

Précision et performances d'une méthode de calcul de l'orientation d'un plan texturé incliné

Orientation computation of an inclined textured plane: accuracy and performances

par S. LELANDAIS¹, J. PLANTIER¹, L. BOUTTÉ^{1,2}

¹ CEMIF-LSC, Université d'Evry, 40 rue du Pelvoux, CE 1455 Courcouronnes, 91020 EVRY Cedex, France, Tél. : 01 69 47 73 03, Fax : 01 69 47 73 06, s.lelandais@iut.univ-evry.fr

² IMASSA, Département Sciences Cognitives, BP 73, 91223 Brétigny sur Orge, France, Tél. : 01 69 88 33 71, Fax : 01 69 88 33 02, jplantier@imassa.fr

résumé et mots clés

Le but de cet article est de présenter une méthode de calcul de l'orientation d'un plan texturé incliné à partir d'une seule vue de ce plan. Cette méthode est constituée de deux étapes. Dans un premier temps on calcule, à partir de l'image initiale, une carte des échelles locales. Ces échelles sont obtenues au moyen d'une décomposition en ondelettes de l'image d'origine. Puis on interpole cette carte des échelles locales par l'équation théorique de leurs variations. On obtient ainsi des paramètres qui permettent de calculer les angles de tilt et de slant, décrivant l'orientation du plan. Pour valider cette démarche, nous avons mené une étude théorique sur la précision qui pouvait être atteinte par une telle méthode. Nous avons pu mettre en évidence que, si la précision sur l'angle de tilt était assez bonne (de l'ordre de 1°), celle sur l'angle de slant n'excédait pas 5°, à condition que cet angle soit suffisamment important (supérieur à 40°). Mais la précision sur l'angle de slant est conditionnée par la connaissance des paramètres de prise de vue. En effet, nous avons mis en évidence que l'utilisation de valeurs erronées des paramètres de la caméra entraînerait une erreur maximum pour un slant entre 40° et 50°, c'est à dire, *a priori*, là où la méthode est la meilleure. Cette étude théorique a été validée par des expérimentations sur des images de synthèse et sur des images de textures réelles. Une base de données d'une centaine d'images a été constituée, au moyen d'une caméra préalablement calibrée, pour évaluer la qualité des résultats fournis par notre méthode.

Analyse de textures, échelles locales, décomposition en ondelettes, extraction de forme à partir de la texture, orientation d'un plan.

abstract and key words

The aim of this paper is to present one method for computing the orientation of an inclined textured plane with only one view of this plane. Two steps are used for this computation. First we build a local scales map by a wavelets decomposition of the image of the plane. Then we have to do an interpolation of this map by use the theoretical equation of the local scales variation. So we obtain features values which allow us to compute the tilt and the slant angles. After developing the computation technique, we do a theoretical study in order to know the precision of the method. For the tilt angle, the precision is about one degree, but for the slant angle the precision is only about five degrees, if the slant angle is over forty degrees. But, we have to know the camera parameters for computing the slant angle. If there is some errors about these parameters, so the slant angle will be bad. After this

study, we build a data base of one hundred images of real textures with different tilt and slant angles. The camera which has been used for acquiring the images has been calibrated. Results on this data base are agree with the theoretical study.

Texture analysis, shape from texture, local scales, wavelets decomposition, plane orientation.

1. introduction

Les méthodes de reconnaissance d'une forme à partir de la texture (Shape From Texture ou SFT) utilisent le principe qu'une texture va être déformée lors de sa projection sur un objet tridimensionnel. L'analyse de cette déformation permet donc de revenir à l'objet initial à partir de l'étude d'une seule image [11]. C'est la force de ces méthodes « SFT » que de faire de la reconstruction tridimensionnelle à partir de la vision monoculaire.

Le travail que nous présentons ici consiste dans le développement et l'expérimentation d'une méthode de recouvrement des angles de tilt et de slant caractérisant l'orientation d'un plan incliné texturé. Cette méthode, basée sur le modèle de la projection en perspective, utilise les trois effets de déformations imposés à la texture lors de sa projection sur le plan incliné que nous présenterons dans la deuxième partie de cet article. Deux étapes principales sont à la base de cette méthode, comme l'illustre la figure 1. En premier, le calcul des échelles locales, à savoir les inverses des fréquences locales, pour lequel nous avons développé une méthode originale fondée sur une décomposition en ondelettes utilisant des fonctions du type DOG (différences de gaussiennes). Puis, un calcul des angles donnant l'orientation du plan, qui est une amélioration de l'algorithme proposé par Lu et coll. Une fois cette méthode opérationnelle, nous nous sommes posé le problème de la qualité des résultats fournis. C'est ce travail que nous présentons dans la suite de cet article. Nous avons d'abord effectué une étude pour estimer les performances des deux méthodes et les comparer. Pour cela nous avons réalisé des tests sur des images de synthèse contenant une texture sinusoïdale. Une telle texture est définie à partir d'une fréquence unique et constitue donc l'exemple parfait du champ d'application de ces méthodes. Puis nous avons conduit une réflexion sur la précision et les facteurs pouvant entraîner une diminution de cette précision. Nous avons également comparé les résultats théoriques à des expérimentations menées par Knill [12] sur la perception humaine des textures dans le but de savoir si les résultats attendus étaient suffisamment performants pour une application dans un environnement réel. Enfin nous avons validé ces résultats théoriques par une expérimentation sur des textures réelles numérisées sous des orientations différentes par une caméra parfaitement calibrée. Nous avons ainsi constitué une base de données d'une centaine d'images sur lesquelles nous avons testé notre méthode. Nous avons pu voir que les résultats

expérimentaux suivaient assez bien les performances théoriques. Ces différents travaux nous ont conduits à un certain nombre de réflexions sur les méthodes « SFT » que nous proposerons en guise de conclusion.

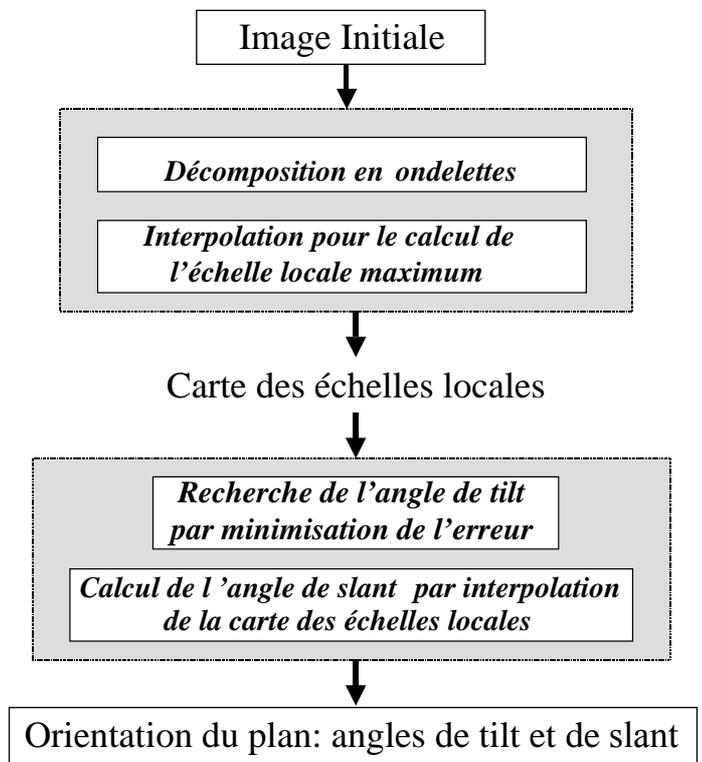


Figure 1. – Schéma de principe de la méthode proposée.

2. présentation des techniques « Shape From Texture »

Le problème du calcul de l'orientation d'une surface plane texturée peut être traité de différentes manières. Les méthodes peuvent être classées selon le modèle de la texture qui est supposé

et la manière dont cette description est utilisée pour remonter à l'orientation. Dans cet article, comme dans la plupart des travaux dans le domaine des méthodes « SFT », nous nous limitons au problème des plans inclinés sur lesquels sont projetées des macrotextures, et nous utilisons le modèle de la projection en perspective. Cette hypothèse prend en compte les trois types de déformations subies par la texture lors de sa projection, tels que Knill [12] les a définis (voir figure 2) :

1. Un effet de distance, qui est responsable du changement d'échelle de la texture dans l'image ;
2. Un effet de position qui est responsable de la variation de répartition des texels dans l'image.
3. Un effet de raccourci qui est responsable du changement de forme des texels dans l'image ;

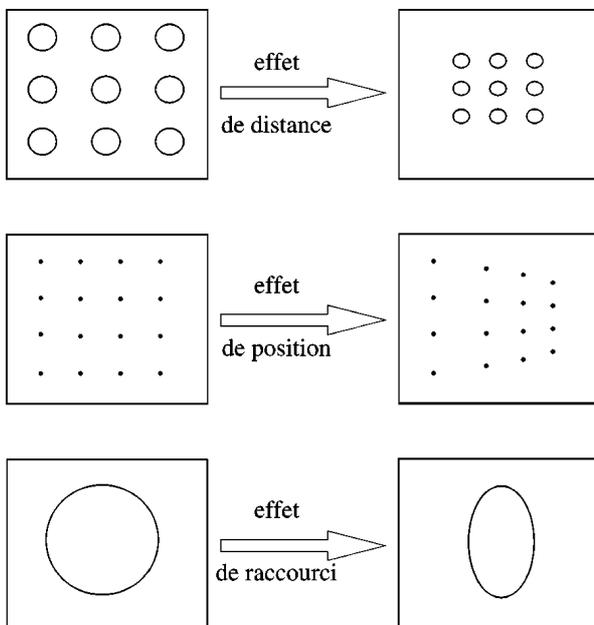


Figure 2. – Les trois déformations subies par la texture lors de sa projection.

L'utilisation de l'un ou plusieurs de ces trois effets permet de retrouver l'orientation du plan incliné [25].

L'effet de distance est responsable de la variation de taille des texels, par conséquent, pour retrouver l'orientation d'un plan texturé en utilisant cet effet, il faut mesurer la variation de taille des texels dans l'image. Les méthodes s'appuyant sur cette distorsion prennent comme hypothèse que la surface des texels reste constante sur toute la texture (avant sa projection dans l'image). Deux groupes de méthodes utilisent cet effet. Bajcsy [2], Brochard [6] et Blostein [3] proposent une interpolation de

la variation de la surface des texels par un modèle mathématique, les paramètres de l'interpolation permettant alors de remonter à l'orientation du plan. Ohta [19] calcule l'équation de la ligne de fuite du plan incliné à partir de la variation de surface des texels. Une fois cette équation connue, il est aisé de remonter à l'orientation du plan. Ces méthodes utilisant l'effet de distance sont dépendantes d'une extraction extrêmement précises des texels de l'image, ce qui n'est pas forcément aisé sur un plan incliné [4].

L'effet de position étant responsable du changement de répartition des texels dans l'image, les algorithmes utilisant cet effet pour retrouver l'orientation du plan supposent connue la répartition des texels avant la projection [1 ; 21]. En effet ces méthodes vont utiliser la variation de densité des texels dans l'image pour remonter à l'orientation. Pour trouver la densité des texels présents dans une zone de l'image, il est nécessaire d'extraire les texels ou du moins leurs contours. La variation théorique de la densité des objets présents dans l'image va être modélisée grâce aux équations de la projection en perspective.

Pour utiliser l'effet de raccourci, il est nécessaire d'avoir un certain nombre d'indications concernant la forme des texels. Par exemple Garding et coll. [10 ; 16] supposent que les textures étudiées sont isotropiques. D'autres auteurs [8 ; 22 ; 23] utilisent les caractéristiques fréquentielles des texels présents dans l'image. En effet, dans le spectre de l'image ces caractéristiques varieront en fonction de l'orientation du plan. Ils appliquent leur méthode à des surfaces courbes en supposant qu'un texel, au moins, est situé dans un plan perpendiculaire à la caméra.

