

Calcul d'un estimé de description minimale du mouvement et du relief dans une scène à partir d'une séquence d'images

S. Hadjres et A. Mitiche

INRS-Télécommunications

16 place du commerce, Ile-des-soeurs

Québec, Canada, H3E 1H6

Résumé

Cet article décrit une méthode directe d'estimation du mouvement et du relief dans une scène à partir d'une séquence d'images de cette scène. La méthode est différentielle et se fonde sur le principe de longueur de description minimale (minimum description length: MDL) pour réduire le problème à la minimisation d'une fonction de coût de description de l'estimé, c'est à dire, du mouvement considéré rigide et du relief. La méthode est directe parce qu'elle ne nécessite pas le calcul d'un estimé du flot optique au préalable. La minimisation se fait par continuation et résulte en un système d'équations résolu par la méthode des itérations de Jacobi.

Mots clés:

mouvement, profondeur, longueur de description minimale, estimation directe du mouvement.

1 Introduction

Le calcul du mouvement et du relief d'une scène par un système de vision est l'objet de recherches accrues. La plupart des méthodes proposées calculent d'abord un estimé du flot optique avant d'utiliser celui-ci pour obtenir un estimé du mouvement et du relief de la scène observée. Une revue des méthodes d'estimation du flot optique peut être trouvée dans [10].

Introduite par Aloimonos et Brown [1], une alternative à ces méthodes consiste à déterminer directement le mouvement 3D et la structure des objets sans passer par une étape intermédiaire de calcul du flot optique [7], [6], [9], [11].

Dans cet article, on se propose de développer une méthode directe d'estimation du mouvement et du relief à partir d'une séquence d'images. C'est une approche différentielle qui se fonde sur le principe de longueur de description minimale pour réduire le problème à la minimisation d'une fonction de coût de description de l'estimé (mouvements dans la scène considérés rigides et relief).

Dans la section qui suit, on pose le problème et on le formule en se fondant sur le principe de longueur de description minimale. Les différents modèles utilisés pour écrire la fonction objectif sont décrits. A la section trois, on traite du problème d'optimisation de cette fonction objectif. La section 4 est consacrée à l'implantation de l'algorithme et à la présentation des résultats. La dernière section conclut l'article.

2 Formulation du problème

2.1 Structure et mouvement 3D

Soit $P(X, Y, Z)$ un point d'un objet se déplaçant dans l'espace, et $p(x, y, f)$ sa projection centrale sur le plan image (figure 1). Les relations de projection sont:

$$\begin{cases} x = f \frac{X}{Z} \\ y = f \frac{Y}{Z} \end{cases} \quad (1)$$

où f est la distance focale.

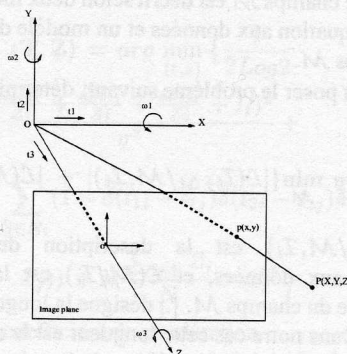


Figure 1: Modèle géométrique d'observation: projection centrale d'objets en mouvement rigide

En supposant que le mouvement dans l'espace est rigide, on le caractérise par son torseur cinématique (Ω, T) où $\vec{\Omega} = (\omega_1, \omega_2, \omega_3)$ représente le vecteur vitesse de rotation autour d'un axe passant par l'origine, et $\vec{T} = (t_1, t_2, t_3)$ représente le vecteur vitesse de translation.

Le vecteur vitesse \vec{V} de P est donné par:

$$\vec{V} = \vec{T} + \vec{\Omega} \times \vec{OP} \quad (2)$$

O étant l'origine du système de référence.

