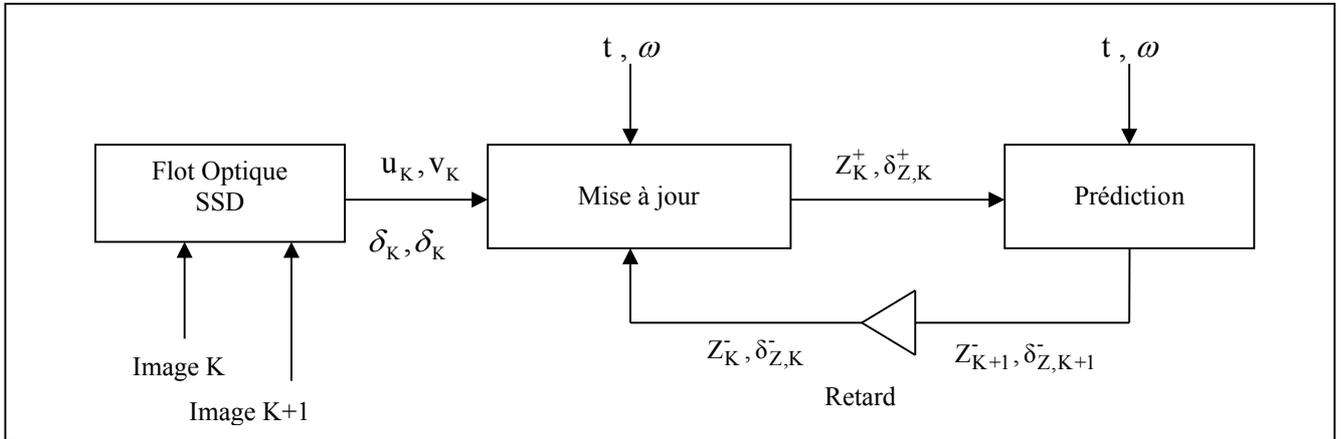
$$R_K = \begin{bmatrix} \delta_u^2 & 0 \\ 0 & \delta_v^2 \end{bmatrix} \quad (12.e)$$

où  $K_K$  est le gain du filtre,  $C_K$  un résultant du développement du filtre de Kalman étendu et  $R_K$  la matrice de covariance de la mesure du flot. Notons que nous avons supposé que les composantes du flot optique mesurées  $u_K$  et  $v_K$  sont indépendantes.

Par opposition au flot optique  $(u_K, v_K)$  qui est mesuré à partir des deux images  $K$  et  $K+1$ , les composantes  $u_K^-$  et  $v_K^-$  dans l'équation de mise à jour de l'état désignent le champ de mouvement prédit, calculé selon (11.a) et (11.b) en utilisant la valeur courante de la profondeur  $Z_K^-$  résultant de l'itération précédente du filtre.

##### b) Composantes de prédiction

Dans cette étape du filtre, l'estimation courante de la profondeur  $Z_K^+$  est simplement utilisée dans l'équation dynamique connue du système pour prédire la profondeur qui va servir pour la prochaine itération  $K+1$ . La variance de la profondeur est ajustée en conséquence. Les valeurs prédites sont, à l'itération  $K$  (traitons les images  $K$  et  $K+1$ ), indiquées par l'indice inférieur  $K+1$  et l'indice supérieur "-". Les équations de prédiction de la profondeur et de sa variance sont respectivement:



**Figure 3:** Schéma fonctionnel du filtre de Kalman.

$$Z_{K+1}^- = -Tt_z + \left[ 1 - T(\omega_x y - \omega_y x) \right] Z_K^+ \quad (13.a)$$

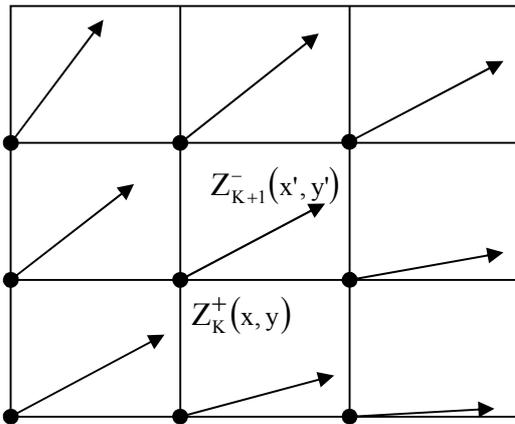
$$\delta_{Z,K+1}^- = \left[ 1 - T(\omega_x y - \omega_y x) \right]^2 \delta_{Z,K}^+ \quad (13.b)$$

Un schéma fonctionnel de l'algorithme de filtrage est présenté sur la figure 3.

En mettant ce filtre en application, nous rencontrons un certain nombre de difficultés. C'est ce que nous allons discuter dans le paragraphe suivant.

## 6- CORRECTION DE LA PROFONDEUR

La difficulté principale à laquelle nous sommes confrontés dans la mise en application de cette technique est la suivante (Fig. 4):



**Figure 4:** Problèmes rencontrés dans l'estimation de la carte de profondeur.

Le point  $(x, y)$ , sur lequel le filtre de Kalman actionne, se déplace entre les images. Par conséquent, la valeur prédite de la profondeur calculée à ce point dans l'itération  $K$  est inadmissible à ce même point dans l'itération suivante, mais plutôt à un certain  $(x', y')$  qui est l'endroit vers lequel  $(x, y)$  se déplace en raison du mouvement relatif de la caméra. Heureusement,  $(x', y')$  est donné par le vecteur du champ de mouvement mesuré dans l'itération  $K$ .