Utiliser simultanément les trois types de déformation pour retrouver l'orientation du plan nécessite de trouver un paramètre de texture se trouvant affecté par ces trois effets. La fréquence dominante de la texture répond parfaitement à cette contrainte. En effet, cette fréquence, que l'on peut définir comme la fréquence la plus importante présente dans le spectre de la texture vue sous une orientation nulle, représente d'une part la fréquence des texels, ce qui implique la prise en compte des effets de distance et de raccourci, et d'autre part la fréquence de la répartition spatiale de ces texels, ce qui implique la prise en compte de l'effet de position. Le problème initial de ces méthodes est la détermination de la fréquence spatiale locale en chaque point de l'image. Krumm et coll. [13 ; 14] utilisent une transformée de Fourier par fenêtres glissantes. Le principe de leur méthode consiste à rechercher la variation des pics de la fréquence locale entre deux zones voisines de l'image de manière à en déduire l'inclinaison locale. L'image peut alors être constituée par plusieurs plans texturés d'orientations différentes. Super et coll. [24] proposent une technique fondée sur l'analyse de la variation des moments du second ordre du spectre local pour retrouver directement la forme d'objets texturés, en utilisant le modèle de la projection orthogonale. Une autre méthode utilisant la variation des fréquences locales pour remonter à l'orientation est celle développée par Lu et coll. [17]. Nous y reviendrons dans la suite de cet article.

3. extraction des échelles locales

La fréquence locale d'un point peut être définie comme la fréquence la plus importante dans le voisinage de ce point. Cette caractéristique peut être obtenue au moyen d'une décomposition en ondelettes de l'image d'origine [8 ; 22 ; 23 ; 24]. En fait, la convolution de l'image et d'une ondelette conduit à la représentation spatiale de la fréquence filtrée par cette ondelette. Le résultat de la convolution est d'autant plus élevé que la fréquence au voisinage du point est proche de la fréquence caractéristique de l'ondelette. Il est ainsi possible de couvrir toutes les fréquences présentes dans l'image en utilisant une base d'ondelettes qui couvre tout le spectre de celle-ci. Cette décomposition est une approximation de la distribution de Wigner. Dans notre étude nous avons utilisé, comme base de décomposition, des ondelettes du type DOG ou « différences de gaussiennes ». Ce choix a été motivé par un critère psychophysique. En effet les DOG donnent une bonne approximation du fonctionnement de la rétine humaine [9 ; 18 ; 20], et, dans la mesure où nous cherchons à voir un phénomène que l'œil perçoit, nous avons voulu nous mettre dans des conditions proches de la vision humaine. Pour s'abstraire des problèmes de variation locale de la luminosité, nous avons utilisé une formule d'ondelettes sensible au contraste [7]. Ces fonctions DOG sont appliquées à des images dans lesquelles les niveaux de gris sont la traduction numérique des luminances de la scène réelle. Dans l'équation (1), qui nous donne la forme générale des DOG utilisées pour la décomposition en ondelettes, $C_\sigma(x, y)$ représente le contraste multi-résolution de l'image $I(x, y)$.

$$C_\sigma(x, y) = \alpha^2 \frac{\left(e^{-\frac{\rho^2}{\sigma^2 K}} - \frac{1}{\alpha^2} e^{-\frac{\rho^2}{\alpha^2 \sigma^2 K}} \right) \otimes I(x, y)}{\left(e^{-\frac{\rho^2}{\alpha^2 \sigma^2 K}} \right) \otimes I(x, y)} \quad (1)$$

Dans cette équation, $\rho^2 = x^2 + y^2$, α , qui définit la largeur de la bande filtrée par la DOG, est pris égal à $\sqrt{2}$, et σ , qui représente les différentes échelles d'ondelettes, varie de 1 à la moitié de la taille de l'image ($1, \sqrt{2}, 2, 2\sqrt{2}, \dots$). L'échelle de l'ondelette représente l'inverse de sa fréquence caractéristique. La figure 3 illustre ce principe. En effet, en 3b on peut voir l'effet du filtrage de l'image 3a obtenu par la convolution d'une fonction DOG, analysant uniquement les variations de luminosité, avec l'image d'origine. L'image présentée en 3c est obtenue en utilisant l'équation (1). On peut constater que les zones de contours, qui expriment un contraste fort, ressortent beaucoup mieux sur cette image 3c.

La figure 4 présente une image complexe en niveaux de gris (4a) et sa décomposition en ondelettes obtenue pour 6 échelles différentes.

Pour obtenir la carte des échelles locales d'une image, il convient d'associer à chaque pixel l'échelle de l'ondelette qui a conduit à la réponse la plus importante lors de la décomposition (figure 5a). Une telle démarche entraîne un problème que l'on désignera sous le nom « d'effet de marches d'escalier ». En effet, le nombre d'ondelettes étant fini, des points voisins, qui possèdent des échelles locales proches comprises entre deux échelles caractéristiques de la base d'ondelettes, peuvent être affectés soit à l'une, soit à l'autre de ces deux échelles, entraînant ainsi une variation importante dans le résultat. Pour pallier cet inconvénient, Sakai [23] propose un lissage de la carte des échelles locales en faisant une convolution avec une gaussienne. Cette méthode permet d'obtenir une variation continue des échelles locales, mais sans obtenir une meilleure information sur la valeur des échelles elles-mêmes. El Ghadi et coll., de même

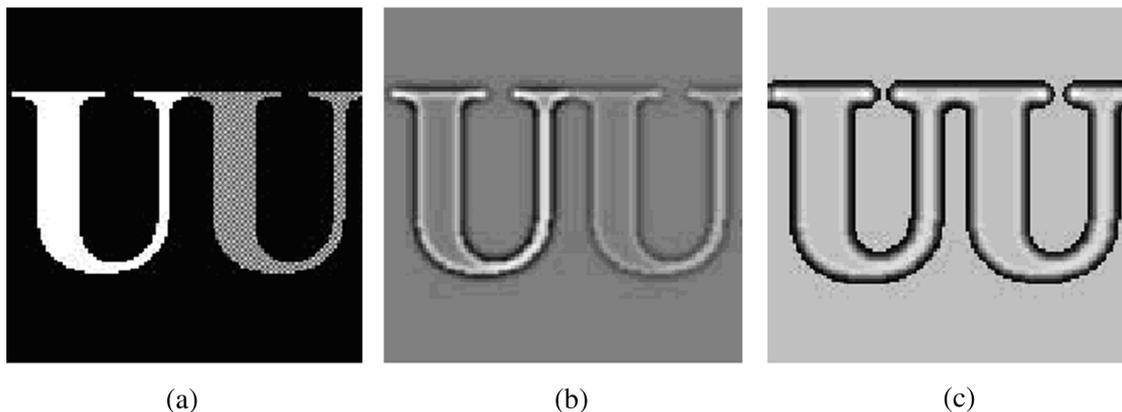
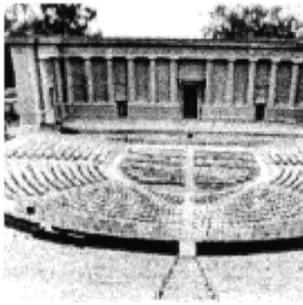


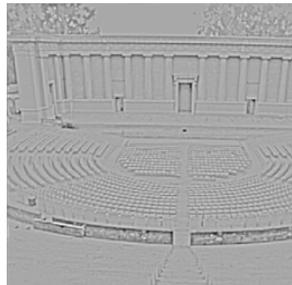
Figure 3. – Exemple de filtrage d'une image (a) 128x128 par une fonction DOG d'échelle 8 pixels et définie en luminosité (b) et en contraste (c).



(a) Image d'origine (256 × 256 pixels)



(b) $\sigma = 4$



(c) $\sigma = 8$



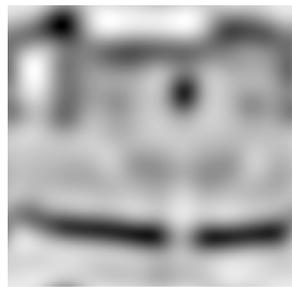
(d) $\sigma = 16$



(e) $\sigma = 32$



(f) $\sigma = 64$

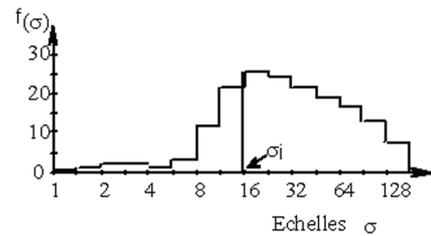


(g) $\sigma = 128$

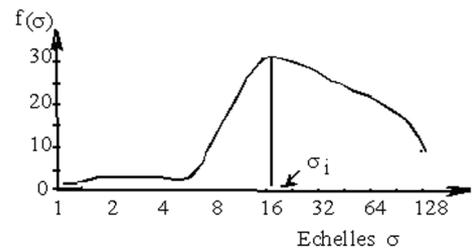
Figure 4. – Exemple de décomposition en ondelettes de contraste d'une image.

que Super et Bovik [8 ; 24], utilisent, pour obtenir la fréquence locale, le rapport de la dérivée de la convolution de l'image avec l'ondelette, par la convolution elle-même. Cette technique permet de combler les intervalles qui existent entre les fréquences centrales des ondelettes utilisées, conduisant ainsi à des valeurs continues pour les échelles ou les fréquences locales. L'inconvénient de cette méthode est la complexité des calculs ainsi que le risque de division par zéro entraînant des valeurs aberrantes. Pour obtenir une carte des échelles locales présen-

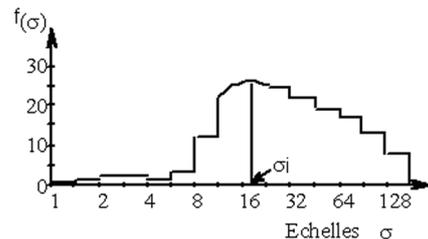
tant des valeurs continues en évitant les écueils précédents, nous avons développé une technique d'interpolation locale basée sur l'idée que le spectre local en chaque point est une courbe continue (figure 5b). Le maximum de cette courbe est défini par l'échelle de l'ondelette qui a une fréquence centrale égale à la fréquence locale du point considéré. La méthode de décomposition en ondelettes utilisée ici ne permet pas d'obtenir directement cette courbe, mais seulement d'en avoir un échantillonnage (figure 5a). Pour en trouver le maximum, nous procédons à une interpolation parabolique que nous restreignons à trois valeurs : l'échelle de l'ondelette ayant conduit au maximum global de la courbe, ainsi que les deux échelles l'entourant (figure 5c). La figure 5 présente les trois exemples de courbe permettant de décrire la variation des échelles en un point. Dans ces trois courbes, σ représente l'échelle de l'ondelette. En 5a, la courbe discrète conduit à l'attribution du maximum, soit 16, comme valeur d'échelle locale pour ce point. En 5b, on observe la courbe continue supposée déduite de la figure 5a. En 5c, on voit la courbe discrète avec le résultat de l'interpolation menée sur l'échelle 16 et ses deux voisines ($8\sqrt{2}$ et $16\sqrt{2}$), conduisant à l'attribution de la valeur 17,37 pour le point concerné.



(a) : Courbe discrète



(b) : Courbe continue



(c) : Courbe interpolée

Figure 5. – Courbes décrivant les réponses des ondelettes en un point particulier d'une image.

La figure 6 présente les résultats de ces différentes méthodes de calcul des échelles locales conduites sur l'image initiale présentée sur la figure 4a. La figure 6a montre le résultat de la méthode du maximum. En 6b, on peut voir l'effet du lissage pratiqué sur 6a. En 6c nous avons les résultats de la méthode de la dérivée, obtenus d'après l'algorithme de Super et Bovik décrit précédemment, et en 6d la méthode d'interpolation locale que nous proposons. Dans ces quatre images, la valeur de chaque pixel est donnée par l'équation (2). Le témoin 6e donne la correspondance entre échelles et niveaux de gris.

$$E(x, y) = \sigma_i \quad f(\sigma_i) = \text{Max}(f(\sigma)) \quad (2)$$

L'amélioration de la continuité des échelles est nettement visible entre l'image 6a et les trois autres. Les images 6c et 6d conduisent à des résultats similaires et il n'est pas possible, dans l'état

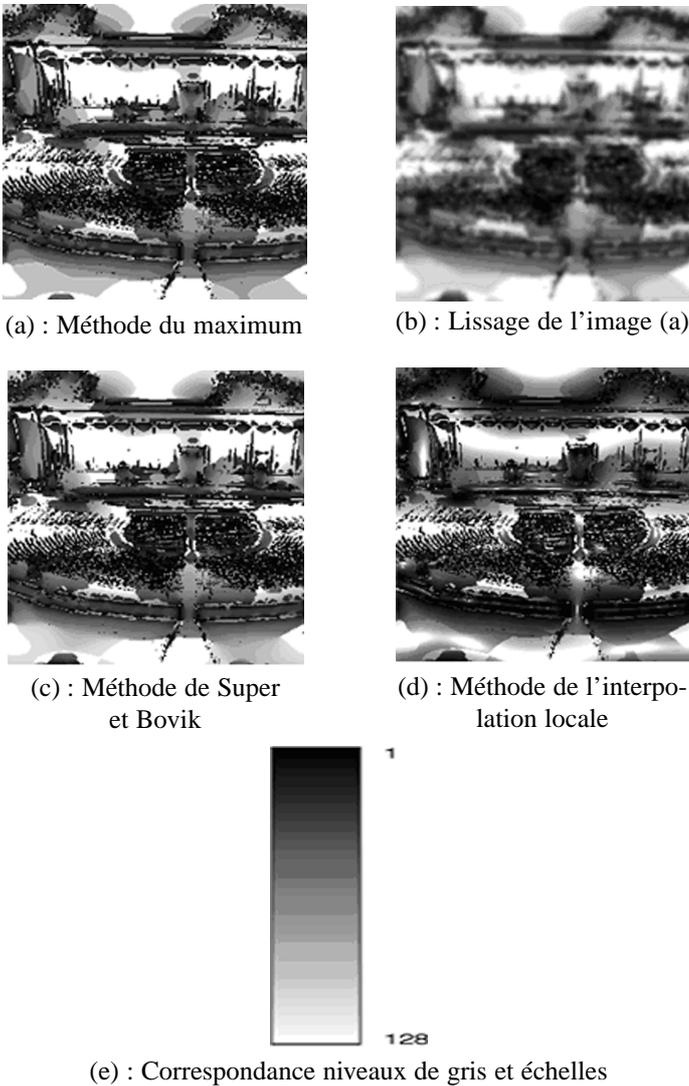


Figure 6. – Illustration des différentes méthodes permettant d'obtenir la carte des échelles locales à partir de l'image initiale présentée sur la figure 4a.