Si $\vec{V} = [\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}]$, on a:

$$\begin{cases} \frac{dX}{dt} = t_1 + \omega_2 Z - \omega_3 Y \\ \frac{dY}{dt} = t_2 - \omega_1 Z + \omega_3 X \\ \frac{dZ}{dt} = t_3 + \omega_1 Y - \omega_2 X \end{cases} \quad (3)$$

En dérivant (1) on obtient:

$$\begin{cases} \frac{dx}{dt} = \frac{f}{Z} \frac{dX}{dt} - \frac{x}{Z} \frac{dZ}{dt} \\ \frac{dy}{dt} = \frac{f}{Z} \frac{dY}{dt} - \frac{y}{Z} \frac{dZ}{dt} \end{cases} \quad (4)$$

En combinant (3) et (4) on obtient l'expression du mouvement 2D, (u, v) , en fonction du torseur cinématique et de la profondeur:

$$\begin{cases} \frac{dx}{dt} = u = \frac{-xy}{f} \omega_1 + \frac{f^2+x^2}{f} \omega_2 - y\omega_2 + \frac{f}{Z} t_1 - \frac{x}{Z} t_3 \\ \frac{dy}{dt} = v = \frac{xy}{f} \omega_2 - \frac{f^2+y^2}{f} \omega_2 + x\omega_3 + \frac{f}{Z} t_2 - \frac{y}{Z} t_3 \end{cases} \quad (5)$$

2.2 Formulation du problème

Etant donné deux images, \mathcal{I}_t au temps t et $\mathcal{I}_{t+\Delta t}$ au temps $t + \Delta t$, on recherche un champs $\mathcal{M} = (\Omega, T, Z)$ qui explique le passage d'une image à l'autre.

Par analogie à la formulation dans [4] pour l'estimation du flot optique, le champs \mathcal{M} est décrit selon deux modèles: un modèle d'adéquation aux données et un modèle de description du champs \mathcal{M} .

Ceci conduit à poser le problème suivant: déterminer $\hat{\mathcal{M}}$ tel que:

$$\hat{\mathcal{M}} = \arg \min_{\mathcal{M}} \{ |\mathcal{L}(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t)| + |\mathcal{L}(\mathcal{M}/\mathcal{I}_t)| \}$$

où $\mathcal{L}(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t)$ est la description du modèle d'adéquation aux données, et $\mathcal{L}(\mathcal{M}/\mathcal{I}_t)$ est la description du modèle du champs \mathcal{M} . $|\cdot|$ désigne la longueur d'une description. Dans notre cas cette longueur est le nombre de bits pour encoder la description. Il s'agit donc pour nous de déterminer \mathcal{M} de longueur de description minimale (MDL: minimum description length), c'est à dire de déterminer \mathcal{M} codé par un nombre minimum de bits.

2.3 Modèle d'adéquation aux données

On suppose que les changements d'intensité le long des trajectoires de mouvement sont dus au bruit d'observation, et que le mouvement entre deux images consécutives est un mouvement de translation. Sous ces deux hypothèses, on peut écrire:

$$I(x + u\Delta t, y + v\Delta t, t + \Delta t) = I(x, y, t) + \mu(x, y) \quad (6)$$

où $\mathbf{w} = (u, v)$, est le vecteur vitesse image au point (x, y) , Δt est l'intervalle de temps entre l'observation des deux images et μ désigne le bruit d'observation que l'on suppose Gaussien de moyenne nulle et de variance σ^2 et indépendant de \mathcal{I}_t . On peut donc écrire:

$$p(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t) = p(\mu = \mathcal{I}_{t+\Delta t}^w - \mathcal{I}_t) \quad (7)$$

où $\mathcal{I}_{t+\Delta t}^w(x, y)$ désigne $\mathcal{I}_{t+\Delta t}(x + u\Delta t, y + v\Delta t)$.

Le développement de Taylor du terme de gauche de l'équation (6) donne:

$$\begin{aligned} \mathcal{I}(x + u\Delta t, y + v\Delta t, t + \Delta t) \\ = \mathcal{I}(x, y, t) + I_x u\Delta t + I_y v\Delta t + I_t \Delta t + o(\Delta t) \end{aligned} \quad (8)$$

où I_x, I_y et I_t sont les dérivées partielles de l'image \mathcal{I}_t par rapport à x, y et t respectivement. Négligeant le terme $o(\Delta t)$, les expressions (6), (7) et (8) nous permettent d'écrire:

$$p(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t) \cong p(\mu_{x,y} = I_x u\Delta t + I_y v\Delta t + I_t \Delta t) \quad (9)$$