L'approche de base pour résoudre ce problème est de considérer les valeurs de profondeur calculées comme des échantillons d'une fonction à deux variables prises aux endroits  $(x', y')$  indiqués par le champ du mouvement. Nous construisons cette surface aux endroits des points de la grille par interpolation entre les échantillons donnés.

## 7- ESTIMATION DU MOUVEMENT DE LA CAMERA

Jusqu'ici nous avons supposé que les paramètres du mouvement sont connus. Dans la pratique, ceci peut être le cas d'une caméra élevée par un robot mobile se déplaçant dans milieu non accidenté. Cependant, dans la plupart des cas, le mouvement exact est inconnu et nous sommes en fait intéressés à l'estimer.

En examinant l'équation (11), on trouve qu'avec au moins 3 mesures du flot optique et les valeurs de la profondeur correspondantes, on peut formuler un problème des moindres carrés pour les paramètres du mouvement. Cependant, la profondeur  $Z$  n'est pas connue avec précision, elle est ce que le filtre estime ( $Z_K^+$ ). L'idée principale est d'utiliser l'estimation courante, dans ce cas  $Z_K^-$ , comme valeur de profondeur.

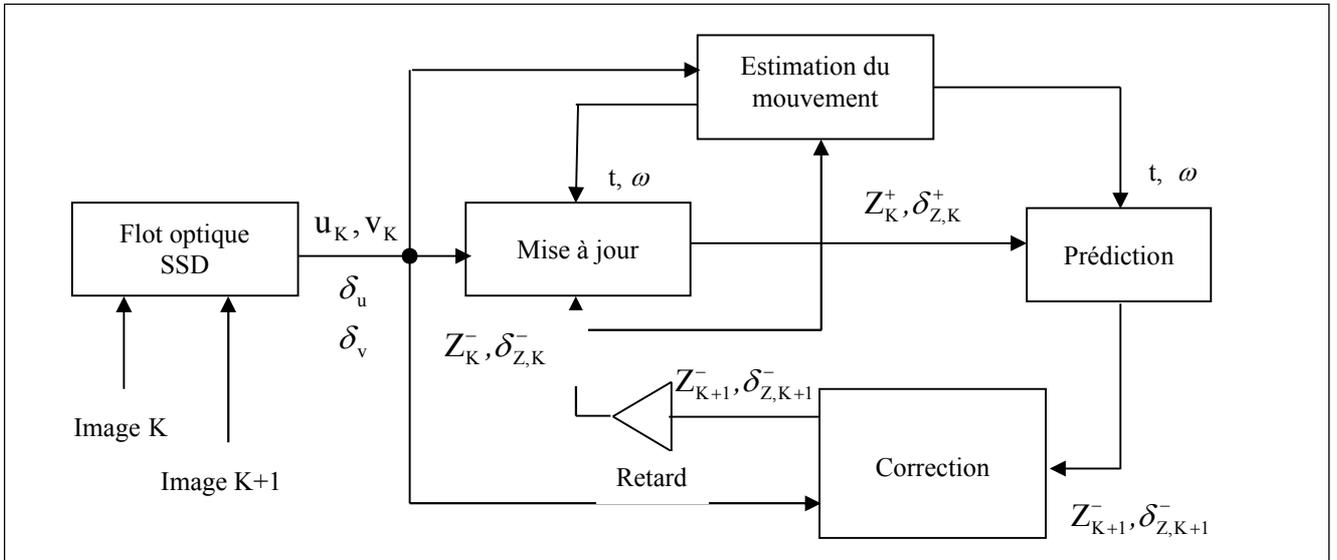
Nous pouvons maintenant présenter notre système dynamique d'estimation de la profondeur (Fig. 5).

## 8- EXPERIENCES

Pour vérifier le caractère pratique de l'approche proposée, nous avons mis notre algorithme à trois niveaux d'essai: sur des séquences de champs synthétiques du flot optique, puis sur des séquences d'images synthétiques et en fin sur des séquences d'images réelles. L'exécution était faite en Matlab sur un ordinateur Pentium 200 Mhz.

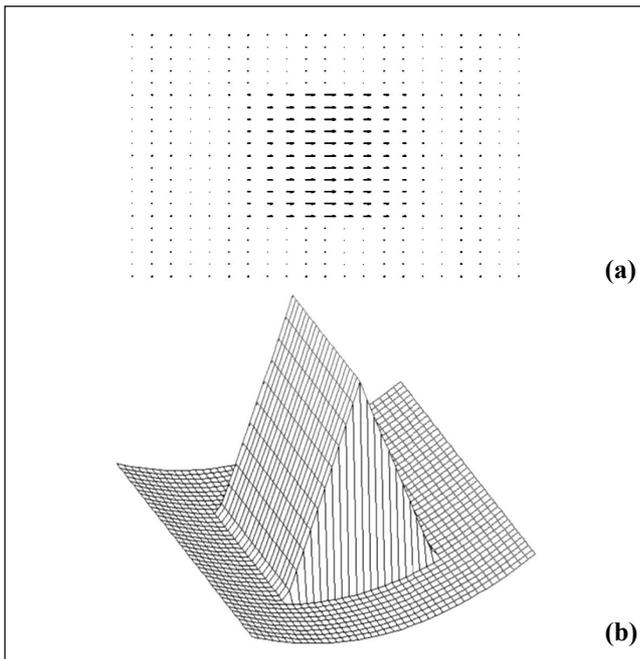
### 8.1- Séquence de champs synthétiques du flot optique