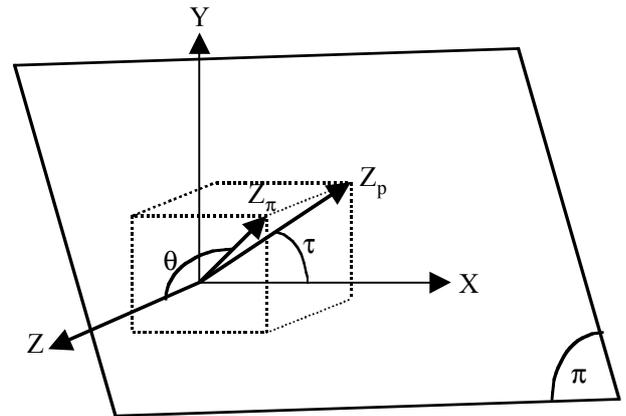
actuel de nos travaux, de décider si les échelles locales obtenues par la méthode de la dérivée décrivent mieux, ou moins bien, les variations réelles de l'image que celles obtenues par notre méthode. Nous avons donc décidé, dans la suite de notre travail, d'utiliser notre méthode d'interpolation locale qui a pour elle l'avantage de la simplicité de mise en œuvre.

4. calcul de l'orientation du plan incliné

4.1. définition des angles de tilt et de slant

Pour obtenir l'orientation du plan incliné texturé à partir d'une seule image nous devons étudier la déformation de la texture lors de sa projection. Cette déformation est décrite par la variation de la fréquence locale dans l'image. Une étude théorique, fondée sur un système de représentation utilisant les angles de tilt et de slant (figure 7), nous permet d'obtenir les équations qui relient l'orientation avec les fréquences locales. L'inclinaison d'une surface dans l'espace peut être parfaitement décrite à l'aide de sa normale, c'est-à-dire au moyen de deux rotations. Les deux angles choisis pour décrire cette orientation sont :

- L'angle de tilt (τ) qui correspond à la projection de la normale du plan texturé dans l'image (Z_p) et l'axe horizontal de l'image (X). Cet angle indique la direction d'inclinaison du plan Π .
- L'angle de slant (θ) qui représente l'angle entre la normale du plan texturé (Z_π) et l'axe optique de la caméra (Z). Cet angle correspond au degré d'inclinaison du plan.



Π : plan texturé, XY : plan image, Z_π : vecteur normal au plan Π ,
 Z : vecteur normal au plan image, Z_p : projection de Z_π sur le plan image XY ,
 θ : angle de slant (angle entre Z et Z_π)
 τ : angle de tilt (angle entre X et Z_p)

Figure 7. – Représentation de l'orientation du plan par les angles de tilt et de slant.

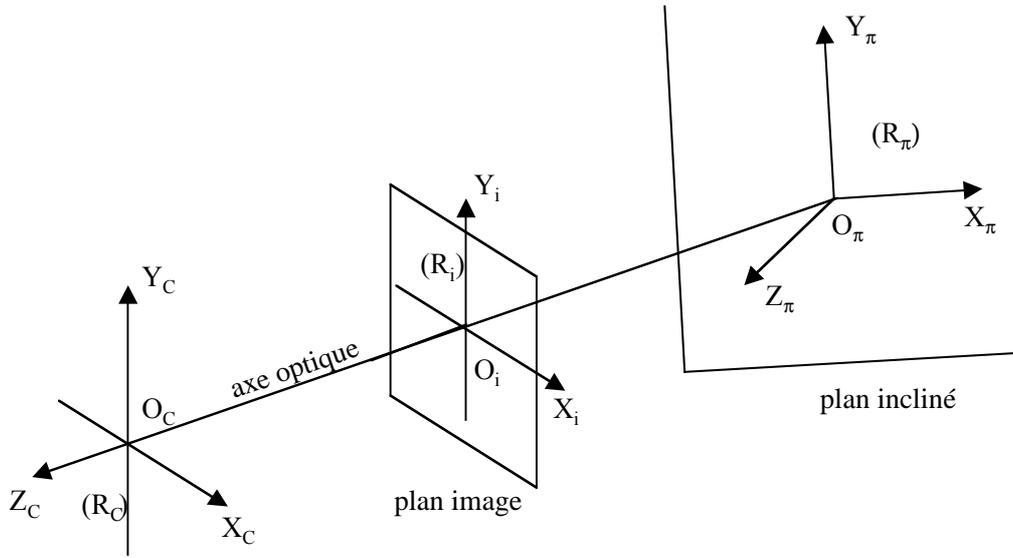


Figure 8. – Représentation des trois repères.

Pour définir les angles de tilt et de slant, nous utilisons trois repères (figure 8) :

- Le premier est le repère objet (R_π). Il est choisi par la personne qui acquiert les images. Il est constitué du plan Π et de l'axe Z_π .
- Le second est le repère caméra (R_C). Il est fixe par rapport à la caméra. Son centre est à l'intersection de l'axe optique et de la lentille de la caméra. Un de ses axes est l'axe optique Z_C et les deux autres sont parallèles à la matrice image (X_C et Y_C).
- Le troisième est le repère image (R_i). C'est un repère bidimensionnel dans lequel les coordonnées des points sont données en pixels. Ces deux axes sont parallèles aux axes X_C et Y_C du précédent repère.

Ω est la matrice de rotation qui permet le passage des coordonnées d'un point du repère objet au repère caméra, à la condition que les centres de ces deux repères soient confondus.

On a :

$$\Omega = \begin{pmatrix} (\sin^2 \tau + \cos^2 \tau \cos \theta) & (\sin \tau \cos \tau (\cos \theta - 1)) & (\cos \tau \sin \theta) \\ (\sin \tau \cos \tau (\cos \theta - 1)) & (\cos^2 \tau + \sin^2 \tau \cos \theta) & (\sin \tau \sin \theta) \\ (-\cos \tau \sin \theta) & (-\sin \tau \sin \theta) & (\cos \theta) \end{pmatrix} \quad (3)$$

4.2. détermination de l'angle de slant

Afin de trouver les équations reliant l'orientation avec les fréquences locales, nous allons commencer par un cas simple, soit

une orientation pour laquelle l'angle de tilt (τ) est égal à zéro. Soit $P_\pi = (x_\pi; y_\pi; 0)^t$ un point du plan incliné dans le repère objet. $P_c = (x_c; y_c; z_c)^t$ représente ce même point dans le repère caméra. $P = (x; y)^t$ est le projeté de P_c dans le repère image. On peut donc écrire :

$$P_c = \Omega \cdot P_\pi = \begin{pmatrix} x_\pi \cos \theta \\ y_\pi \\ -x_\pi \sin \theta \end{pmatrix} \quad (4a)$$

et :

$$P = \begin{pmatrix} \frac{x_\pi \cos \theta}{(f + Z) + x_\pi \sin \theta} \cdot f \\ \frac{y_\pi}{(f + Z) + x_\pi \sin \theta} \cdot f \end{pmatrix} \quad (4b)$$

Pour calculer la variation de la fréquence locale dans l'image, on prend une texture ne possédant qu'une seule fréquence spatiale. Soit $S(x_\pi, y_\pi)$ une texture sinusoïdale parfaite projetée sur le plan Π . Avec l'hypothèse d'un angle de tilt nul, la texture projetée gardera une valeur constante quelque soit « y ». L'équation de cette texture dans le plan Π est :

$$S(x_\pi, y_\pi) = k_1 \cos(\omega x_\pi) + k_2 \quad (5)$$

Dans cette équation (5), k_1 et k_2 sont des constantes permettant d'obtenir la dynamique voulue pour l'image. À partir des équations (4a) et (4b), on obtient :

$$x = \frac{x_\pi \cos \theta}{(f + Z) + x_\pi \sin \theta} \cdot f \quad (6)$$

Orientation d'un plan texturé incliné

Dans cette équation, « f » représente la distance focale de la caméra, « Z » la distance entre le plan incliné et le centre optique de la caméra et « θ » l'angle de slant. À partir de (6), on peut écrire la valeur de la coordonnée « x_π », soit :

$$x_\pi = \frac{(f + Z)}{f} \cdot \frac{1}{\cos \theta} \cdot \frac{x}{1 - \frac{x}{f} \cdot \tan \theta} \quad (7)$$

Nous obtenons alors la relation :

$$S(x, y) = k_1 \cos(\gamma(x_\pi)) + k_2$$

$$\text{avec } \gamma(x_\pi) = \omega \cdot x_\pi = \omega \left[\frac{(f + Z)}{f} \cdot \frac{1}{\cos \theta} \cdot \frac{x}{1 - \frac{x}{f} \tan \theta} \right] \quad (8)$$

La fréquence locale peut alors être définie par l'équation suivante :

$$\begin{aligned} F(x, y) &= \frac{1}{2\pi} \frac{d\gamma(x_\pi)}{dx} \\ &= k_1 \cdot \frac{\omega}{2\pi} \cdot \frac{(f + Z)}{f} \cdot \frac{1}{\cos \theta} \cdot \frac{1}{\left(1 - \frac{x}{f} \tan \theta\right)^2} \end{aligned} \quad (9)$$

À partir de l'équation (9), il est aisé d'obtenir l'échelle locale pour chaque point de l'image, celle-ci étant l'inverse de la fréquence. L'intérêt de cette inversion réside dans la simplification du modèle mathématique à interpoler pour remonter à l'orientation du plan. En effet, si les fréquences locales subissent une variation de type hyperbolique, la variation des échelles locales peut, quant à elle, être modélisée par un polynôme du second degré qui est plus facile à manipuler.

$$E(x, y) = \frac{1}{F(x, y)} = \frac{1}{k_1} (\beta_2 x^2 + \beta_1 x + \beta_0) \quad (10)$$

avec :

$$\beta_0 = \frac{2\pi}{\omega} \left(\frac{f}{f + Z} \right) \cos \theta \quad (11a)$$

$$\beta_1 = \beta_0 \left(\frac{-2}{f} \right) \tan \theta \quad (11b)$$

$$\beta_2 = \beta_0 \left(\frac{1}{f^2} \right) \tan^2 \theta \quad (11c)$$

À partir de ces trois équations, il nous est possible d'obtenir l'angle de slant de trois manières différentes :

$$\theta_1 = \arctan \left(-2f \frac{\beta_2}{\beta_1} \right) \quad (12a)$$

$$\theta_2 = \pm \arctan \left(\sqrt{f^2 \frac{\beta_2}{\beta_0}} \right) \quad (12b)$$

$$\theta_3 = \arctan \left(-\frac{f}{2} \cdot \frac{\beta_1}{\beta_0} \right) \quad (12c)$$

Chacune de ces trois équations donne, dans le cas d'une interpolation parfaite, le même résultat. Nous verrons dans la suite de cet article que des différences existent lors de l'interpolation d'échelles locales obtenues à partir d'images réelles. Ceci nous permettra de justifier le choix de l'équation retenue pour le calcul de l'angle de slant.

