

# Une fonction de corrélation localisée pour la mise en correspondance des images stéréoscopiques

## A localized correlation function for stereoscopic image matching

par Jérôme PERRIN, Bruno TORRÉSANI et Philippe FUCHS

Centre de Robotique de l'École Nationale Supérieure des Mines de Paris,  
60 bd Saint-Michel, 75006 PARIS, {jerome,fuchs}@caor.ensmp.fr  
Laboratoire d'Analyse, Topologie et Probabilités, CMI, Université de Provence,  
39 rue Joliot-Curie, 13453 Marseille Cedex, France, Bruno.Torresani@sophia.inria.fr

### *résumé et mots clés*

Nous proposons et étudions une fonction de corrélation localisée permettant la mise en correspondance d'images stéréoscopiques. Celle-ci est fondée sur la décomposition en échelles des images d'entrée. À l'inverse des algorithmes de type « coarse-to-fine », celui-ci traite simultanément les informations des différentes bandes de fréquence. Pour cela, nous définissons en chaque point des deux images une fonction de corrélation localisée par intégration sur le paramètre d'échelle, dont nous montrons qu'elle est équivalente à la définition d'un noyau de corrélation extrêmement fin dans les paramètres de position et de disparité. La définition peut ensuite être modifiée pour prendre en compte la composition fréquentielle locale des images d'entrées. Enfin, nous nous intéressons au problème de la normalisation préalable des images à apparier et justifions le choix du contraste multirésolution associé à une normalisation quadratique.

Vision stéréoscopique, appariement, corrélation, transformation en ondelettes, analyse d'images.

### *abstract and key words*

We propose and study a localized correlation function for stereoscopic matching. The latter is based on the wavelet decomposition of the input images. Contrarily to « coarse-to-fine » algorithms, this one simultaneously processes information at the different scales. The localized correlation function is defined by locally integrating with respect to the scale variable. We show that it is equivalent to the definition of a correlation kernel, which is extremely precise in terms of position and disparity. The definition can then be modified in order to account for the local frequency content of the images. Then, we suggest pre-processings of the images : we argue in favour of the use of multiresolution contrast techniques, associated to a quadratic normalization.

Stereoscopic vision, matching, correlation, wavelet transform, image analysis.

## 1. introduction

Le but de la *réalité virtuelle* est de permettre à toute personne des expériences sensorielles et cognitives dans un monde artificiel, qui peut être réaliste (dans le cas de la téléopération, par exemple) ou

imaginaire. Pour permettre à l'utilisateur une immersion naturelle dans cet environnement, il est nécessaire de créer une interaction en temps réel à l'aide d'interfaces comportementales sans entraver la naturalité du monde virtuel et l'immersion de l'utilisateur. Les écrans stéréoscopiques, particulièrement utilisés en téléopération, sont le noyau central de ces nouvelles interfaces, de par leur capacité à reproduire un monde visuel tridimensionnel.

Toutefois, ces interfaces modifient les modalités de la perception visuelle et nécessitent d'être contrôlées, soit au niveau matériel, soit au niveau logiciel, pour s'assurer de l'ergonomie du système, notamment par une limitation de la profondeur de champ. Mais, limiter la profondeur de champ, au moins du point de vue logiciel, nécessite de connaître la profondeur de chaque point des images restituées. Partant de ce besoin de la réalité virtuelle, nous nous sommes intéressés au problème de l'appariement stéréoscopique (ou mise en correspondance).