$\mu = I_x u\Delta t + I_y v\Delta t + I_t \Delta t$ est le résidu de Horn and Schunk. En substituant à u et v leur expression (5) en fonction de la profondeur Z et du torseur (Ω, T) et en posant $\Delta t = 1$ unité de temps, l'expression de ce résidu est:

$$a_1 \omega_1 + a_2 \omega_2 + a_3 \omega_3 + a_4 \frac{t_1}{Z} + a_5 \frac{t_2}{Z} + a_6 \frac{t_3}{Z} + I_t \quad (10)$$

avec:

$$\begin{aligned} a_1 &= \frac{-xyI_x - I_y(f^2 + y^2)}{f} \\ a_2 &= \frac{I_x(f^2 + x^2 + xyI_y)}{f} \\ a_3 &= -yI_x + xI_y \\ a_4 &= I_x f \\ a_5 &= I_y f \\ a_6 &= -I_x x - I_y y \end{aligned}$$

Avec des images discrètes et des valeurs de luminance quantifiées, le terme associé au modèle d'adéquation aux données devient, selon un codage d'entropie et tenant compte des hypothèses sur le bruit d'observation:

$$\begin{aligned}
& |\mathcal{L}(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t)| \\
&= -\text{Log}_2 p(\mathcal{I}_{t+\Delta t}/\mathcal{M}, \mathcal{I}_t) \\
&= -\text{Log}_2 p(a_1\omega_1 + a_2\omega_2 + a_3\omega_3 + a_4\frac{t_1}{Z} + a_5\frac{t_2}{Z} + a_6\frac{t_3}{Z} + I_t) \\
&= -\text{Log}_2 \prod_{i \in \Gamma} p(a_1\omega_{1i} + a_2\omega_{2i} + a_3\omega_{3i} + a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t) \\
&= -\sum_{i \in \Gamma} \text{Log}_2 p(a_1\omega_{1i} + a_2\omega_{2i} + a_3\omega_{3i} + a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t) \\
&= -\frac{N}{4\text{Log}_2} \text{Log}(2\pi\sigma^2) + \frac{1}{2\text{Log}_2} \\
&\sum_{i \in \Gamma} \frac{(a_1\omega_{1i} + a_2\omega_{2i} + a_3\omega_{3i} + a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t)^2}{\sigma^2} \quad (11)
\end{aligned}$$

où (w_{1i}, w_{2i}, w_{3i}) , (t_{1i}, t_{2i}, t_{3i}) et Z_i sont, respectivement, la vitesse de rotation, la vitesse de translation et la profondeur au point i ; Γ est le domaine discret de l'image et N est le nombre de points de l'image.

2.4 Modèle de description du champs de mouvement

Le champs de mouvement est posé constant par morceaux. Une région de mouvement constant peut être décrite par sa frontière et la valeur unique de son mouvement. Le nombre de bits nécessaire pour décrire chaque région est donc proportionnel au nombre d'éléments dans le code de chaîne de sa frontière plus une constante pour spécifier le nombre de bits pour représenter le premier élément de cette chaîne et la valeur des deux composantes du mouvement de cette région. Puisqu'une frontière est placée entre deux points de l'image où le mouvement ou la profondeur diffèrent, un bon estimé du nombre de bits nécessaire pour spécifier le champs de mouvement est donné par:

$$|\mathcal{L}(\mathcal{M})| = \frac{b}{2} \sum_{i \in \Gamma} \sum_{j \in N_i} (1 - \delta(w_{1i} - w_{1j})\delta(w_{2i} - w_{2j})\delta(w_{3i} - w_{3j})\delta(t_{1i} - t_{1j})\delta(t_{2i} - t_{2j})\delta(t_{3i} - t_{3j})\delta(Z_i - Z_j)) \quad (12)$$

où b est le coût moyen en bits d'un segment de code de chaîne, N_i est l'ensemble des 4 voisins du point i , et

$$\delta(x) = \begin{cases} 1 & \text{si } x = 0 \\ 0 & \text{sinon} \end{cases} \quad (13)$$

2.5 La fonction objectif

Considérant les expressions (11) et (12), le problème est de déterminer un estimé $\hat{\mathcal{M}}$ de \mathcal{M} tel que:

$$\begin{aligned}
\hat{\mathcal{M}} = (\hat{\Omega}, \hat{\mathbf{T}}, \hat{\mathbf{Z}}) &= \arg \min_{(\mathbf{w}, \mathbf{t}, \mathbf{z})} \left\{ \frac{1}{2\text{Log}_2} \right. \\
&\sum_{i \in \Gamma} \frac{(a_1\omega_{1i} + a_2\omega_{2i} + a_3\omega_{3i} + a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t)^2}{\sigma^2} + \\
&\frac{b}{2} \sum_{i \in \Gamma} \sum_{j \in N_i} (1 - \delta(w_{1i} - w_{1j})\delta(w_{2i} - w_{2j})\delta(w_{3i} - w_{3j}) \\
&\delta(t_{1i} - t_{1j})\delta(t_{2i} - t_{2j})\delta(t_{3i} - t_{3j})\delta(Z_i - Z_j)) \left. \right\} \quad (14)
\end{aligned}$$

Le premier terme est le coût (en terme de longueur de code de description) de description de l'adéquation aux données, et le second terme est le coût de la description du champs.