4.3. détermination de l'angle de tilt

Pour généraliser ces équations à n'importe quel angle de tilt, il faut faire subir à l'image inclinée d'un angle θ , une rotation de l'angle de tilt τ choisi. En posant $(x_r, y_r)^t$, les coordonnées après cette rotation, soit :

$$x = x_r \cos \tau + y_r \sin \tau,$$

l'équation (10) devient alors :

$$\begin{aligned} E(x_r, y_r) &= \beta_0 + \beta_1 x_r \cos \tau + \beta_1 y_r \sin \tau \\ &+ \beta_2 (x_r^2 \cos^2 \tau + y_r^2 \sin^2 \tau + 2x_r y_r \cos \tau \sin \tau) \end{aligned} \quad (13)$$

que l'on peut également écrire :

$$E(x_r, y_r) = a_1 x_r^2 + a_2 x_r y_r + a_3 y_r^2 + a_4 x_r + a_5 y_r + a_6 \quad (13')$$

À partir des équations (13) et (13'), on peut obtenir les valeurs des paramètres a_i :

$$(14a) \quad a_1 = \beta_2 \cos^2 \tau \quad (14b) \quad a_2 = 2\beta_2 \cos \tau \sin \tau$$

$$(14c) \quad a_3 = \beta_2 \sin^2 \tau \quad (14d) \quad a_4 = \beta_1 \cos \tau$$

$$(14e) \quad a_5 = \beta_1 \sin \tau \quad (14f) \quad a_6 = \beta_0$$

Nous obtenons à nouveau un système de trois équations nous permettant de déterminer les valeurs de l'angle de tilt :

$$\tau_1 = \arctan \left(\frac{a_5}{a_4} \right) \quad (15a)$$

$$\tau_2 = \arctan \left(\frac{a_2}{2a_1} \right) \quad (15b)$$

$$\tau_3 = \arctan \left(\frac{2a_3}{a_2} \right) \quad (15c)$$

L'angle de slant est donné par l'une des équations (12) en prenant les coefficients β_0 , β_1 et β_2 définis par :

$$\beta_2 = \frac{a_1}{\cos^2 \tau} \quad (16a)$$

$$\beta_1 = \frac{a_4}{\cos \tau} \quad (16b)$$

$$\beta_0 = a_6 \quad (16c)$$

4.4. implémentation pratique de la méthode et discussion

Pour effectuer ces calculs et trouver l'orientation du plan, il faut faire une interpolation de la courbe théorique (équation 13') à partir de la carte des échelles locales obtenues. La courbe à interpoler n'est pas un polynôme général du second degré, mais un paraboloïde orienté suivant la direction du tilt, comme l'illustre la figure 9.

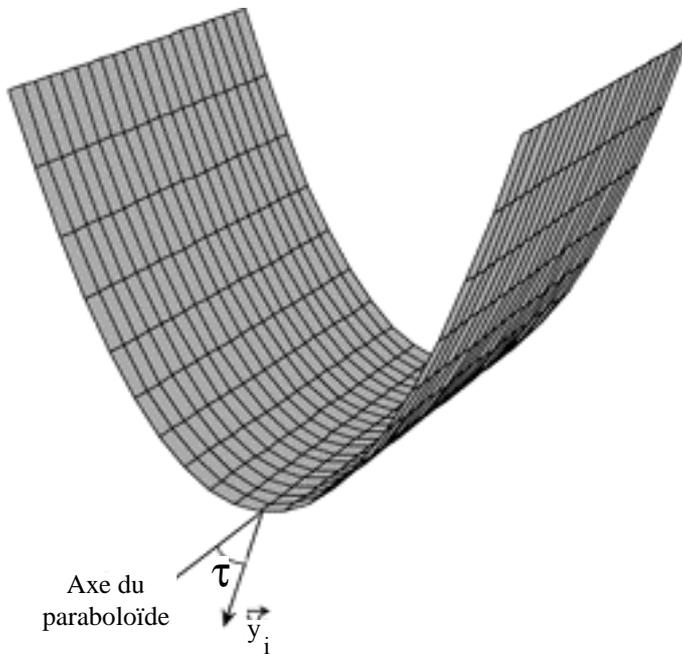


Figure 9. – Paraboloïde obtenu.

Les équations (12a), (12b) et (12c) montrent que le calcul de l'angle de slant est tributaire de la connaissance de la focale de la caméra. Il est donc nécessaire de calibrer la caméra, responsable de l'acquisition des images, de manière à connaître ses paramètres et à pouvoir obtenir exactement l'angle de slant. Nous discuterons plus loin l'imprécision due à des erreurs sur les paramètres de la caméra. Par contre, le calcul de l'angle de tilt n'est pas lié aux paramètres de la caméra ayant réalisé l'acquisition de l'image comme le montre les équations (15a), (15b) et (15c). Ceci peut s'expliquer par le fait que l'angle de tilt indique simplement le sens de l'inclinaison du plan, ce qui cor-

respond à la direction de variation de la texture. Comme la valeur de la focale de la caméra est responsable de l'importance de la déformation subie par la texture lors de sa projection dans l'image mais ne privilégie aucune direction, l'angle de tilt peut être déterminé sur des images pour lesquelles on ne connaît pas la focale utilisée lors de la prise de vue.

Une fois en possession de ces équations, il nous faut procéder à l'interpolation de la carte des échelles locales. Malheureusement, l'équation de la variation des échelles locales dans l'image (13) n'est pas linéaire à cause des termes « $\cos \tau$ » et « $\sin \tau$ » intervenant dans les paramètres de cette équation. Il n'est donc pas possible de faire directement une interpolation du paraboloïde ainsi décrit en utilisant la méthode des moindres carrés. Pour contourner ce problème, Lu et coll. [17] ont choisi d'interpoler la carte des échelles locales du plan texturé par l'équation générale d'un polynôme du second degré. L'avantage de cette méthode est sa grande simplicité, mais cette solution pose un problème car elle suppose que l'image de texture à étudier ne présente réellement qu'une seule fréquence dominante et que l'extraction des échelles locales est parfaite. La non-vérification de ces deux conditions lors de l'analyse de textures réelles entraîne des erreurs non négligeables sur les résultats, comme nous le montrerons plus loin. Par ailleurs, Lu et coll. utilisent l'équation (12a) pour le calcul de l'angle de slant. Cette équation conduit à des résultats très imprécis lorsque l'angle de slant est faible. Ceci est dû au fait qu'elle est basée sur la recherche de l'extremum du paraboloïde. Plus le paraboloïde devient plat, ce qui correspond à des θ faibles, plus il est difficile de situer avec précision la position de son extremum. Ce problème sera mis en évidence dans le paragraphe suivant où nous présenterons des résultats comparatifs entre les deux méthodes.

Pour corriger les inconvénients que nous venons de citer, nous avons fait les choix suivants. Tout d'abord nous avons retenu l'équation (12b) pour le calcul de l'angle de slant ; cette équation est la seule des trois qui permet de s'abstraire de la position de l'extremum du paraboloïde. Nous nous sommes ensuite intéressés au problème de l'interpolation de la carte des échelles locales. Pour pouvoir obtenir l'équation d'un « vrai » paraboloïde, et donc des résultats plus justes, il nous faut linéariser l'équation (13). Ceci est possible à condition de connaître la valeur de l'angle de tilt. L'équation (17) reprend l'équation (13) en posant :

$$X = x \cos \tau \text{ et } Y = y \sin \tau$$

$$E(X, Y) = \beta_2(X^2 + 2XY + Y^2) + \beta_1(X + Y) + \beta_0 \quad (17)$$

On peut alors obtenir les paramètres β_0 , β_1 et β_2 en résolvant le système d'équations (18), donné par la méthode des moindres carrés, après avoir introduit les notations suivantes :

$$c = \cos \tau ; \quad s = \sin \tau ; \quad A_n = x_n \cos \tau + y_n \sin \tau$$

$$\begin{aligned}
 & \beta_2 \sum_{n=1}^N (cx_n^2 + sy_n^2)^2 + \beta_1 \sum_{n=1}^N (cx_n + sy_n) + \beta_0 N = \sum_{n=1}^N z_n \\
 & \beta_2 \sum_{n=1}^N A_n (cx_n^2 + sy_n^2)^2 + \beta_1 \sum_{n=1}^N A_n (cx_n + sy_n) + \beta_0 \sum_{n=1}^N A_n \\
 & = \sum_{n=1}^N (z_n A_n) \\
 & \beta_2 \sum_{n=1}^N A_n^2 (cx_n^2 + sy_n^2)^2 + \beta_1 \sum_{n=1}^N A_n^2 (cx_n + sy_n) + \beta_0 \sum_{n=1}^N A_n^2 \\
 & = \sum_{n=1}^N (z_n A_n^2)
 \end{aligned} \tag{18}$$

Dans ce système d'équations, (x_n, y_n) sont les coordonnées d'un point de l'image des échelles locales, z_n est l'échelle de ce point, et N désigne le nombre total de points.

Nous allons donc, pour toutes les valeurs de tilt possibles, faire une interpolation de la carte des échelles locales, et nous résoudrons à chaque fois le système d'équations (18) de manière à obtenir les valeurs des paramètres β_0 , β_1 et β_2 . Puis, pour chaque interpolation nous calculerons l'erreur d'interpolation commise, en utilisant l'équation suivante :

$$e_i(\tau) = \sum_{n=1}^N [\beta_2 A_n^2 + \beta_1 A_n + \beta_0 - z_n]^2 \tag{19}$$

L'angle de tilt retenu sera celui conduisant à la valeur minimum de l'erreur e_i . L'avantage de cette technique est de toujours interpoler un paraboloïde. Son inconvénient est d'augmenter le nombre de calculs nécessaires pour obtenir la bonne orientation. Pour limiter les calculs, nous discrétisons l'espace des angles de tilt avec une résolution de 1 degré. Une approche par dichotomie serait également possible pour limiter les calculs, mais, dans cette phase d'étude, nous avons préféré travailler sur toutes les valeurs possibles de l'angle de tilt. La précision maximum obtenue sur l'angle de tilt sera donc de 1°. Toutefois, si une plus grande précision s'avérait nécessaire, il serait possible d'obtenir une précision plus grande sur « τ », sans augmenter considérablement le nombre de calculs, en procédant en deux temps : une recherche de la zone où se situe l'angle de tilt en utilisant une discrétisation « grossière » de l'espace des « τ » suivie d'une détermination précise du tilt en utilisant un échantillonnage plus fin, uniquement dans cette zone.

Nous allons maintenant présenter et comparer les résultats obtenus par ces deux méthodes sur des images de synthèse correspondant parfaitement au champ d'application de la méthode. Nous commenterons ces résultats en analysant les facteurs qui influent sur la précision. Nous ferons ensuite un calcul des erreurs provoquées par les paramètres de prise de vue.

5. précision théorique de la méthode et discussion

5.1. résultats théoriques sur des images de synthèse

Dans un premier temps nous avons voulu estimer la précision, sur les valeurs des angles de tilt et de slant, fournie par cette méthode. Sachant que celle-ci est basée sur l'analyse des variations de la fréquence caractéristique de la texture, nous avons fabriqué une image de texture monofréquentielle. Si nous reprenons l'équation (8), nous obtenons directement l'image d'un plan incliné contenant une sinusoïde. L'inclinaison de ce plan est uniquement définie par l'angle de slant, l'angle de tilt étant pris égal à zéro. Pour obtenir toutes les orientations possibles, nous devons simplement faire une rotation, de la valeur de l'angle de tilt souhaité, de l'image contenant la sinusoïde déjà inclinée de l'angle de slant voulu. Cela peut se faire en calculant une nouvelle intensité pour chaque pixel de l'image, en remplaçant « x » dans l'équation (8) par :

$$(x - c_x)\cos \tau + (c_y - y)\sin \tau \tag{20}$$

où c_x et c_y représentent les coordonnées du centre optique de la caméra. Le pixel de coordonnées (x, y) aura donc pour niveau de gris :

$$\begin{aligned}
 & I(x, y) \\
 & = k_1 \cos \left(\omega \left[\frac{(f + Z)}{f} \frac{1}{\cos \theta} \frac{(x - c_x)\cos \tau + (c_y - y)\sin \tau}{1 - \frac{(x - c_x)\cos \tau + (c_y - y)\sin \tau}{f} \tan \theta} \right] \right) \\
 & + k_2
 \end{aligned} \tag{21}$$

où k_1 et k_2 jouent le même rôle que dans l'équation (5). La figure 10 montre trois exemples du résultat de la synthèse d'une image de sinusoïde obtenus par cette méthode. Il est donc possible de générer tous les angles θ et τ voulus afin de tester notre algorithme et de comparer ses résultats à ceux obtenus par la méthode de Lu et *coll.* Dans la phase de test nous avons fait varier θ de 0° à 70°, et τ de 0° à 90°, avec un pas de 5°.