Au premier niveau, nous avons utilisé des champs synthétiques du flot optique afin de supprimer la quantité d'erreur introduite par le calcul du flot. Dans notre expérience, une séquence de champs synthétiques du flot optique, constituée de 5 champs, a été construite comme si



**Figure 5:** Système dynamique pour l'estimation de la profondeur.

la scène était regardée par un observateur passant au-dessus d'une structure triangulaire. Un regard au champ du flot optique montré dans la figure 6a, confirme l'idée intuitive que les points d'image correspondant au bout de la cale, qui sont les plus proches de l'observateur, ont le plus grand mouvement entre les images, ce qui se traduit, en ces points, par le plus grand flot optique. Le fond est choisi de forme plate. Donc, il est évident que l'estimation initiale de la structure soit plate. La figure 6 montre un des champs du flot optique ainsi que la structure récupérée par l'algorithme après une itération (correspondante à 2 images) et qui ne change pas de manière significative pendant les itérations restantes. Les erreurs constatées, telle que la forme du fond qui n'est plus plate, sont principalement dues aux erreurs d'estimation des paramètres inconnus du mouvement.



**Figure 6:** (a) Un champ synthétique du flot optique.  
(b) La structure rétablie à près une itération.

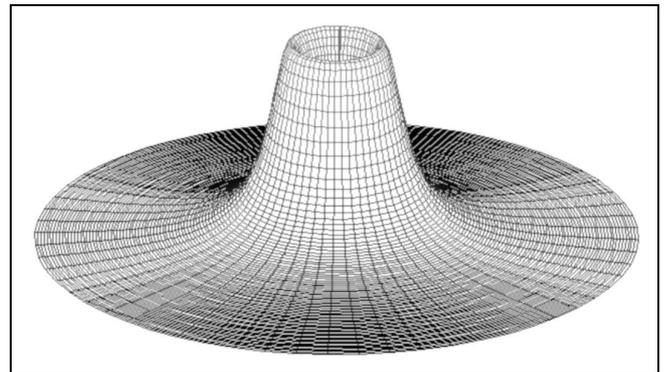
## 8.2- Séquences d'images synthétiques

L'étape suivante de notre essai expérimental était d'utiliser des images synthétiques générées par Matlab 5.3. Ce logiciel permet de générer des formes en représentation 3D, sélectionner le point de prise de vue, choisir les conditions d'illumination et produire l'image vue par l'observateur imaginaire placé dans l'endroit choisi.

Nous avons créé des séquences d'images synthétiques pour différentes formes d'objets prises par différents types de mouvements.

### a) Séquence "cratère"

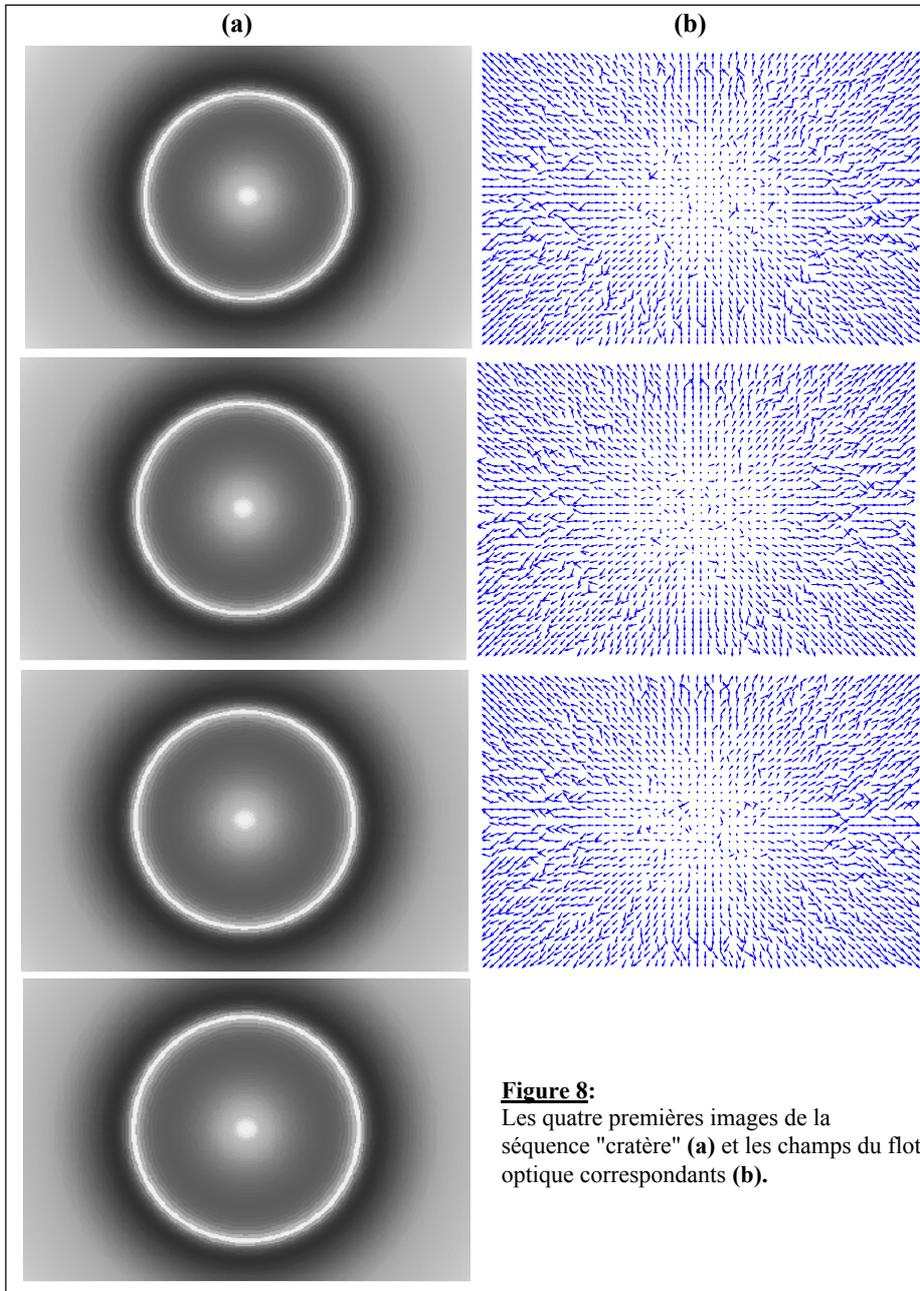
Nous avons créé un cratère rotationnellement symétrique. La figure 7 montre la représentation 3D de la structure construite que nous avons utilisé pour générer la séquence d'images.



