

Détermination de l'orientation d'un plan texturé à partir du calcul des échelles locales

Laurent BOUTTÉ¹, Justin PLANTIER², Sylvie LELANDAIS¹

¹CEMIF - Groupe Système Complexes, Université d'Évry Val d'Essonne
40 rue du PELVOUX, CE 1455 COURCOURONNES, 91020 ÉVRY CEDEX, France

²IMASSA – Département Sciences Cognitives
BP 73 91223 BRÉTIGNY sur ORGE, France

boutte@cemif.univ-evry.fr, lboutte@imassa.fr,
jplantier@imassa.fr, s.lelandais@iut.univ-evry.fr

Résumé – Dans cet article, nous proposons une méthode de « Shape from Texture » basée sur une nouvelle technique d'extraction des fréquences locales. La méthode que nous utilisons ici est une amélioration de celle introduite par Lu et co. [3]. Elle consiste à faire une interpolation des fréquences locales, contenues dans une image d'un plan incliné texturé, pour retrouver l'orientation de ce plan en utilisant le modèle de projection en perspective. La méthode d'extraction des fréquences locales dans l'image est obtenue par une décomposition en ondelettes. La détermination de la fréquence locale de chaque point de l'image est réalisée à l'aide d'une interpolation de la réponse des ondelettes calculées à différentes échelles.

Abstract – In this article, we propose a "Shape From Texture" method based on a new technique of local frequency extraction. The method used is an improvement of the one introduced by Lu and al [3]. It consists in making a local frequency interpolation, of a textured slope plan's picture, to recover its orientation under the perspective projection model. The method of local frequency extraction is based on a wavelet decomposition. The determination of the local frequency of each point of the picture is achieved with an interpolation of wavelets' values.

1. Introduction

Le problème du calcul de l'orientation d'un plan incliné texturé peut être étudié de plusieurs façons. Les méthodes développées jusqu'ici se distinguent par les hypothèses retenues pour effectuer ce travail. La première différence qui peut être faite repose sur le modèle de la texture étudiée. Les motifs élémentaires qui composent la texture peuvent être pris avec une surface constante [7]. Les textures peuvent aussi être vues de telle sorte que la répartition de leurs motifs élémentaires soit uniforme sur la surface du plan incliné [1], ou la texture est considérée comme isotrope [7], ou encore il est possible de supposer que la texture ne possède qu'une seule fréquence spatiale dominante [3]. La deuxième différence qui existe entre ces différentes méthodes est le modèle de caméra utilisé. Certains utilisent une hypothèse de projection en perspective [2], [3], alors que d'autres personnes utilisent des hypothèses simplificatrices comme la paraperspective [1], ou bien la projection orthogonale [6].

Dans cette étude, les hypothèses de perspective et de textures ne possédant qu'une seule fréquence spatiale dominante vont être utilisées. Le travail présenté ici montre une méthode permettant de retrouver l'orientation d'un plan incliné texturé à partir de la variation des échelles locales dans une image, l'échelle locale correspondant à l'inverse de la fréquence locale. La technique proposée se distingue surtout par le fait d'utiliser une nouvelle méthode d'extraction

des échelles locales. Les différentes étapes de la méthode consistent en la détermination des échelles locales et au calcul des angles caractérisant l'orientation du plan dans un espace tridimensionnel.

2. Détermination de l'échelle locale

La convolution d'une image avec une ondelette permet de faire un filtrage fréquentiel de celle-ci. Il est ainsi possible de séparer l'ensemble des fréquences présentes dans une image à partir d'un ensemble d'ondelettes couvrant la totalité du spectre présent dans l'image. L'avantage de l'analyse en ondelettes par rapport à une analyse de Fourier, est de conserver une information spatiale. En effet, le résultat de cette convolution donne la répartition de la fréquence filtrée dans l'image (plus la réponse de l'ondelette est importante plus la fréquence présente autour du point considéré est proche de la fréquence filtrée par l'ondelette). Pour notre étude, les images seront décomposées sur plusieurs échelles à l'aide d'une transformée en ondelettes utilisant des filtres de type DOG (Difference of Gaussians) [4].