Les figures 11a et 11b présentent la précision obtenue, pour le calcul des angles de tilt et de slant, par notre méthode et celle de LU sur ces images de synthèse. Les erreurs sur l'angle de tilt (figure 11a) sont très proches entre les deux méthodes, aussi n'avons nous donné que la courbe de nos résultats. Pour notre méthode, ces erreurs sont en moyenne toujours inférieures à 1° ce qui est la limite de précision que nous pouvons atteindre. Pour ce qui est des résultats sur l'angle de slant (figure 11b) notre méthode apporte une nette amélioration, spécialement pour des angles de slant faibles. Lorsque θ augmente, nous notons encore une erreur plus faible pour notre méthode, mais l'amélioration est d'autant moins significative que θ est important.

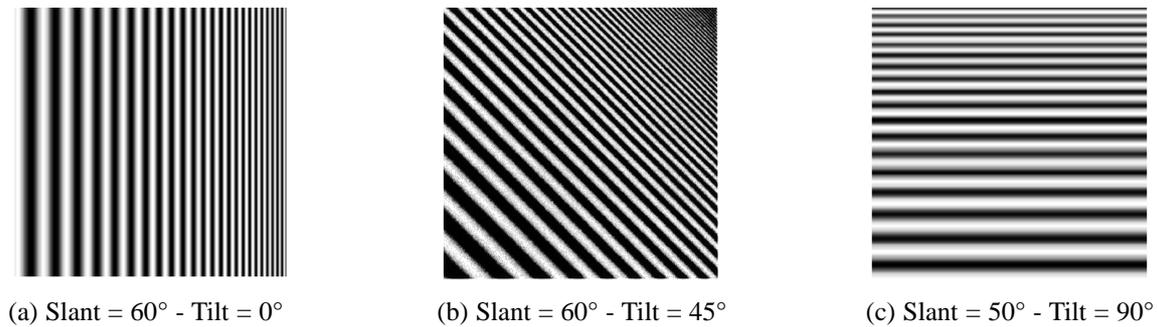


Figure 10. – Exemples d'images de synthèse contenant une sinusoïde orientée.

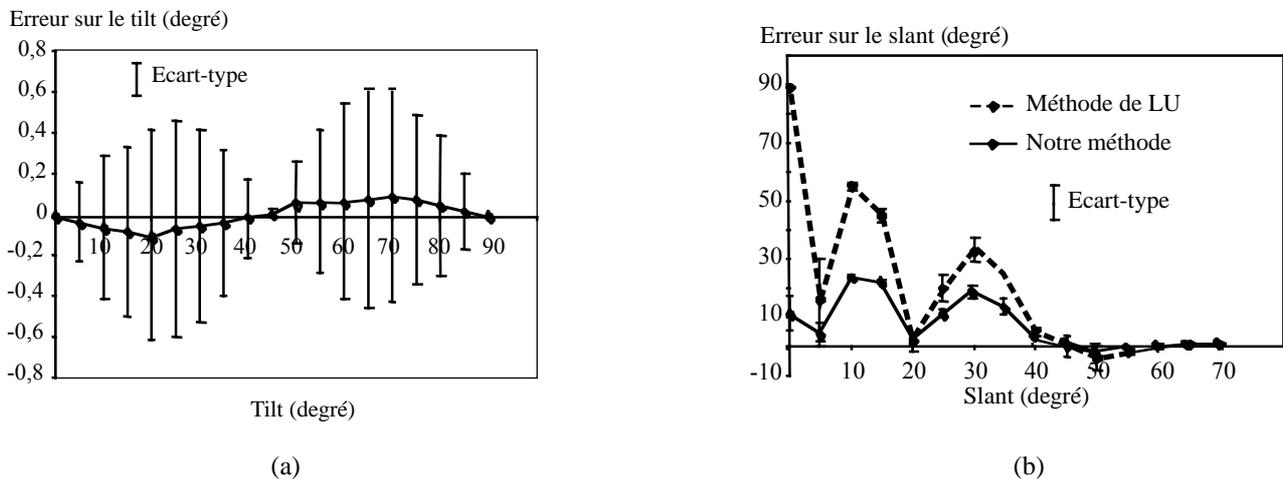


Figure 11. – Comparaison des erreurs commises par les deux méthodes, sur les angles de tilt (a) et de slant (b), pour les images de synthèse.

En conclusion, nous pouvons dire que sur des images de sinusoïdes parfaites, correspondant exactement au champ d'application de la méthode, l'estimation de l'orientation pourra se faire avec une précision de l'ordre de 1° sur l'angle de tilt. La précision sur l'angle de slant sera nettement moins bonne. Elle pourra atteindre 5° pour des angles de slant égaux ou supérieurs à 40° . Pour des angles plus faibles les erreurs peuvent atteindre 20° . Les améliorations essentielles apportées par notre méthode résident dans la diminution de l'erreur commise sur θ et dans une meilleure stabilité des résultats.

Nous avons complété ces calculs d'erreurs par une analyse de la variation de l'estimation de l'inclinaison entre deux images inclinées du même angle de slant [5]. Cette démarche vise à expliquer les écarts-types importants donnés par les deux méthodes. Dans un premier temps, nous avons vérifié que la valeur de l'angle de tilt avait une influence sur la stabilité des résultats. En effet lorsque τ varie de 0° à 45° , la taille de l'axe de l'interpolation augmente, entraînant l'augmentation du nombre de données permettant d'estimer la courbure du paraboloïde, et ainsi améliorant la précision des résultats. La compa-

raison des résultats expérimentaux sur l'angle de slant nous a permis de vérifier la véracité de cette assertion. Nous nous sommes également penché sur le problème de l'isotropie de la texture. Nous avons pu constater que, si une texture n'est pas isotrope, selon l'angle de tilt utilisé pour faire la rotation (équation 20), les déformations imposées à sa fréquence dominante ne vont pas être perçue de la même manière et ne vont donc pas entraîner les mêmes erreurs lors de l'estimation du tilt. Ceci explique également les variations d'estimation de l'angle de slant. La variation de la phase de la sinusoïde est également à l'origine de la variation de l'estimation de l'orientation. En effet, ce phénomène, que nous pouvons assimiler à une translation de la texture dans l'image, va entraîner une translation de la carte des échelles locales, et par la suite une modification de l'estimation de l'orientation du plan. L'influence de cette variation de phase est d'autant plus importante que la valeur de θ est faible. Ceci s'explique par le fait que, plus le slant est important, plus le nombre de périodes de la sinusoïde observée est important entraînant ainsi une extraction plus précise des échelles locales. L'échelle de la sinusoïde, qui correspond à la taille du motif élé-

mentaire de la texture, est également responsable des différences d'estimation de l'orientation. Nous avons réalisé l'estimation moyenne de l'angle de slant pour six périodes différentes (de 20 à 70 pixels). L'analyse de ces résultats nous a permis de remarquer que les extrema locaux, visibles sur la figure 8b, ne se trouvaient pas aux mêmes emplacements, et que leurs intensités variaient. Ceci suggère une dépendance entre la taille du texel et les échelles des ondelettes analysantes. Par ailleurs, l'erreur obtenue pour des angles de slant faibles est plus importante pour des sinusoides de fortes périodes. Une des explications de ce phénomène réside dans le fait que, pour des périodes élevées, pour obtenir une information suffisante pour calculer l'orientation du plan, il faudrait observer un grand nombre d'oscillations qui occuperaient une surface supérieure à celle de l'image analysée. D'autre part, on rencontre un problème sur les sinusoides de périodes très faibles qui va se traduire par une forte compression de la texture pour des orientations importantes. L'alternance entre les zones claires et sombres de la sinusoides ne sera plus perceptible. L'échelle locale ne pourra plus être extraite convenablement et la reconstruction correcte de l'orientation du plan ne sera plus possible. Nous avons vérifié ceci sur une sinusoides de période 10 pixels. Pour un angle de slant supérieur à 60°, les résultats de l'estimation se dégradent, ce qui n'a jamais été observé dans les résultats précédents.

Tous ces commentaires montrent que ces méthodes « SFT », si elles permettent d'obtenir des résultats convenables dans un certain nombre de situations, sont aussi très sensibles à de nombreux paramètres.

5.2. influence des paramètres de la prise de vue sur la précision

L'objectif de la méthode que nous venons de décrire est d'être utilisée sur des scènes réelles et non pas sur des images de synthèse. Les équations données précédemment montrent que dans le cas des images réelles, il est nécessaire de connaître certains paramètres de la caméra utilisée pour l'acquisition des images pour pouvoir procéder à la reconstruction de l'orientation du plan. Ces paramètres sont :

- la distance focale de l'objectif de la caméra, qui nous donne la relation entre la courbure du paraboloides et l'inclinaison du plan ;
- la position du centre optique de la caméra, qui représente le centre du repère de référence utilisé pour les calculs.

Nous avons analysé, dans un premier temps, l'influence d'une valeur erronée de la focale. Cette étude n'a été faite que pour l'angle de slant car, comme nous l'avons vu précédemment, l'angle de tilt ne dépend pas de f (équations 15). En supposant que la variation des échelles locales est parfaitement connue, c'est-à-dire qu'il n'y a pas d'erreur de mesure sur les paramètres β_0 , β_1 et β_2 , alors l'angle de slant est donné par la formule suivante :

$$\theta' = \arctan((1 + e)\tan \theta) \tag{22}$$

Où « e » représente l'erreur relative sur la focale, θ l'angle de slant réel et θ' l'angle de slant mesuré. L'erreur commise sur cet angle de slant s'exprimera alors par :

$$e_1 = |\theta' - \theta| \tag{23}$$

Comme l'illustre la figure 12, si l'incertitude sur la focale est faible ($e < 5\%$), son effet sur l'erreur commise sur l'angle de slant est inférieure à 1°. Par contre, cette erreur peut dépasser les 5° pour des valeurs de « e » supérieures à 20%. Notons aussi que cette erreur est maximale pour $\theta = 45^\circ$.

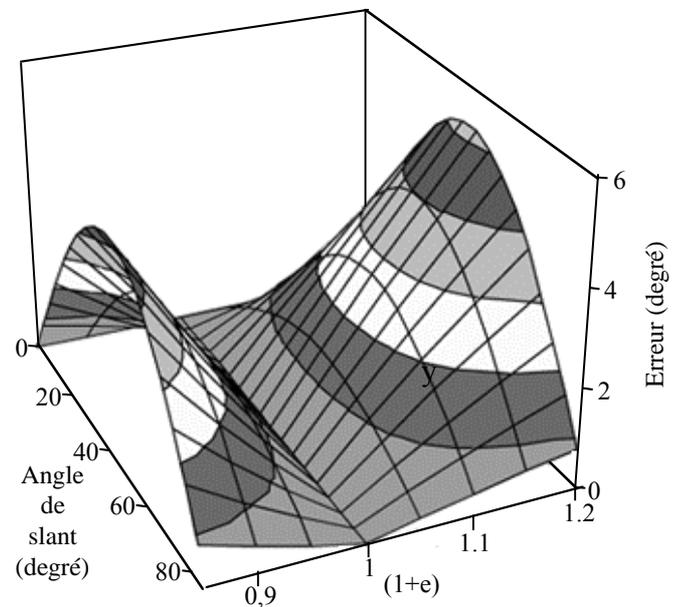


Figure 12. – Erreur provoquée par une focale erronée.