**Figure 7:** La structure créée pour générer la séquence "cratère".

Pour la construction de la séquence "cratère", nous avons procédé comme suit:

- Le cratère est à une distance de 175 cm de la caméra dans sa position initiale et le fond est à 250 cm.
- Pour tester l'efficacité du procédé d'estimation du mouvement, nous avons choisi un mouvement non-uniforme de la caméra qui se translate de 5 cm, selon son axe optique  $Z$  (qui coïncide avec l'axe de symétrie du



**Figure 8:**  
Les quatre premières images de la séquence "cratère" (a) et les champs du flot optique correspondants (b).

cratère) à l'exception des images 5 et 6 où la translation entre elles est de 3 cm.

- Les distances et les mouvements ont été choisis de telle sorte que le déplacement entre les images soit au maximum de 6 pixels.

Les quatre premières images de la séquence "cratère", qui est constituée de 10 images, ainsi que les champs du flot optique correspondants sont montrées dans la figure 8.

Nous signalons que nous initialisons notre algorithme par une carte de profondeur plate de 212.5 cm (une valeur moyenne de la profondeur actuelle qui peut être obtenue par une sonde simple).

La figure 9a montre l'évolution de la composante de translation  $t_z$  mesurée du mouvement en la comparant avec le vrai mouvement (les autres composantes du mouvement tendent vers zéro). Bien qu'aucune prétention concernant

le mouvement n'a été prise, la non-uniformité est détectée. La figure 9.b présente l'erreur relative dans le rétablissement de la composante de translation du mouvement. Nous notons, cependant, que malgré que l'estimation initiale de la structure est plate, l'erreur sur le mouvement initial estimé est au dessous de 5.5%. Ceci nous permet d'affirmer que notre procédé d'estimation du mouvement, concernant les erreurs de la profondeur, est robuste. Notons aussi que malgré l'utilisation de la méthode des moindres carrés pour l'estimation du mouvement, qui est intercalée avec les étapes du filtre de Kalman, la figure 9.b montre une convergence rapide du procédé d'estimation du mouvement.

Enfin, la figure 10 montre la structure rétablie après traitement de la séquence en entier. Les valeurs des paramètres du mouvement et de la profondeur sont naturellement rétablies à un facteur près, qui dépend du choix de la valeur initiale de  $Z$  (voir illustration pour un point à la figure 11).

#### b) Séquence "vase"

Pour vérifier l'efficacité de notre algorithme pour d'autres types de scènes, nous avons estimé qu'une structure à faibles variations de profondeur est un bon test. Pour cela, nous avons créé un vase rotationnellement symétrique dont la forme est représentée dans la figure 12.

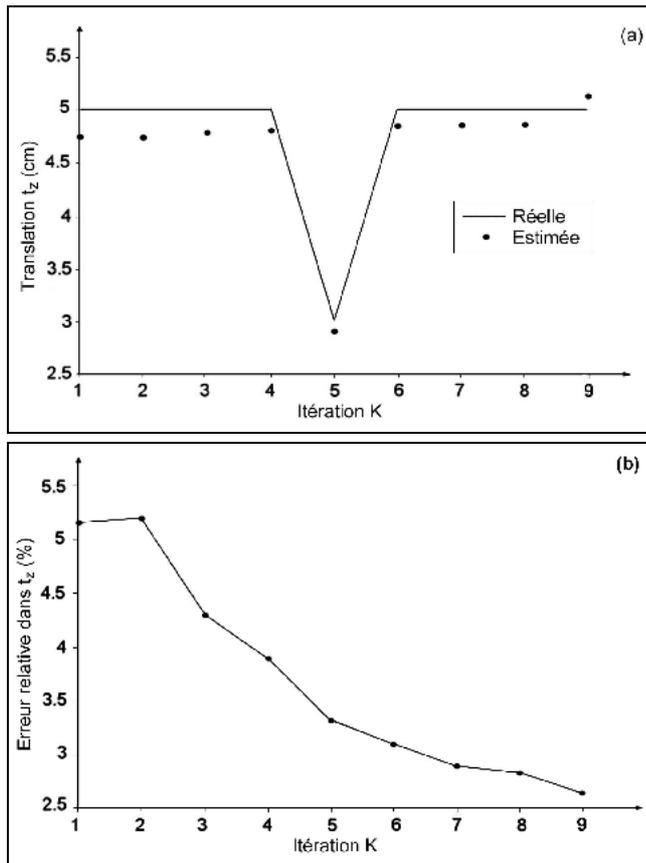
Pour la construction de la scène, nous avons opté pour les choix suivants:

- La distance du fond de la scène à la caméra est de 500 cm.
- Le point de la scène le plus proche à la caméra se situe à une distance de 455 cm.
- Le mouvement de la caméra, selon son axe  $x$ , est constant et de valeur 1 cm entre images.
- Les distances et les mouvement ont été choisis de façon à limiter les composantes du flot optique au-dessous de 5 pixels.