Ainsi, nous pouvons définir l'échelle locale d'un point de l'image comme la valeur de l'échelle de l'ondelette ayant conduit à l'amplitude maximum en ce point.

Des travaux antérieurs [5] utilisaient la notion de maximum de la fréquence locale.

Le principal problème de cette méthode est de créer un phénomène d'« escalier » dans l'image. En effet, le nombre de fréquences présent dans l'image résultat est égal au nombre d'ondelettes utilisé pour la décomposition (figure 1a). Pour améliorer ces résultats, nous avons utilisé une méthode basée sur le principe de l'interpolation parabolique autour de ce maximum (figure 1b). L'hypothèse de base est qu'en chaque point d'une image la réponse des différentes ondelettes suit une courbe continue. Pour une question de simplicité et une plus grande rapidité d'exécution, le calcul d'interpolation est restreint au maximum de cette courbe et à ses deux voisins. Cette technique permet d'interpoler les fréquences comprises entre deux ondelettes.

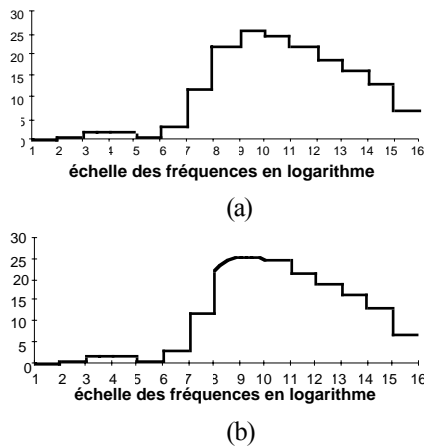
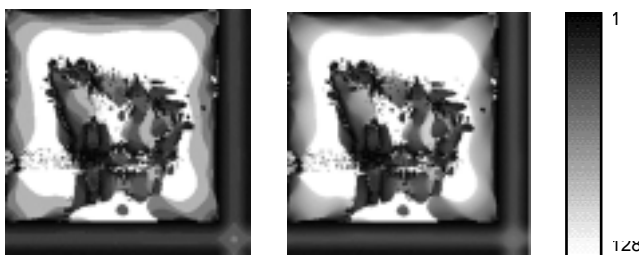


FIG. 1 : Courbes de réponse des ondelette en un point de l'image 1a : courbe avant interpolation ; 1b : courbe après interpolation.

La figure 2 montre la carte des échelles locales sur une image complexe (2a). En 2b on a extrait l'échelle par la méthode directe du maximum et en 2c par la méthode d'interpolation parabolique. On constate aisément que cette deuxième technique induit un lissage de la carte en limitant les « effets d'escalier » visibles en 2b.



(a)



(b)

(c)

FIG. 2 : Détermination des échelles locales. 2a : image d'origine ; 2b : méthode du maximum ; 2c : méthode d'interpolation ; 2d : répartition des échelles.

3. Calcul de l'Orientation d'un plan texturé incliné

La projection d'un plan incliné texturé dans une image provoque la déformation de la texture. La texture peut être caractérisée par les formes élémentaires qui la composent (texels) ou bien par ses fréquences locales. Si l'on choisit de décrire la texture avec les texels, on peut retrouver l'orientation de la surface grâce à la variation de répartition des texels dans l'image et à leurs déformations [1]. Dans la description fréquentielle, l'inclinaison du plan va se traduire par une variation de la fréquence locale. L'avantage de cette approche est que l'on obtient des informations en tous points de l'image, contrairement à la description par les texels.

Une étude théorique utilisant le système de représentation par les angles de slant et de tilt (figure 3) permet d'obtenir les équations de ces angles en fonction de la fréquence locale. Pour trouver ces équations, nous allons tout d'abord commencer par analyser un cas simple, c'est-à-dire un plan dont le tilt est nul.