Sous l'hypothèse que le mouvement 3D est localement un mouvement de translation, (14) se réduit à:

$$\begin{aligned}
\hat{\mathcal{M}} = (\hat{\mathbf{T}}, \hat{\mathbf{Z}}) &= \arg \min_{(\mathbf{t}, \mathbf{z})} \left\{ \frac{1}{2\text{Log}_2} \right. \\
&\sum_{i \in \Gamma} \frac{(a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t)^2}{\sigma^2} + \\
&\frac{b}{2} \sum_{i \in \Gamma} \sum_{j \in N_i} (1 - \delta(t_{1i} - t_{1j})\delta(t_{2i} - t_{2j})\delta(t_{3i} - t_{3j}) \\
&\delta(Z_i - Z_j)) \left. \right\} \quad (15)
\end{aligned}$$

Remarquons que le terme de droite de (15) est homogène: si (t_1, t_2, t_3, Z) est une solution, $(\alpha t_1, \alpha t_2, \alpha t_3, \alpha Z)$ l'est aussi. Ceci traduit l'incertitude d'échelle: on ne peut récupérer que la profondeur relative des objets en mouvement. On fixe cette échelle en donnant à la moyenne de l'estimé des profondeurs une valeur arbitraire: $\frac{\sum_{i \in \Gamma} Z_i}{N} = c$. Ceci nous amène à:

$$\begin{aligned}
\hat{\mathcal{M}} = (\hat{\mathbf{T}}, \hat{\mathbf{Z}}) &= \arg \min_{(\mathbf{t}, \mathbf{z})} \left\{ \frac{1}{2\text{Log}_2} \right. \\
&\sum_{i \in \Gamma} \frac{(a_4\frac{t_{1i}}{Z_i} + a_5\frac{t_{2i}}{Z_i} + a_6\frac{t_{3i}}{Z_i} + I_t)^2}{\sigma^2} + \\
&\frac{b}{2} \sum_{i \in \Gamma} \sum_{j \in N_i} (1 - \delta(t_{1i} - t_{1j})\delta(t_{2i} - t_{2j})\delta(t_{3i} - t_{3j}) \\
&\delta(Z_i - Z_j)) + \lambda \left(\frac{\sum_{i \in \Gamma} Z_i}{N} - c \right)^2 \left. \right\} \quad (16)
\end{aligned}$$

où λ est un coefficient de pondération.

Nous résolvons ce problème de minimisation par continuation [2], une méthode qui consiste à inscrire le problème dans une famille de problèmes dont le paramètre est l'écart standard d'une Gaussienne approximant la fonction δ (13).

3 Optimisation: méthode de continuation

La méthode de continuation consiste à plonger la fonction objectif $F(\mathbf{T}, \mathbf{Z})$ dans une famille de fonctions $F(\mathbf{T}, \mathbf{Z}, s)$ paramétrée par s , pour laquelle il existe un seul minimum pour une valeur suffisamment grande de s , et pour laquelle le nombre et la position des minimums locaux convergent vers ceux de la fonction objectif originale $F(\mathbf{T}, \mathbf{Z})$ quand s tend vers zéro.

Après avoir fixé s à une valeur suffisamment grande, le minimum local de la fonction $F(\mathbf{T}, \mathbf{Z}, s)$ est calculé en utilisant un des algorithmes de descente connus, puis s est diminué et le minimum recalculé. Cette procédure est répétée jusqu'à ce que le s atteigne une valeur suffisamment faible.