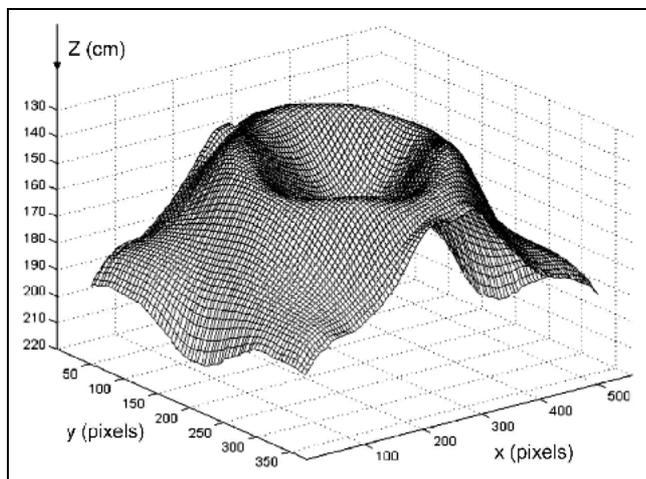
La scène est constituée de 8 images dont les 4 premières sont représentées dans la figure 13 ainsi que leurs champs du flot optique. Nous remarquons que les régions d'intensité uniforme (fond de la scène des images synthétiques) ne permettent pas une estimation correcte dans l'estimation du flot, ce qui engendre des erreurs dans l'estimation du mouvement (comme cela est montré dans la figure 14).

Cependant, il faut signaler que nous avons commencé notre algorithme avec une carte de profondeur plate d'une valeur moyenne de 477.5 cm.

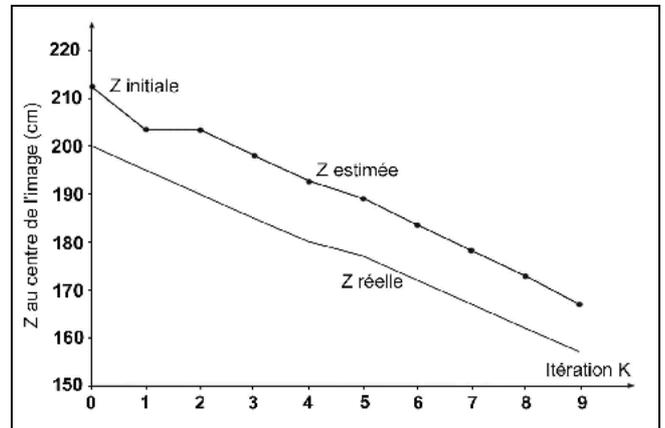
Enfin, la figure 15 montre la structure rétablie après traitement de la séquence en entier. En la comparant avec la structure initiale générée par le logiciel graphique, on note que seule une partie du vase (endroit brillant) n'est pas bien détectée.



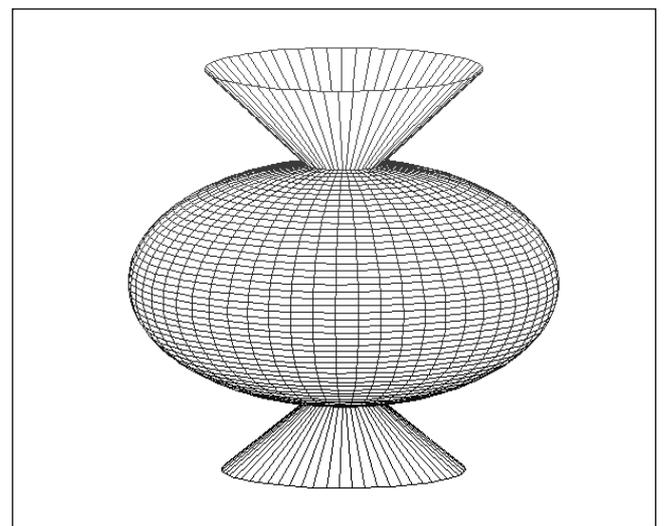
**Figure 9:** (a) Comparaison entre la composante du mouvement  $t_z$  réelle et estimée. (b) Erreur relative commise dans l'estimation de  $t_z$ .