Nous nous sommes ensuite intéressés à l'erreur induite par une mauvaise connaissance de la position du centre optique. L'origine du repère utilisé pour faire l'interpolation du paraboloides est positionnée sur le centre optique de la caméra. Une variation de la position du centre optique entraîne donc une translation du repère de calcul. Soit le changement de repère :

$$x'_i = x_i - \Delta x_i \tag{24a}$$

$$y'_i = y_i - \Delta y_i \tag{24b}$$

où Δx_i et Δy_i représentent les erreurs commises en x et en y sur la mesure du centre optique. L'équation (13) du paraboloides devient donc :

$$\begin{aligned}
 E_i(x'_i, y'_i) &= a_1(x'_i + \Delta x_i)^2 + a_2(x'_i + \Delta x_i)(y'_i + \Delta y_i) \\
 &+ a_3(y'_i + \Delta y_i)^2 + a_4(x'_i + \Delta x_i) \\
 &+ a_5(y'_i + \Delta y_i) + a_6 \\
 &= a'_1 x_i'^2 + a'_2 x'_i y'_i + a'_3 y_i'^2 + a'_4 x'_i + a'_5 y'_i + a'_6
 \end{aligned} \quad (25)$$

avec :

$$a'_1 = a_1 \quad (26a)$$

$$a'_2 = a_2 \quad (26b)$$

$$a'_3 = a_3 \quad (26c)$$

$$a'_4 = 2\Delta x_i a_1 + a_2 \Delta y_i + a_4 \quad (26d)$$

$$a'_5 = 2\Delta y_i a_3 + a_2 \Delta x_i + a_5 \quad (26e)$$

$$a'_6 = (\Delta x_i)^2 a_1 + a_2 \Delta x_i \Delta y_i + a_3 (\Delta y_i)^2 + a_4 \Delta x_i + a_5 \Delta y_i + a_6 \quad (26f)$$

Comme on peut le constater, les coefficients a_1 , a_2 et a_3 ne sont pas modifiés par la translation du repère de travail. Par conséquent le calcul de l'angle de tilt s'effectuera en utilisant les équations (15b) ou (15c) qui ne seront pas affectées par une possible erreur sur la position du centre optique. Par contre le calcul de l'angle de slant est tributaire de la connaissance de ce centre optique. Il nous faut donc calculer l'influence d'une position erronée de ce point. Trois équations (12a, 12b et 12c) nous permettent d'obtenir la valeur de « θ ». Les calculs menés à partir de chacune de ces trois formules conduisent à une erreur quasi identique [5]. Comme nous avons précisé que nous utilisons l'équation (12b) pour le calcul de l'angle de slant, nous développerons le calcul dans ce cas uniquement. Cette équation est modifiée par les valeurs des paramètres obtenus dans les équations (26). On obtient alors une nouvelle formule pour l'angle de slant :

$$\theta' = \arctan \left(\sqrt{f^2 \frac{\frac{a'_1}{\cos^2 \tau}}{a'_6}} \right) \quad (27)$$

En supposant que les paramètres β_0 , β_1 et β_2 sont parfaitement connus et en posant la relation suivante :

$$A = (\Delta x_i \cos \tau + \Delta y_i \sin \tau) \quad (28)$$

donc $\forall \Delta x_i, \Delta y_i, A \leq \sqrt{(\Delta x_i)^2 + (\Delta y_i)^2}$
Le coefficient a'_6 peut être mis sous la forme :

$$a'_6 = \beta_2 A^2 + \beta_1 A + \beta_0 \quad (29)$$

Nous obtenons alors l'angle de slant par l'équation :

$$\theta' = \arctan \left(\sqrt{f^2 \frac{\tan^2 \theta}{A^2 \tan^2 \theta - 2fA \tan \theta + f^2}} \right) \quad (30)$$

L'erreur commise sur l'angle de slant va s'exprimer par e_2 :

$$e_2 = |\theta' - \theta| \quad (31)$$

Sur la figure 13, nous exprimons l'erreur sur l'angle de slant en fonction de A qui représente l'incertitude sur la position du centre optique de la caméra. Cette erreur devient rapidement importante et peut atteindre plusieurs degrés pour des valeurs de A dépassant 50 pixels. On constate également que l'erreur croît régulièrement pour θ variant de 0° à 90° .

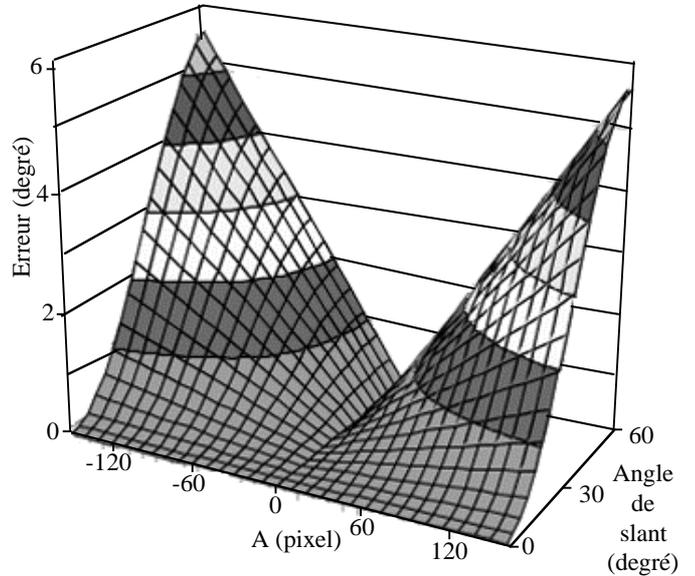


Figure 13. – Erreur provoquée par une position erronée du centre optique.

Le dernier calcul réalisé sur ces erreurs consiste en une analyse de l'effet cumulé de ces deux incertitudes. En utilisant l'équation (12b), nous arrivons à une formule pour l'angle de slant qui est :

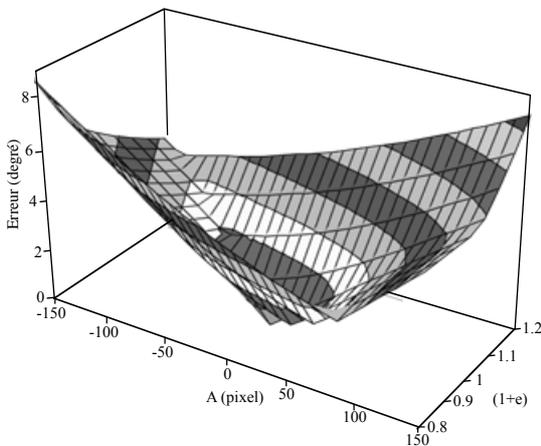
$$\theta' = \arctan \left(\sqrt{(1+e)f^2 \frac{\tan^2 \theta}{A^2 \tan^2 \theta - 2fA \tan \theta + f^2}} \right) \quad (32)$$

L'erreur cumulée, que nous appellerons e_3 , sera donnée par l'équation suivante :

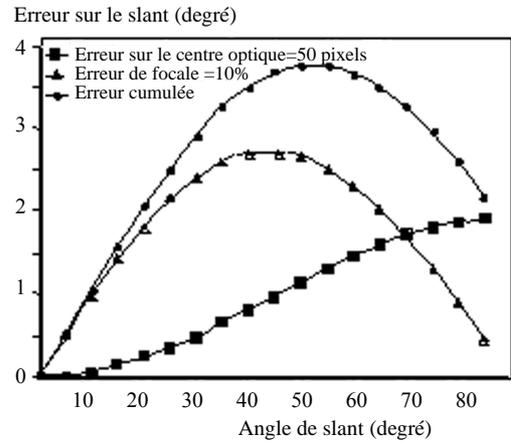
$$e_3 = |\theta' - \theta| \quad (33)$$

La figure 14a présente, en fonction des valeurs de A et de $(1+e)$, le maximum de l'erreur e_3 obtenue pour l'ensemble des

Orientation d'un plan texturé incliné



(a) : Valeurs maximales de e_3 pour différents angles de slant



(b) : Erreurs e_1 , e_2 et e_3 pour une valeur particulière de A et $(1 + e)$

Figure 14. – Effets de l'incertitude sur la valeur de la focale et sur la position du centre optique.

orientations possibles du plan. C'est-à-dire que nous avons calculé e_3 pour chaque valeur de slant, et que nous avons conservé uniquement son maximum. La figure 14b récapitule la forme de la variation des erreurs e_1 , e_2 , et e_3 pour des valeurs particulières de A , soit 50 pixels, et de e , soit 10 %.

Sur la figure (14a), on constate que l'erreur e_3 dépasse rapidement quelques degrés. Les valeurs de A et $(1 + e)$, sont importantes, mais limitées à des erreurs potentielles sur la focale et la position du centre optique. Si ces erreurs augmentaient, l'erreur e_3 augmenterait également. La figure (14b) montre que l'erreur provoquée par le cumul des erreurs sur la focale et le centre optique (e_3) peut être approchée par la somme de ces deux types d'erreurs (e_1 et e_2). Nous pouvons constater que cette erreur est maximum pour des angles de slant compris entre 40° et 70° , ce qui correspond à la plage angulaire dans laquelle la méthode est censée donner les meilleurs résultats. On comprend donc aisément l'importance de calibrer la caméra qui servira à l'acquisition des images servant au calcul d'orientation, de manière à optimiser la précision de la valeur du slant.

5.3. comparaison avec le système visuel humain

Après avoir étudié les différents types d'erreurs qui pouvaient survenir, nous avons souhaité comparer les résultats théoriques attendus de ces méthodes, avec ceux procurés par le système visuel humain. Pour cela nous avons utilisé les travaux de Knill [12]. Il a réalisé l'expérience suivante. Il a proposé à des sujets d'estimer la différence d'orientation entre deux images, chacune de ces images consistant en une vue d'un plan incliné. Les deux plans supportaient la même texture. La courbe moyenne obtenue à partir des résultats de Knill est présentée sur la figure 15. Elle

montre une amélioration de la précision de l'appréciation de la différence lorsque l'angle de slant augmente. Cette amélioration est particulièrement sensible pour des angles de slant supérieurs à 50° . L'erreur devient alors inférieure à 10°. La deuxième courbe présentée sur la figure 15 est celle obtenue à partir des résultats de notre algorithme appliqué aux images des sinusoides.

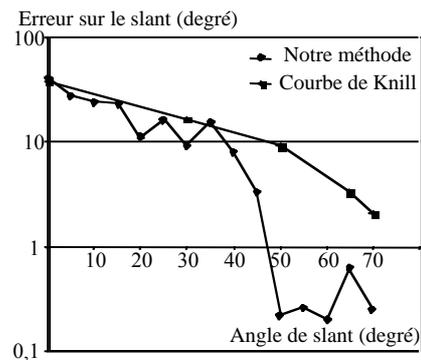


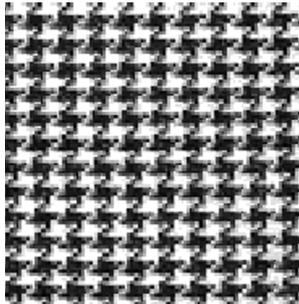
Figure 15. – Comparaison des performances de notre méthode et du système visuel humain (échelle logarithmique).

Nous voyons sur cette figure 15 que, même si nos résultats n'ont pas la régularité de ceux obtenus par Knill, les deux courbes ont la même tendance décroissante. Les erreurs dues au système visuel humain sont toujours supérieures à celles commises par l'algorithme d'interpolation des échelles locales. La précision relative de cet algorithme est malgré tout du même ordre de grandeur que celle délivrée par la vision humaine. Ceci ne signifie pas que les résultats obtenus sont les meilleurs possibles, mais permet de valider la méthode utilisée par rapport à une référence connue : l'homme.

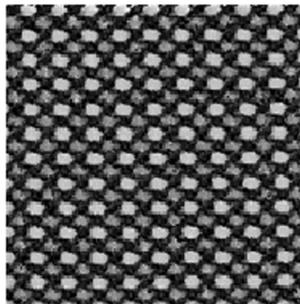
6. résultats expérimentaux

Pour finir cette étude, il fallait vérifier que les résultats espérés étaient réellement atteints sur des images réelles. Pour ce faire, nous avons acquis 90 images de macrotextures variées sous des

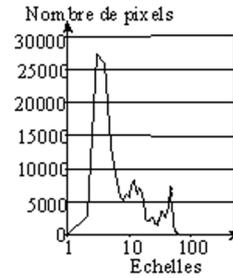
orientations différentes. La caméra utilisée pour la constitution de cette base de données a été soigneusement calibrée de manière à ce que ses paramètres soient connus avec précision. Nous avons choisi dix textures différentes que nous pouvons diviser en quatre catégories. Les figures 16, 17, 18 et 19, présentent un exemple de texture de chacune de ces catégories. Ces



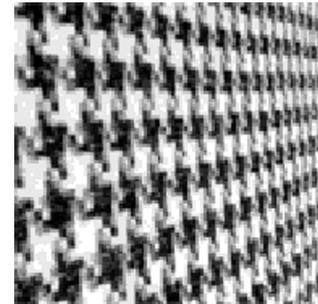
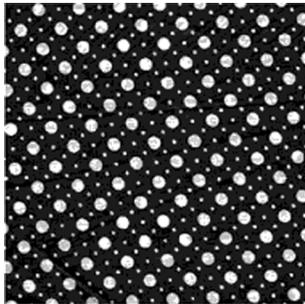
(16a) : Originale



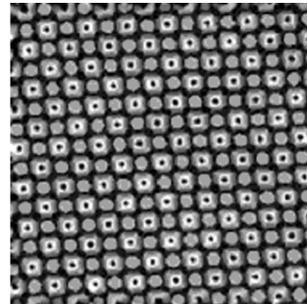
(16b) : Échelles locales



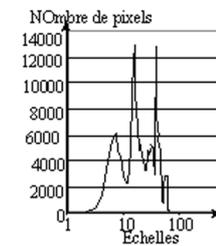
(16c) : Histogramme

(16d) : $\theta = 55^\circ ; \tau = 10^\circ$.