Dans notre cas, la famille de fonctions est obtenue en remplaçant $\delta(t_{1i} - t_{1j})\delta(t_{2i} - t_{2j})\delta(t_{3i} - t_{3j})\delta(Z_i - Z_j)$ par:

$$e_{ij} = e^{-\frac{(t_{1i} - t_{1j})^2 - (t_{2i} - t_{2j})^2 - (t_{3i} - t_{3j})^2 - (z_i - z_j)^2}{(s\sigma)^2}}$$

Pour un s donné la fonction objectif est donc:

$$F(\mathbf{T}, \mathbf{Z}, s) = \left\{ \frac{1}{2L\log 2} \sum_{i \in \Gamma} \frac{(a_4 \frac{t_{1i}}{Z_i} + a_5 \frac{t_{2i}}{Z_i} + a_6 \frac{t_{3i}}{Z_i} + I_i)^2}{\sigma^2} + \frac{b}{2} \sum_{i \in \Gamma} \sum_{j \in N_i} (1 - e_{ij}) + \lambda \left(\frac{\sum_{i \in \Gamma} Z_i}{N} - c \right)^2 \right\}$$

Le minimum de cette fonction est atteint quand ses dérivées par rapport à $\omega_{1i}, \omega_{2i}, \omega_{3i}, t_{1i}, t_{2i}, t_{3i}$ et Z_i sont nulles. Donc, en écrivant ces dérivées et en les égalisant à zéro, on obtient les équations de Jacobi suivantes:

$$\left\{ \begin{aligned} t_{1i}^{n+1} &= \frac{-a_4(a_5 t_{2i}^n + a_6 t_{3i}^n + I_i Z_i^n) + \frac{b L \log 2}{s^2} \sum_{j \in N_i} t_{1j}^n e_{ij}^n}{a_4^2 + \frac{b L \log 2}{s^2} \sum_{j \in N_i} e_{ij}^n} \\ t_{2i}^{n+1} &= \frac{-a_5(a_4 t_{1i}^n + a_6 t_{3i}^n + I_i Z_i^n) + \frac{b L \log 2}{s^2} \sum_{j \in N_i} t_{2j}^n e_{ij}^n}{a_5^2 + \frac{b L \log 2}{s^2} \sum_{j \in N_i} e_{ij}^n} \\ t_{3i}^{n+1} &= \frac{-a_6(a_4 t_{1i}^n + a_5 t_{2i}^n + I_i Z_i^n) + \frac{b L \log 2}{s^2} \sum_{j \in N_i} t_{3j}^n e_{ij}^n}{a_6^2 + \frac{b L \log 2}{s^2} \sum_{j \in N_i} e_{ij}^n} \\ Z_i^{n+1} &= \frac{-I_i(a_4 t_{1i}^n + a_5 t_{2i}^n + a_6 t_{3i}^n) + \frac{b L \log 2}{s^2} \sum_{j \in N_i} Z_j^n e_{ij}^n}{I_i^2 + \frac{b L \log 2}{s^2} \sum_{j \in N_i} e_{ij}^n + \frac{2\sigma^2 \lambda L \log 2}{N^2}} \\ &\quad \frac{2\sigma^2 \lambda L \log 2 \left(\frac{\sum_{i \in \Gamma} Z_i^n}{N} - c \right)}{I_i^2 + \frac{b L \log 2}{s^2} \sum_{j \in N_i} e_{ij}^n + \frac{2\sigma^2 \lambda L \log 2}{N^2}} \end{aligned} \right.$$

n étant l'indice des itérations.

4 Implantation et discussion

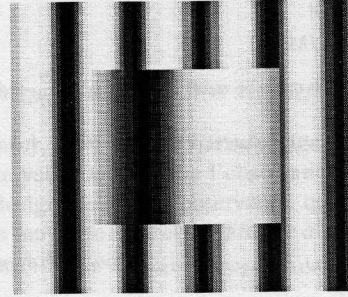


Figure 2: Séquence sinusoïde (image 1)