**Figure 10:** La structure rétablie après traitement de la séquence "cratère" en entier.



**Figure 11:** Comparaison entre la profondeur estimée et réelle en un point de l'image.



**Figure 12:** La structure créée pour générer la séquence "vase".

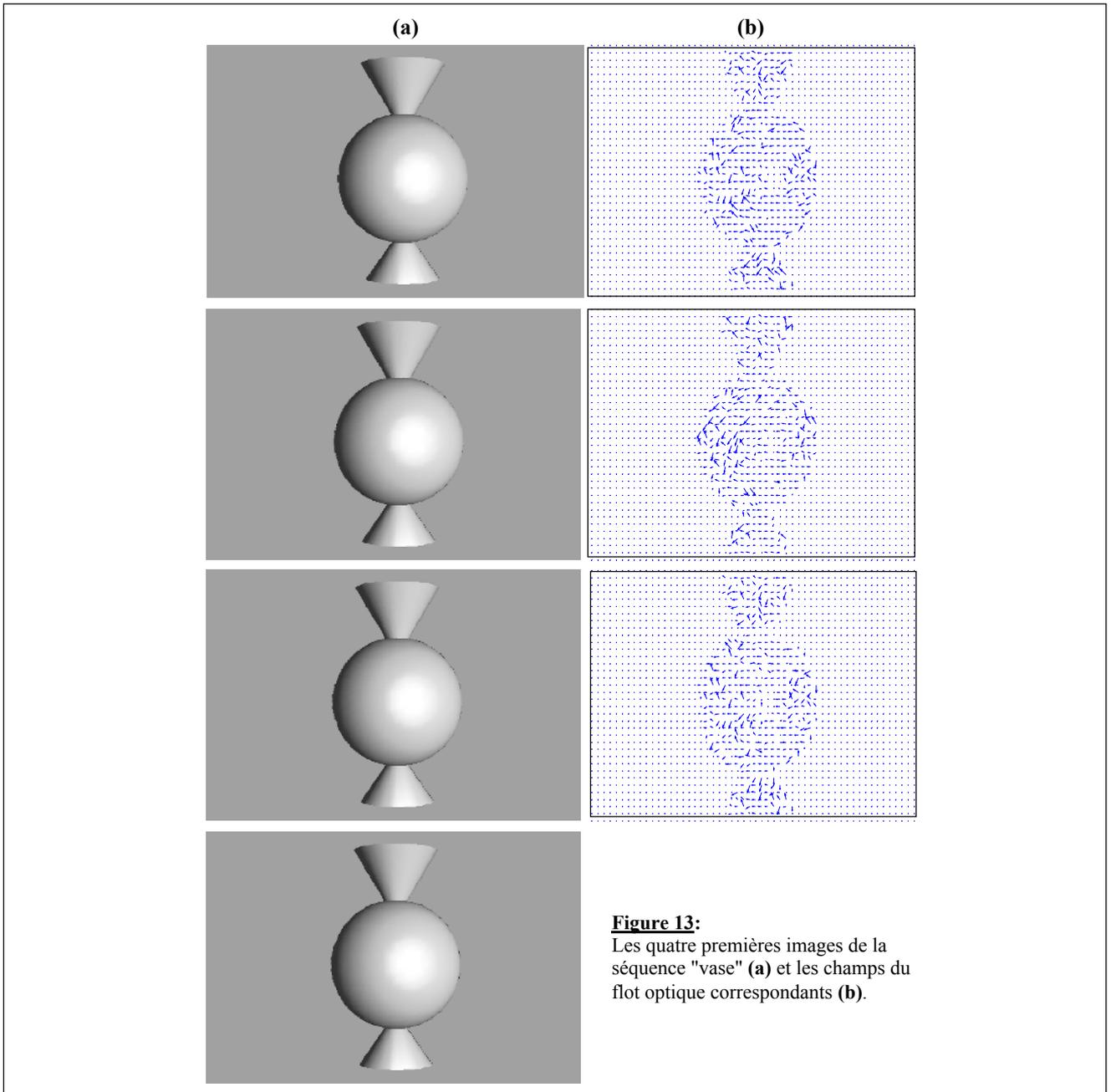
### 8.3- Séquence d'images réelles

La série suivante d'expériences est consacrée aux images réelles où une caméra CCD de type Olympus C-420L se déplace en translation pure au dessus d'une scène constituée d'une petite bouteille placée devant un fond plat.

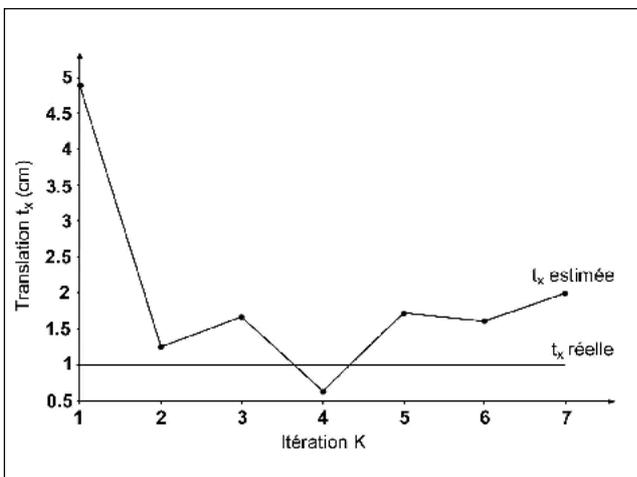
La séquence est seulement constituée de 4 images, en raison des difficultés rencontrées lors de la prise de vues, que ce soit en ce qui concerne les conditions d'illumination ou de déplacement de la caméra qui étaient mal adaptées pour une telle expérience, exigeant une grande précision. La séquence réelle "bouteille" a été prise avec les conditions suivantes:

- La bouteille est à 56 cm de distance de la caméra et le fond est à 71 cm.
- La caméra se déplace selon son axe  $x$  avec un mouvement non uniforme de 2 mm, pour générer les trois premières images, puis de 4 mm pour la quatrième.
- Les mouvements sont choisis de telle façon à limiter le flot à un maximum de 5 pixels.