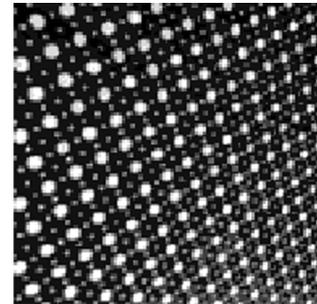
(17a) : Originale



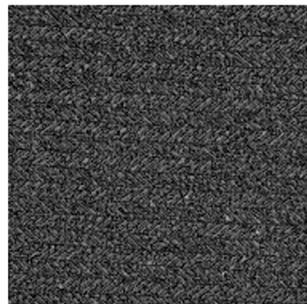
(17b) : Échelles locales



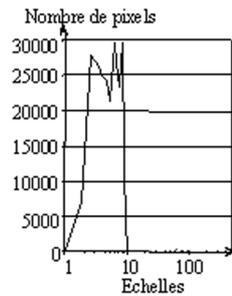
(17c) : Histogramme

(17d) : $\theta = 40^\circ ; \tau = -20^\circ$.

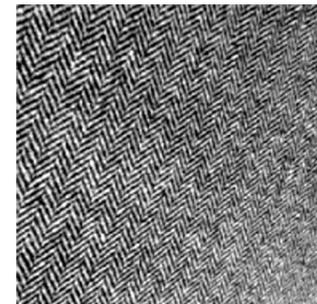
(18a) : Originale



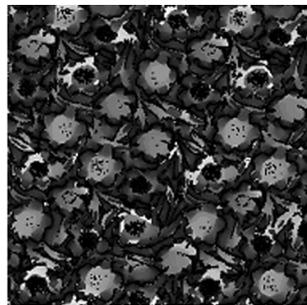
(18b) : Échelles locales



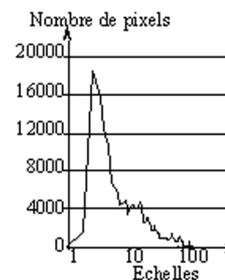
(18c) : Histogramme

(18d) : $\theta = 40^\circ ; \tau = -20^\circ$.

(19a) : Originale



(19b) : Échelles locales



(19c) : Histogramme

(19d) : $\theta = 50^\circ ; \tau = -20^\circ$.

Figures 16, 17, 18 et 19. – Illustration des quatre catégories de textures retenues.

Orientation d'un plan texturé incliné

figures sont organisées de la même manière. En (a) on trouve la texture d'origine scannée à plat, c'est à dire pour un tilt et un slant nuls. La vue (b) montre la carte des échelles locales obtenues à partir de (a). La courbe (c) donne l'histogramme de ces échelles locales (l'abscisse est en échelle logarithmique). Enfin, en (d) on peut voir la texture avec une certaine orientation. Le premier groupe de quatre textures correspond aux conditions idéales d'applications de notre méthode. Ce sont des macrotextures régulières dont la taille des motifs élémentaires et la distance entre ces motifs sont très proches. Une fréquence identique caractérise ces deux phénomènes (figure 16). Sur la figure 17, on voit un élément du deuxième groupe qui est constitué de deux tissus présentant un motif très régulier sur un fond uni. La surface des motifs est petite devant l'espacement entre ceux-ci. Cette construction introduit donc une deuxième composante dans l'histogramme des échelles locales. La troisième catégorie, qui comprend également deux tissus, est illustrée en 18. C'est un cas limite de la méthode proposée. En effet, si ces textures présentent une structure régulière évidente, il est néanmoins difficile d'isoler un motif élémentaire, que l'on pourrait assimiler au texel de base. Cette catégorie est à la frontière entre micro et macrotextures. Enfin, la figure 19 présente une des deux macrotextures de la quatrième catégorie. Cette catégorie a été créée pour vérifier le comportement de notre algorithme sur des structures particulières. En effet, si la taille des fleurs est assez constante, de même que l'espacement entre elles, il est malgré tout difficile de retrouver une organisation régulière de l'ensemble.

Le dispositif de prise de vue est constitué par une caméra noir et blanc fixe positionnée au-dessus de la scène. Une correction de distorsion est appliquée aux images acquises. L'inclinaison due à l'angle de slant a été réalisée au moyen d'un support orientable autour d'un axe. La rotation autour de cet axe peut être bloquée

Tableaux 1. – Comparaison des moyennes et écart-types des résultats obtenus par les deux méthodes.

	Moyenne	Écart-type
Méthode de Lu	12,9	28,4
Notre méthode	10,05	13,31

(1a) : Erreur sur l'angle de tilt

	Moyenne	Écart-type
Méthode de Lu	26	15,9
Notre méthode	13,8	11,26

(1b) : Erreur sur l'angle de slant

pour effectuer les prises de vue et possède une graduation précise au degré. L'obtention de l'angle de tilt désiré s'est faite par une rotation du support sur le plan de travail. La mesure de cet angle a été faite au rapporteur. Sa précision a été estimée à 5°. 90 images ont été ainsi obtenues. Sur toutes ces images nous avons appliqué l'algorithme de Lu et *coll.* et notre technique pour calculer les angles de tilt et de slant. Les tableaux 1a et 1b présentent la comparaison des résultats obtenus pour ces deux angles sur l'ensemble de la base de données.

Ces deux tableaux montrent que notre méthode procure toujours une meilleure précision. Ceci est d'autant plus sensible pour l'angle de slant. De plus, la stabilité de nos résultats est plus importante, ce qui se traduit par un écart-type plus faible. La figure 20 conforte ce résultat en comparant les deux méthodes pour différents angles de slant, en prenant un angle de tilt égal à -90° , valeur retenue car elle correspond à la moitié des prises de vue réalisées.

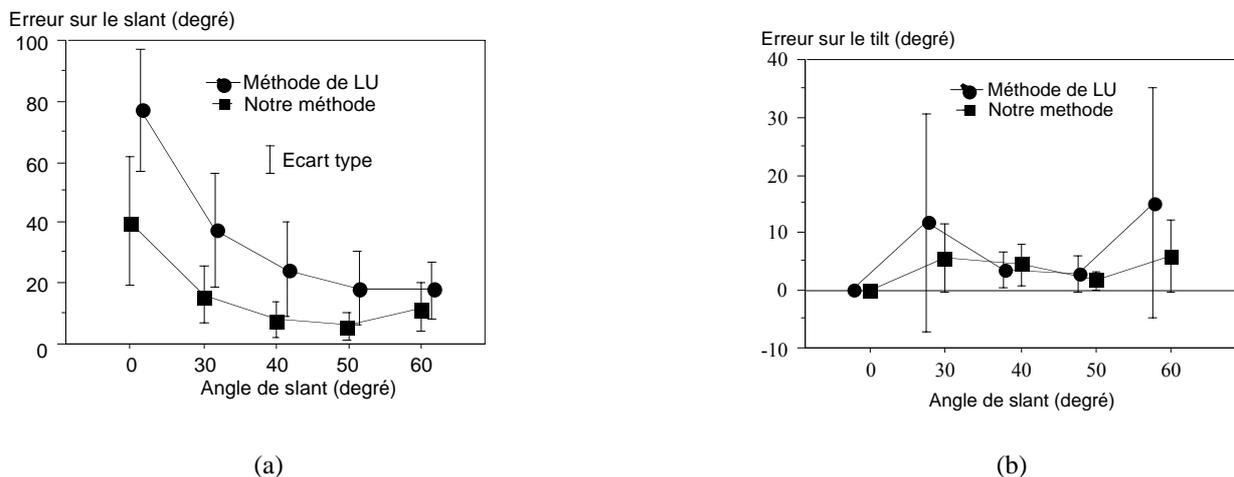
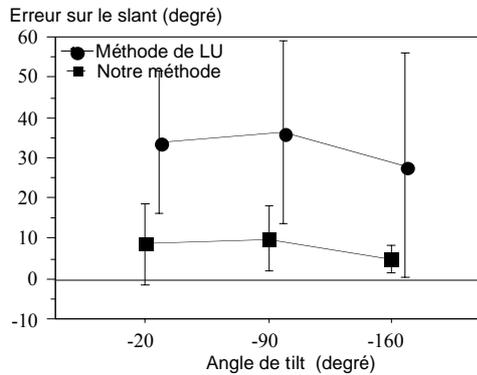
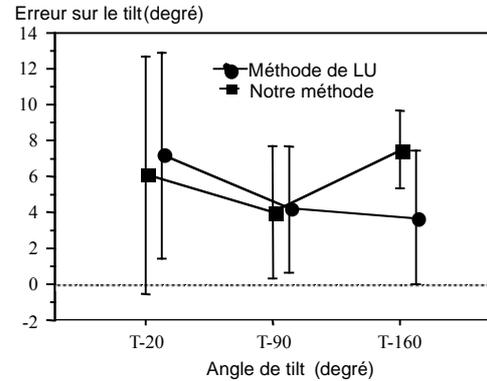


Figure 20. – Comparaison des erreurs commises sur les calculs du slant et du tilt, en fonction de l'angle de slant et pour un tilt = -90° .



(a)



(b)

Figure 21. – Comparaison des erreurs commises sur les calculs du slant et du tilt, en fonction de trois angles de tilt et pour un slant = 40°.

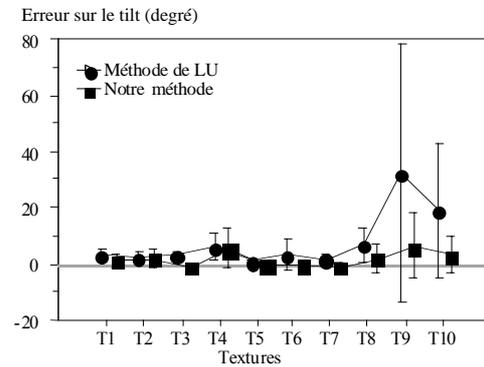
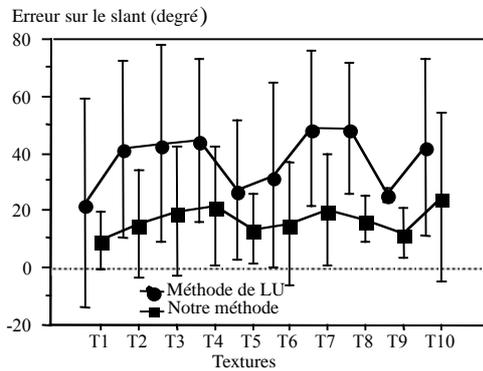


Figure 22 : Comparaison des résultats procurés par les deux méthodes pour le calcul des angles de slant et de tilt à travers les 10 textures sélectionnées.

On constate le comportement analogue des deux techniques, avec une amélioration des résultats au fur et à mesure de la croissance de l'angle de slant, mais on voit bien également que l'erreur obtenue par notre technique est toujours inférieure, ceci étant d'autant plus vrai que l'angle de slant est faible. Nous validons ainsi les améliorations apportées à la méthode de Lu *et coll.* Nous constatons également que, comme pour les images de synthèse, l'erreur commise sur l'angle de slant devient inférieure à 10° lorsque cet angle dépasse une quarantaine de degrés. Ce résultat est vrai pour toutes les images.

La figure 21 présente les erreurs sur le slant et le tilt pour trois valeurs particulières de l'angle de tilt et pour une valeur de θ égale à 40°. Ces résultats sont intéressants car on se place dans un cas favorable pour ces méthodes. On constate nettement l'amélioration que nous avons apportée sur la précision de l'angle de slant, puisque notre méthode conduit toujours à une erreur inférieure à 10°. Pour l'angle de tilt les résultats sont plus homogènes entre les deux algorithmes.