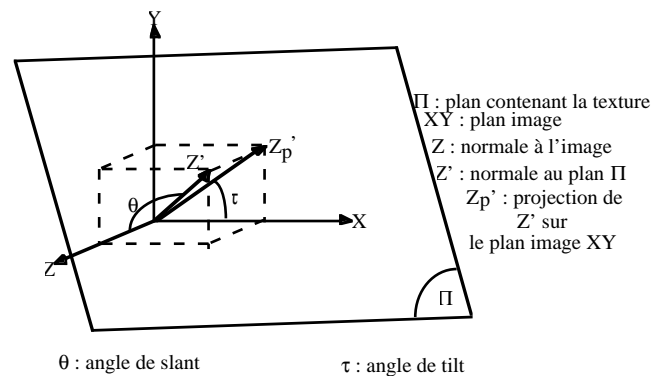


FIG. 3 : Représentation des angles de slant et de tilt

Soit une texture sinusoïdale parfaite (figure 4), dans le repère du plan incliné, l'équation de cette texture est:

$$S(x, y) = \cos(\omega_i x) \quad (1)$$

En prenant l'hypothèse de projection en perspective, l'équation de la texture dans le repère image devient:

$$S(x_i, y_i) = \cos(\gamma_i) = \cos \left(\omega_i \left[\frac{(f + Z)}{f} \right] \frac{1}{\cos \theta} \frac{x_i}{1 - \frac{x_i}{f} \tan \theta} \right) \quad (2)$$

où f est la distance focale de la caméra, Z la distance du plan incliné au plan image de la caméra et θ l'angle de slant.

La fréquence locale dans l'image étant définie comme la dérivée de γ_i par rapport à x_i , on obtient l'équation suivante:

$$F_i = \frac{\omega_i}{2\pi} \frac{(f+Z)}{f} \frac{1}{\cos\theta} \frac{1}{\left(1 - \frac{x_i}{f} \tan\theta\right)^2} \quad (3)$$

L'échelle locale étant l'inverse de la fréquence, il suit :

$$E_i(x_i, y_i) = \frac{2\pi}{\omega_i} \left(\frac{f}{f+Z}\right) \cos\theta \left(1 - \frac{x_i}{f} \tan\theta\right)^2 \quad (4)$$

$$= \beta_2 x_i^2 + \beta_1 x_i + \beta_0$$

avec

$$\begin{aligned} \beta_0 &= \frac{2\pi}{\omega_i} \left(\frac{f}{f+Z}\right) \cos\theta \\ \beta_1 &= \beta_0 \frac{-2}{f} \tan\theta \\ \beta_2 &= \beta_0 \frac{1}{f^2} \tan^2\theta \end{aligned} \quad (5)$$

A partir de cette équation, l'angle de slant peut être trouvé avec la relation suivante :

$$\theta = \arctan\left(-2 \frac{f\beta_2}{\beta_1}\right) \quad (6)$$

En généralisant ce résultat pour l'ensemble des angles de tilt possible, l'équation (4) devient:

$$E_i(x_i, y_i) = a_1 x_i^2 + a_2 x_i y_i + a_3 y_i^2 + a_4 x_i + a_5 y_i + a_6 \quad (7)$$

avec

$$\begin{aligned} a_1 &= \beta_2 \cos^2 \tau & a_2 &= \beta_2 \cos\tau \sin\tau & a_3 &= \beta_2 \sin^2 \tau \\ a_4 &= \beta_1 \cos\tau & a_5 &= \beta_1 \sin\tau & a_6 &= \beta_0 \end{aligned} \quad (8)$$

où τ représente l'angle de tilt.

A partir de cette équation, l'angle de tilt et les deux variables β_1 et β_2 peuvent être trouvés de la façon suivante :

$$\tau = \arctan\left(\frac{a_2}{2a_1}\right) \text{ ou } \tau = \arctan\left(\frac{2a_3}{a_2}\right) \text{ ou } \tau = \arctan\left(\frac{a_5}{a_4}\right) \quad (9)$$

$$\beta_1 = \frac{a_4}{\cos\tau} \quad \text{et} \quad \beta_2 = \frac{a_1}{\cos^2\tau} \quad (10)$$

4. Résultats expérimentaux

La méthode a d'abord été testée sur des images synthétiques dont on connaît la fréquence élémentaire et l'orientation. La figure 4a présente une de ces images. Elle est composée d'une sinusoïde projetée sur un plan dont l'inclinaison correspond à un angle de tilt de 50° et un angle de slant de 60° .