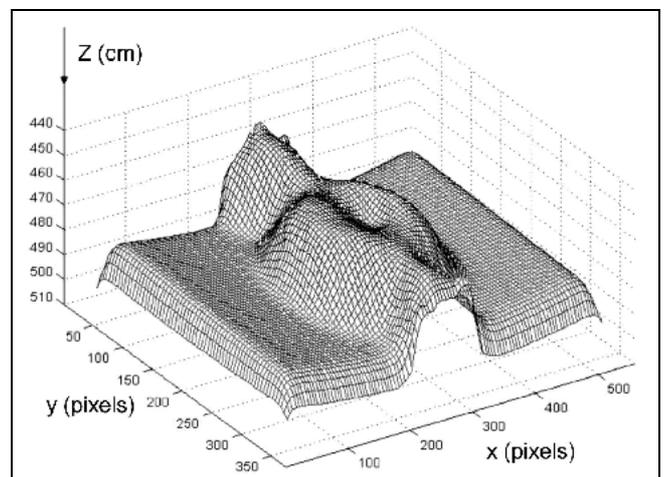
La figure 16 montre les quatre images constituant la séquence réelle "bouteille" et les trois champs du flot optique correspondants.



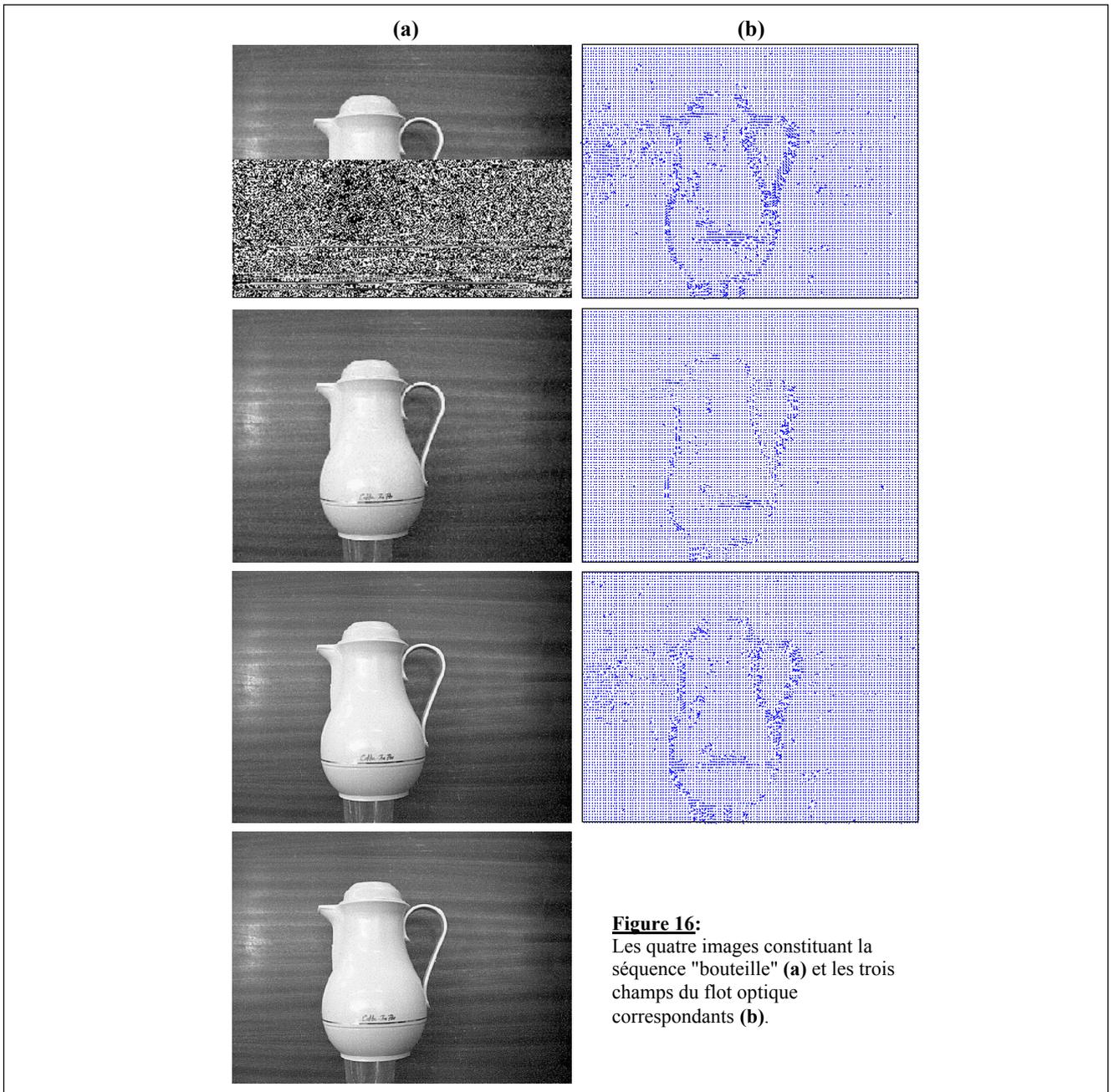
**Figure 13:** Les quatre premières images de la séquence "vase" (a) et les champs de flot optique correspondants (b).



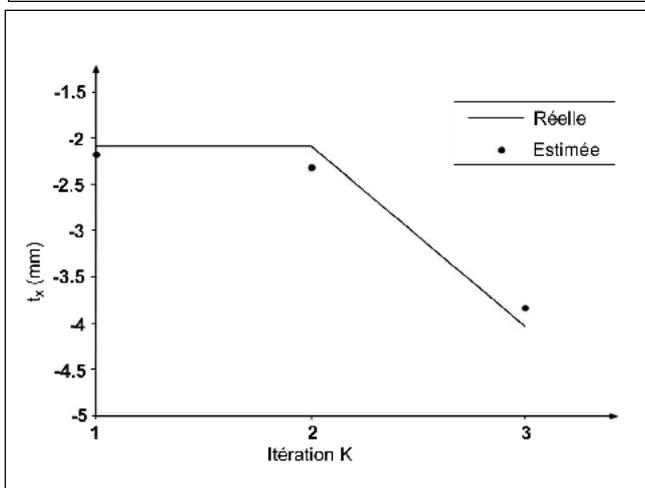
**Figure 14:** Comparaison entre la composante du mouvement  $t_x$  réelle et estimée.