Enfin, la figure 22 propose une analyse des résultats de l'erreur commise sur les angles de slant et de tilt, en fonction des diffé-

rentes textures. Les résultats pris en compte pour tracer ces courbes sont ceux obtenus avec un tilt égal à -90° et des angles de slant de 0°, 30°, 40°, 50° et 60°.

On peut constater que les deux méthodes réagissent de manière très symétrique aux caractéristiques des différentes textures, sauf dans le cas de la catégorie 4 (T9 et T10). En effet, dans ce cas, la méthode de Lu conduit à des erreurs très importantes sur l'angle de tilt.

Pour compléter cette analyse des résultats obtenus selon les différentes textures, nous présentons, dans les tableaux 2a et 2b, le résumé des erreurs pour chacune des catégories. La catégorie 1, qui regroupe des textures monofréquentielles nous donne l'ordre de grandeur des résultats que l'on peut attendre de ces méthodes. Pour chacun des algorithmes et des angles, on retrouve dans la catégorie 2 les résultats les meilleurs en terme de moyenne. Ce résultat peut paraître surprenant car c'est dans cette catégorie que l'on trouve des textures présentant une double fréquence. Cela dit, les motifs de ces textures sont extrêmement réguliers et présentés sur un fond uni. Les variations dues à l'orientation sont donc plus aisément perçues dans un tel contexte. Les deux

Tableaux 2. – Analyse par catégorie de textures des résultats obtenus par les deux méthodes.

Catégorie	Erreur sur le slant (degré)			
	Méthode de LU		Notre méthode	
	Moy.	Écart type	Moy.	Écart type
1	34,9	31,1	16,5	17,5
2	27,3	27,5	14,5	17,1
3	45,9	23	18,4	13,3
4	34,7	26,6	20,9	25,1

(2a) : Erreur sur l'angle de slant

Catégorie	Erreur sur le tilt (degré)			
	Méthode de LU		Notre méthode	
	Moy.	Écart type	Moy.	Écart type
1	3,3	3,1	3,9	4,5
2	2	4	1,2	1,1
3	4	5	2,4	3,9
4	21,8	26,8	6	7,1

(2b) : Erreur sur l'angle de tilt

fréquences sont affectées de façon identique par l'orientation, ce qui perturbe peu l'étape d'interpolation. Pour les catégories 3 et 4, qui contiennent des textures correspondant également à des cas limites d'application (microtextures ou textures aléatoires), notre algorithme obtient malgré tout des résultats assez voisins de ceux obtenus pour la catégorie 1. Pour ce qui est des écarts-type, on peut remarquer que la catégorie 3 (T7 et T8), présente des variations plus faibles pour les résultats de l'angle de slant. Ceci s'explique par la compression des texels distants de la caméra. L'erreur sur le tilt commise sur les images de la catégorie 4 est nettement supérieure aux résultats des autres catégories pour les deux méthodes. L'organisation aléatoire des texels est certainement à l'origine de cette détérioration de la précision. Lors de l'extraction des échelles locales, la non isotropie de la texture va induire un bruit indépendant de l'orientation du plan. L'algorithme de Lu qui suppose une extraction parfaite des échelles locales sera donc très sensible à ce phénomène. Dans notre méthode, la recherche de l'erreur minimum commise sur le tilt peut aboutir à des minima locaux qui sont toutefois peu éloignés de la valeur réelle de l'angle. Ceci explique l'altération moins grande de nos résultats.

Tous ces résultats expérimentaux montrent l'intérêt des améliorations que nous avons apportées permettant ainsi l'application de cette méthode à une variété de textures plus importantes que celles relevant du seul modèle monofréquentiel.

Pour conclure dans cette analyse nous devons signaler que nous avons observé quelques résultats en complet décalage avec la majorité des cas, que ce soit par leur excellente ou très mauvais qualité. Nous n'avons pas d'explications satisfaisantes pour ces phénomènes isolés.

7. conclusion

Dans cet article nous avons proposé une nouvelle méthode pour calculer l'orientation d'un plan incliné texturé. Cette méthode utilise une extraction des échelles locales suivie d'une interpolation de la carte des échelles locales obtenues. Les paramètres de cette interpolation permettent d'obtenir les valeurs des angles de tilt et de slant donnant l'orientation du plan. Une telle méthode permet d'obtenir une précision de l'ordre de quelques degrés sur l'angle de tilt. Par contre, sur l'angle de slant, les erreurs peuvent atteindre 5° et même dépasser 10°, si le plan est faiblement incliné. Nous avons mis en évidence qu'il était indispensable de calibrer la caméra servant à l'acquisition des images afin de limiter l'erreur commise sur l'angle de slant. Nous avons pu également constater que les résultats que pouvaient fournir ces méthodes étaient du même ordre de grandeur que les performances obtenues par le système visuel humain, lorsqu'il est confronté à des situations analogues. Cette constatation est importante car elle permet de valider l'idée qu'un système utilisant de tels algorithmes serait apte à se déplacer dans un environnement comportant des textures orientées. Les estimations théoriques de performances ont été validées par des tests sur des images synthétiques, mais aussi par une expérimentation sur une centaine d'images de textures réelles, vues sous des orientations différentes et acquises au moyen d'une caméra calibrée. La comparaison de notre méthode de calcul à celle développée par Lu *et coll.* a mis en évidence l'intérêt des améliorations réalisées. Nous avons pu également apporter des éléments de réponse concernant l'instabilité des résultats pour les angles de slant faibles. Nous avons vu que la phase de la texture, sa fréquence, l'isotropie... pouvaient expliquer certains résultats, *a priori*, surprenants.

Cette méthode de calcul de l'orientation nous a permis de mettre en évidence l'intérêt des méthodes de « Shape From Texture » pour des applications variées comme la segmentation des images, la caractérisation de la texture, la recherche de défauts [4, 15]. Nous avons également pu montrer qu'il était possible d'améliorer nettement la précision sur l'angle de slant pour des macrottextures régulières en combinant notre méthode initiale à un calcul de la ligne de fuite de l'image [5].

Des travaux complémentaires pourraient permettre d'augmenter la compréhension de ces méthodes et d'en étendre leur portée, comme le test d'ondelettes analysantes différentes des DOG, l'étude de textures multifréquentielles et l'analyse d'autres surfaces que le plan, telles des sphères ou des cylindres.

BIBLIOGRAPHIE

- [1] J. Aloimonos, « Shape from texture », *Biological Cybernetics*, Vol. 58, pp. 345-360, 1988.
- [2] R. Bajcsy, « Texture gradient as a depth cue », *Computer Graphics and Image Processing*, Vol. 5, pp.52-67, 1976.
- [3] D. Blostein, N. Ahuja, « Shape from texture : integrating texture-element extraction and surface estimation », *IEEE PAMI*, Vol. 11, n°12, pp.1233-1251, 1989.
- [4] L. Boutte, J. Plantier, S. Lelandais, P. Checchin, « Texel extraction on inclined textures by adaptive thresholding of local scales », *ICPR 2000*, pp. 560-563, Vol. 3, septembre 2000, Barcelone, Espagne.
- [5] L. Boutte, « Extraction des échelles locales appliquée au calcul de l'orientation d'un plan texturé », Doctorat de robotique, Université d'Évry Val d'ESSONNE, novembre 2000.
- [6] J. Brochard, C. Lecomte, M. Leard, « Analyse de texture sous divers facteurs d'échelle : applications à la reconnaissance de textures et à la mesure de profondeurs », *Congrès RFIA'98*, pp. 153-159, janvier 1998, Clermont-Ferrand, France.
- [7] D. Duval-Destin, « Analyse spatiale et spatio-temporelle de la stimulation visuelle à l'aide de la transformée en ondelettes », Doctorat de physique théorique, Université des sciences de Luminy, Marseille, 1991.
- [8] M. El Ghadi, A. Guerin-Dugue, « Forme des objets texturés par estimation des fréquences locales », *Congrès ORASIS'99*, pp. 181-189, avril 1999, France.
- [9] C. Enroth-Cugell, J.-G. Robson, « The contrast sensitivity of retinal ganglion cells in the cat », *J. of Physiology*, London, Vol. 187, pp. 517-552, 1966.
- [10] J. Garding, « Direct estimation of shape from texture », *IEEE PAMI*, Vol. 15, n°11, pp. 1202-1208, Novembre 1993.
- [11] K. Kanatani, T. Chou, « Shape from texture: general principle », *Artificial Intelligence*, Vol. 38, n°1, pp. 1-48, 1989.
- [12] D.-C. Knill, « Surface orientation from texture: ideal observers, generic observers and the information content of textures cues », *Vision Research*, Vol. 38, n°11, pp. 1655-1682, juin 1998.
- [13] J. Krumm, S. Shafer, « Local Spatial frequency analysis of image texture », *3rd International Conference on Computer Vision*, pp. 354-358, Mai 1990.
- [14] J. Krumm, S. Shafer, « Segmenting textured 3D surfaces using the space frequency representation », *Spatial Vision*, Vol. 8, n°2, pp. 281-308, 1994.
- [15] S. Lelandais, L. Boutte, J. Plantier, « Defects detection on inclined textured planes », *QCAV 2001*, Le Creusot, France, 21-23 mai 2001.
- [16] T. Lindeberg, J. Garding, « Shape from texture from a multi-scale perspective », *4th International Conference on Computer Vision*, pp. 683-690, juillet 1993, Berlin, Allemagne.
- [17] C.-S. LU, W.-L. Hwang, H.-Y. Mark Lio, P.-C. Chung, « Shape from texture based on the ridge of continuous wavelet transform », *IEEE International Conference on Image Processing*, septembre 1996, Suisse.
- [18] D. Marr, « Vision », *Freemans Editors*, New York, 1982.
- [19] Y. Ohta, K. Moenobu, T. Sakai, « Obtaining surface orientation from texels under perspective projection », *7th IJCAI*, pp. 746-751, 1981, Vancouver, Canada.
- [20] J. Plantier, J.-P. Menu, « Analyse des contrastes locaux de luminance dans des images complexes », *Rapport CERMA N°93-12*, 1993, France.
- [21] E. Ribiero, E.-R. Hancock, « Estimating the 3D orientation of texture planes using local spectral analysis », *Image and Vision Computing*, Vol. 18, n°8, pp. 619-631, 2000.
- [22] K. Sakai, L.-H. Finkel, « A Shape-From-Texture algorithm based on human visual psychophysics », *Congrès CVPR'94*, pp. 527-532, 1994.
- [23] K. Sakai, L.-H. Finkel, « Characterization of the spatial-frequency spectrum in the perception of shape from texture », *Journal of Opt. Soc. Am. A.*, Vol. 12, n°6, pp. 1208-1224, 1995.
- [24] B. Super, A. Bovik, « Shape From Texture using local spectral moments », *IEEE PAMI*, Vol. 17, n°4, pp. 333-343, avril 1995.
- [25] F. Tomita, S. Tsuji, « Computer analysis of visual textures », *Kluwer Academic Press*, Boston, USA, 1990.

Manuscrit reçu le 12 avril 2001

LES AUTEURS

Sylvie LELANDAIS



Sylvie LELANDAIS a obtenu un Doctorat en Automatique et Traitement du Signal à l'Université Technologique de Compiègne en 1984. Sa thèse portait sur la numérisation 3-D. De 1985 à 1990, elle a travaillé dans le Laboratoire Traitement d'Images de l'IRESTE, Ecole d'Ingénieur à Nantes. Actuellement, elle est dans le Laboratoire Systèmes Complexes de l'Université d'Évry. Ses recherches s'inscrivent dans les domaines suivants :

d'images, analyse de texture, ondelettes, vision en robotique, imagerie couleur et traitement de données.

Laurent BOUTTÉ



Laurent BOUTTÉ a obtenu un Doctorat en Robotique à l'Université d'Évry-Val d'Essonne en 2000. Le sujet de sa thèse était la reconstruction 3-D de surfaces texturées utilisant uniquement des informations monoculaires. Il est actuellement ingénieur de développement dans les systèmes embarqués et il travaille particulièrement pour la TV numérique.

Justin PLANTIER



Justin PLANTIER a obtenu un Doctorat Spécialité Instrumentation et Mesures Optiques à l'Institut National Polytechnique de Grenoble (INPG), en 1985. Actuellement, il est Ingénieur de Recherche à l'IMASSA (Institut de Médecine Aérospatiale du Service de Santé). Ses domaines de recherches sont l'analyse d'image et la modélisation de la perception visuelle humaine en particulier la vision des contrastes, la perception des formes et la vision 3-D.