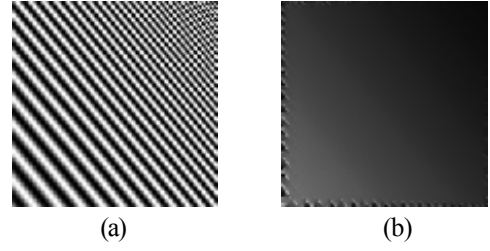


FIG. 4 : Image test. 4a : image synthétique ; 4b : carte des échelles locales.

Pour retrouver l'orientation des plans inclinés des images tests, on réalise la carte des échelles de chaque image. Notre but est ensuite faire une interpolation de cette carte par le cylindre parabolique dont nous venons de donner les équations (7 et 8).

Cette famille de courbes possède quatre inconnues: β_0 , β_1 , β_2 et τ . On veut trouver la courbe de cette famille qui se rapproche le plus de tous les points de la carte. Du fait des non-linéarités introduites par les cosinus, les sinus et les multiplications entre les différents paramètres, on ne peut trouver de solution simple avec une méthode d'interpolation de type des moindres carrés. On est donc obligé de faire l'interpolation avec la famille générale des courbes « x^2 », c'est-à-dire les courbes dont l'équation peut s'écrire de la manière suivante:

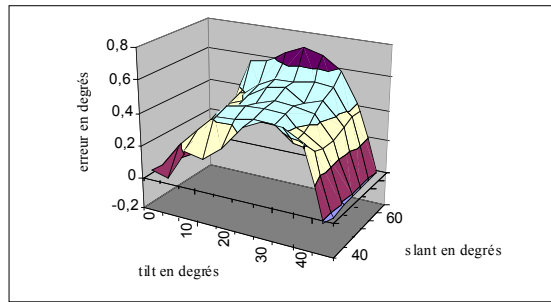
$$f(x, y) = ax^2 + bxy + cy^2 + dx + ey + f \quad (7)$$

En effet, cette famille de courbes peut être facilement interpolée par la méthode des moindres carrés, et elle inclut la famille de courbes que l'on souhaite trouver. On obtient donc six paramètres au lieu de quatre. Comme les cartes des échelles doivent en théorie représenter un paraboloidé, on peut supposer que l'interpolation faite à l'aide des courbes générales « x^2 » donnera une réponse très proche des valeurs souhaitées. En utilisant cette méthode générale d'interpolation, les erreurs trouvées sur l'angle de tilt sont assez faibles quelque soit l'orientation du plan incliné, contrairement aux erreurs faites sur l'angle de slant qui sont importantes d'autant plus que l'inclinaison réelle du plan incliné est faible.

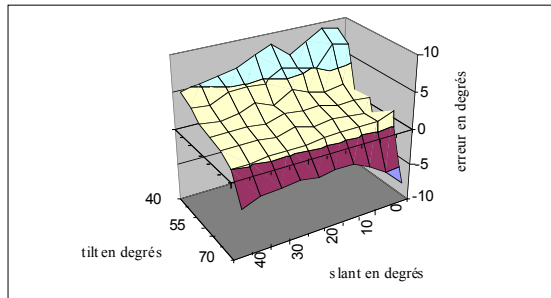
L'algorithme final, permettant de calculer l'angle de slant avec une erreur minimum, se compose de trois interpolations qui se déroulent de la manière suivante :

- Interpolation par des courbes générales « x^2 » avec suppression des pixels situés au bord de l'image.
- Interpolation par des courbes générales « x^2 » avec suppression des points éloignés de plus de 10% des points interpolés avec la méthode précédente.
- Interpolation par des cylindres paraboliques grâce à la connaissance de l'angle de tilt issu de la précédente interpolation.