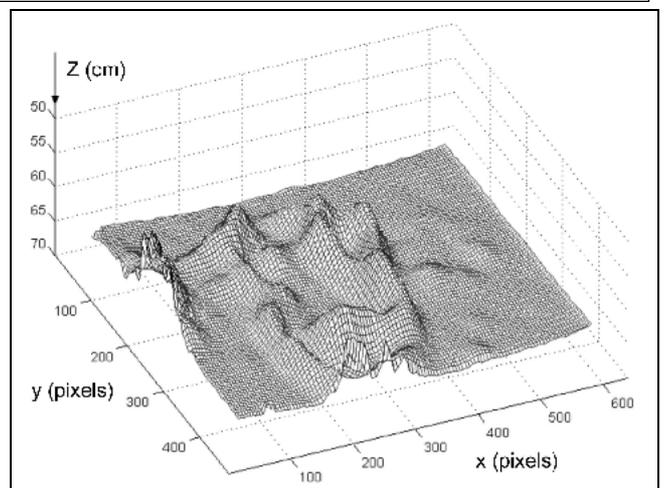
**Figure 15:** La structure rétablie après traitement de la séquence "vase" en entier.



**Figure 16:** Les quatre images constituant la séquence "bouteille" (a) et les trois champs du flot optique correspondants (b).



**Figure 17:** Comparaison entre la composante de translation  $t_x$  réelle et estimée.



**Figure 18:** La structure rétablie après traitement de la séquence "bouteille".

La profondeur est initialisée avec une carte plate de valeur moyenne 63.5 cm.

Bien que le flot optique mesuré (Fig. 16.b) apparaisse comme trompeur (flot nul pour les régions d'intensité uniforme), le mouvement est récupéré avec succès (Fig.17), ce qui confirme à nouveau de la robustesse du procédé d'estimation du mouvement.

Enfin, et après trois itérations du filtre, la structure rétablie est montrée dans la figure 18. En la comparant avec la structure réelle, on remarque que seule la frontière est restituée. Cela est dû au nombre insuffisant d'images constituant la séquence.

## 9- TEMPS D'EXECUTION

L'exécution est raisonnablement rapide: pour une taille d'image synthétique de  $375 \times 560$  pixels, une itération prend approximativement 1 minute sur un Pentium 200 MHz, en supposant que le flot optique est déjà calculé. Notons, cependant, qu'en raison de la nature locale des filtres de Kalman, cet algorithme est idéal pour une exécution parallèle.

## CONCLUSION

Dans cet article, nous avons proposé une solution à deux problèmes fondamentaux dans la vision par ordinateur; en l'occurrence le rétablissement de la structure et l'estimation du mouvement 3D à partir de séquences d'images 2D. En se basant sur la théorie des systèmes dynamiques, nous avons conçu un filtre de Kalman pour le rétablissement de la structure de l'environnement sous forme de cartes de profondeur. Les performances du système ainsi proposé sont fortement dépendantes de la qualité du flot optique et de la variance associée qui en constituent les données d'entrée.

La caractéristique essentiel de notre approche est qu'elle n'est pas limitée à l'estimation de la profondeur et du mouvement, mais elle peut être étendue à d'autres problèmes de la vision dans lesquels l'information est acquise à travers des séquences d'images telles que la segmentation, la couleur, la texture, etc.

## REFERENCES

- [1]- Ayache N., Faugeras O., "Building, registering and fusing noisy visual Maps", *The International Journal of Robotics Research*, vol. 7, N° 6, December (1988).
- [2]- Koch R., "Automatic reconstruction of buildings from stereoscopic image sequences", *Eurographics'93*, vol. 12, N° 3, (1993).
- [3]- Aggarwal J.K., Nadhakaran N., "On the computation of motion from sequences of images", *In Proc. IEEE*, vol. 76, N° 8, August (1988), pp. 917-935.
- [4]- Matthies L., Szeliski R., Kanade R., "Kalman filter-based algorithms for estimating depth from image sequence", *In Proc. DARPA Image Understanding Workshop*, Cambridge, MA, April (1988).
- [5]- Zhang Z. , "A Stereovision System for a Planetary Rover: Calibration, Correlation, Registration, and Fusion", *In: Proceedings IEEE Workshop on Planetary Rover Technology and Systems*, Minneapolis, Minnesota, USA, avril (1996).
- [6]- Henriksen L. and Krotkov E., "Natural terrain hazard detection with a laser rangefinder", *In: IEEE International Conference on Robotics and Automation*, Albuquerque, New Mexico (USA), April (1997), pp. 968-973.
- [7]- Boumaza R., "Perception multisensorielle pour la reconnaissance d'objets tridimensionnelles", Thèse de l'Université de Paul Sabatier de Toulouse, Juillet (1994).
- [8]- Heel J., "Dynamic motion vision", Technical Report 2139, MIT Artificial Intelligence Laboratory, (1990).
- [9]- Boubidi A., "Estimation des cartes de profondeur à partir d'une caméra mobile", Mémoire de Magister du C.U-Guelma, soutenu le 27 septembre (2000). □