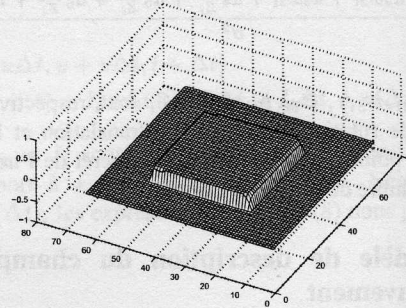


Figure 3: profondeur obtenue pour la séquence sinusoïde

L'algorithme développé a été testé sur des séquences d'images monoscopiques. La première, la séquence *sinusoïde*, (figure 2) est une séquence d'images synthétiques de 64×64 pixels, et est composée d'un fond fixe et d'un objet en premier-plan se déplaçant d'un pixel vers la droite. Cet exemple simule, en accord avec notre modèle de prise de vue, le mouvement d'un objet à 35cm devant un arrière-plan à 65cm. Le résultat après 15 itérations de s , avec pour chacune d'elles deux itérations de Jacobi, est montré à la figure (3). Dans cet exemple le paramètre s a été initialisé à 0.1 et sigma à 1. Le relief a été fidèlement reproduit à part à quelques points dans les zones éparées de discontinuité de mouvement. Le mouvement 2D reconstitué à partir des paramètres 3D obtenus est aussi fidèlement reproduit sauf en quelques points se trouvant sur les discontinuités de mouvement. La composante horizontale u est égale à 1 pixel et la composante verticale v est nulle. Ce résultat est présenté dans la figure 4. La deuxième séquence test est la séquence *papillon* (figure 5) de dimension 284×268 . Elle

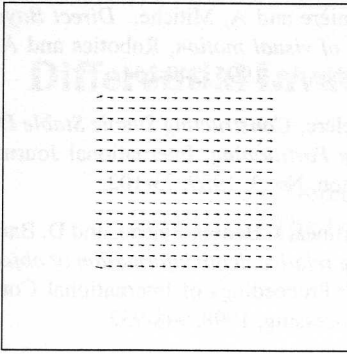


Figure 4: mouvement image reconstitué pour la séquence sinusoïde

a été obtenue à partir d'une image réelle dans laquelle on a déplacé une région rectangulaire d'un pixel vers la droite. Le résultat obtenu pour cette séquence est montré à la figure 6. Il a été obtenu avec le paramètre s initialisé à 1. et σ à 4. Ce résultat a été obtenu après 50 itérations et 2 itérations de Jacobi pour chaque s . Le relief est reproduit fidèlement. Le mouvement 2D reconstitué à partir des paramètres 3D obtenus est présenté à la figure 7. Dans ce cas aussi, le mouvement 2D calculé est reproduit fidèlement (u est égale à 1 pixel et v est nul).

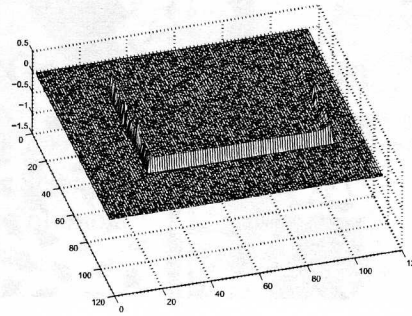


Figure 6: profondeur obtenue pour la séquence papillon

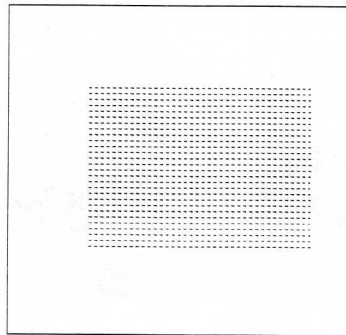


Figure 7: mouvement image reconstitué pour la séquence papillon

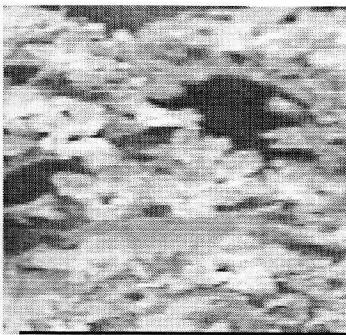


Figure 5: Séquence papillon (image 1)

Une troisième séquence test, la séquence *canette*, a été utilisée (figure 8), c'est une séquence réelle de dimension 320×200 comportant une canette de Coke et une boîte cubique disposées devant une affiche murale. Le mouvement de la caméra dans cette séquence est un mouvement de translation latéral. la profondeur obtenue pour les deux premières images de la séquence est montrée à la figure 9. Ce résultat a été obtenu après 200 itérations de s et 10 itérations de Jacobi pour chacune d'elles.

5 Conclusion

Dans cet article, une approche directe d'estimation du mouvement 3D et du relief à partir d'une séquence d'images a été présentée. L'utilisation du principe de la longueur de description minimale a permis une formulation simple du problème d'estimation. Pour valider l'algorithme, nous l'avons appliqué à plusieurs séquences d'images monoscopiques. Pour chacune de ces séquences, en utilisant la profondeur obtenue pour une image donnée, on a pu créer l'image stéréoscopique lui correspondant. La visualisation des paires d'images stéréoscopiques ainsi obtenues, nous a montré que l'algorithme développé est très promettant.