Cet algorithme a été testé sur des images de synthèse comprenant une seule sinusoïde (figure 5a) et pour lesquelles nous avons fait varier les angles de tilt et de slant. Les résultats calculés à partir de la carte des échelles locales (figure 5b) sont les suivants :



(a)



(b)

FIG. 5 : Courbe de résultat. 5a : erreur sur l'angle de slant ; 5b : erreur sur l'angle de tilt.

Nous pouvons remarquer grâce à ces deux courbes que l'erreur sur l'angle de tilt est très faible ($<1^\circ$) alors que l'angle de slant est trouvé avec une erreur de quelques degrés. Cette différence peut être attribuée au fait que l'angle de tilt ne dépend que de l'orientation de la courbe, alors que l'angle de slant dépend de l'interpolation précise de la courbe.

Après avoir effectué ces calculs sur des images de synthèse, nous avons testé nos algorithmes sur des images réalistes (voire figure 6a et 6b). Avec ces images, nous obtenons les résultats donnés dans le tableau 1.

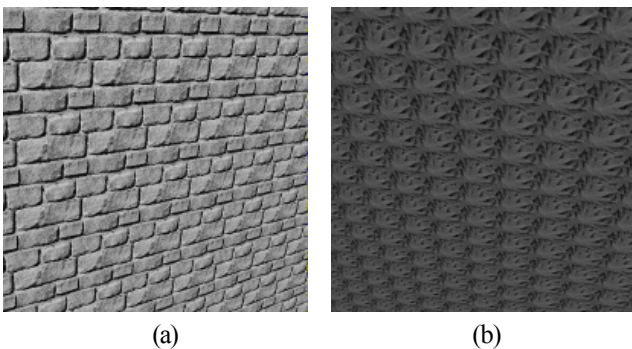


FIG. 6 : Images réalistes. 6a : briques; 6b : feuilles.

Sur ces images, nous voyons que l'erreur sur les angles de tilt reste faible (moins de 5°) mais que l'erreur sur les angles de slant augmente, sans toutefois dépasser les 15° .

TAB. 1 : Résultats avec des images réalistes

images		briques		feuilles	
Slant d'origine	Tilt d'origine	Slant	Tilt	Slant	Tilt
50	0	63,2	-3	44	-4
60	90	65,3	94	70	92
55	38	65,7	43	55,9	37
50	-35	60,5	-41	62,8	-29
60	-110	68,8	-106	71,6	-108

5. Conclusion

Nous présentons une méthode de calcul de l'orientation d'un plan texturé incliné par l'extraction de la carte des échelles locales. Le lissage de cette carte et une triple interpolation ont permis d'obtenir des résultats satisfaisants pour l'angle de tilt et une estimation raisonnable pour l'angle de slant. Ces résultats obtenus sur des images de test, ont été confirmés sur des images réalistes. La suite de ce travail porte sur deux aspects : l'application de cette technique à des images réelles (ce qui implique la prise en compte du modèle de la caméra) et la généralisation de ce travail à des surfaces de type sphère ou cylindre.

Références

- [1] Shape from texture ;J. Aloimonos ; Biol. Cyb., vol. 58, pp 345-360, 1988
- [2] Shape from texture: general principle ; K. Kanatani, T. Chou ; Artificial intelligence, vol. 38, n°1, pp 1-419, 1989
- [3] Shape from texture based on the ridge of continuous wavelet transform ; C.S. Lu, W.L. Hwong, H.Y. Mark Lio, P.C. Chung ; Proc. IEEE International conference on Image Processing, Switzerland, septembre 1996
- [4] Apport des contrastes locaux de luminance dans l'extraction de formes ; J. Plantier, S. Weibel ; Rapport de recherche du CERMA n°96-05, janvier 1996, FRANCE
- [5] Characterization of the spatial-frequency spectrum in the perception of shape from texture ; K. Sakai, L. H. Finkel ; J. Opt. Soc. Am. A., vol. 12 n°6, pp 1208-1224, 1995
- [6] Planar surface orientation from texture spatial frequencies ; B. Super, A. Bovik ; Pattern Recognition, vol. 28 n°5, pp 729-743, 1995
- [7] Computer analysis of visual textures ; F.TOMITA, S.TSUJI ; Kluwer academic pub., Boston