References

- [1] J. Aloimonos and C. M. Brown, *Direct Processing of Curvilinear Sensor Motion From a Sequence of Perspective Images*, Workshop on Computer Vision: Representation and Control, Annapolis, MD 1984, 72-77.
- [2] G. Dahlquist and A. Björck, *Numerical methods*, N. Anderson (trans.), Prentice-Hall, inc. Englewood Cliffs,

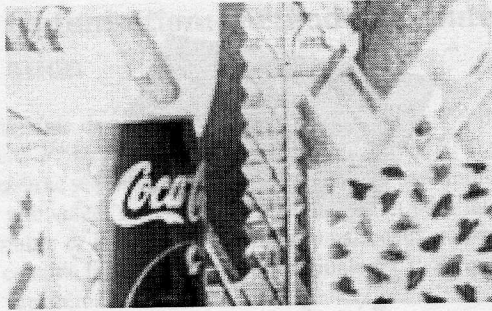


Figure 8: Séquence canette (image 1)

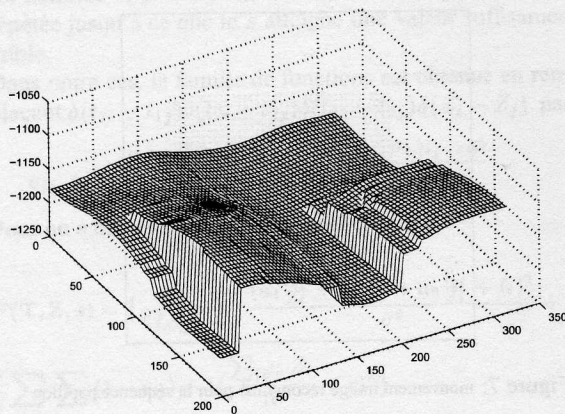


Figure 9: profondeur obtenue pour la séquence canette

New Jersey, 1974.

- [3] H. Gu, Y. Shirai, and M. Asada, *MDL-Based Segmentation and Motion Modeling in a long Image Sequence of scene with multiple independently Moving objects*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18, No. 1, January 1996, 58-63.
- [4] S. Hadjres, A. Mitiche, and A. Mansouri, *Computation of an Optical Flow Estimate of Minimum Description Length*, the 1999 International Conference on Imaging Science, Systems, and Technology (CISST'1999), Las Vegas, Nevada, June 1999, 411-415.
- [5] B. Horn and B. Schunck, *Determining Optical Flow*, Artificial Intelligence, Vol. 17, 1981, 185-203.
- [6] Y. S. Hung and H. T. Ho, *A Kalman Filter Approach to Direct Depth Estimation Incorporating surface Structure*, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 6, June 1999, 570-575.
- [7] R. Laganière and A. Mitiche, *Direct Bayesian interpretation of visual motion*, Robotics and Autonomous Systems, No. 14, 1995, 247-254.
- [8] Y. G. Leclerc, *Constructing Simple Stable Descriptions for Image Partitioning*, International Journal of Computer Vision, No. 3, 1989, 73-102.
- [9] F. X. Martinez, J. Benois-Pineau, and D. Barba, *Extraction of the relative depth information of objects in video sequences*, Proceedings of International Conference on Image Processing, 1998, 948-952.
- [10] A. Mitiche and P. Boutheymy, *Computation and Analysis of Image Motion: A synopsis of Current Problems and Methods*, International Journal of Computer Vision, 19(1) 1996, 29-55.
- [11] F. Morier, H. Nicholas, J. Benois, D. Barba, H. Sanson, *Relative Depth Estimation of Video Objects for Image Interpolation*, Proceedings of International Conference on Image Processing, 1998, 953-957
- [12] P. C. Wareham and S. D. Blostein, *Region-oriented video coding using the MDL Principle and Quad-tree Optimisation*, Proceedings of International Conference on Image Processing, vol. 1, 1996, 665-668.
- [13] H. Zheng and S. D. Blostein, *Motion-based Object Segmentation and Estimation Using the MDL principle*, IEEE Transactions on Image Processing, vol. 4, No 9, sept 1995, 1223-1235.