De nombreuses approches ont été proposées pour le résoudre, basées sur différentes stratégies, et utilisant différentes procédures algorithmiques. Une classification de ces méthodes est nécessairement quelque peu arbitraire. En les regroupant par stratégie générale, nous mentionnerons : l'appariement de caractéristiques (points riches d'informations [Gri81, Mar82] ou primitives de type coins, régions,...), l'appariement global (mise en correspondance de chaque point des deux images par minimisation d'une fonctionnelle globale et des techniques de moindres carrés, ou d'optimisation stochastique [OLL96],...), l'appariement « orienté objets » (utilisant un modèle *a priori* de la scène [Hel88] et prolongé par une recherche dans des graphes d'appariement) et l'appariement par corrélation (mettant en correspondance des régions de l'image). Par souci de mettre en cohérence notre démarche d'adaptation de l'interface stéréoscopique avec le fonctionnement du système visuel, il était logique de fonder notre approche du problème sur des études psychophysiques chez l'homme.

La mise en évidence de la possibilité de fusion des stéréogrammes de points aléatoires [Jul71], ainsi que d'autres études psychophysiques chez l'homme [CSS91] laissent penser que dans les stades précoces du traitement de l'information visuelle, la corrélation des informations de chaque œil pourrait être utilisée. Cet aspect a motivé plusieurs travaux sur les méthodes de corrélation, qu'elles soient linéaires [Nis87] ou non-linéaires [Fle94, LNN94]. Mais comme nous le verrons plus loin, l'information obtenue par corrélation est une information globale, c'est-à-dire qu'elle renseigne sur la translation d'une région entière de l'image. Pour localiser l'information, les techniques usuelles travaillent avec des sous-régions de l'image de tailles décroissantes en reprenant les idées d'appariement *coarse-to-fine* de [Mar82]. Elles prennent ainsi en compte un autre aspect de la vision humaine : le traitement *multiéchelle*.

En effet, après les premières mises en évidence d'une décomposition fréquentielle de l'information visuelle chez l'homme à la fin des années 60 [CR68], faisant suite à la théorie des champs récepteurs [ECR66], fut développée l'analyse en canaux correspondant approximativement à la notion d'espace-échelle, qui explique l'intérêt porté à la décomposition en ondelettes dans le domaine de la vision [Mar82, Mal89, DD91]. Mais, paradoxalement, les propriétés de localisation des décompositions temps-échelle n'ont que rarement été exploitées pour donner une version localisée de la fonction de corrélation utilisant *simultanément* les informations aux différentes échelles [JonMal91]. Pourtant, certains auteurs

ont remis en cause par des expériences de psychophysique l'hypothèse de traitement binoculaire hiérarchique [Sma95, Mal96]. Par ailleurs, l'interdépendance des canaux fréquentiels dans le système visuel a été mise en évidence par Roumes [Rou95] et Perrin [Per98]; enfin, d'un point de vue algorithmique, de par sa sequentialité, ce modèle n'est pas optimal.

Nous introduisons donc ici une fonction de corrélation localisée, construite à partir des décompositions en ondelettes d'images stéréoscopiques. Cette fonction, obtenue par réduction d'une formule de transformation en ondelettes inverse, présente l'avantage de ne pas privilégier d'échelle particulière dans les images. En tant que telle, elle fournit une alternative intéressante aux fonctions de corrélation par blocs couramment utilisées. Nous montrons que la fonction de corrélation localisée permet une mise en correspondance des images stéréoscopiques relativement fiable, *via* un simple algorithme d'optimisation locale. Nous illustrons la discussion par des résultats obtenus sur des images de synthèse (figures 1).

Afin de simplifier la présentation, la fonction de corrélation localisée est dans un premier temps construite à partir de la transformation en ondelettes. Cependant, la fonction de corrélation localisée s'adapte tout aussi bien à des variantes de l'analyse par ondelettes, et en particulier à une variante non-linéaire appelée analyse par contraste multirésolution, forme « multiplicative » de l'analyse par ondelettes. Le contraste multirésolution permet de réhausser le contraste des images de façon naturelle, tant du point de vue mathématique [DD91, DDMT93] que du point de vue des données psychophysiques sur la perception du contraste [DD91].

Il va de soi que la fonction de corrélation localisée ne fournit pas en elle-même un algorithme d'appariement, pas plus que ne le ferait la corrélation par blocs sans pré-traitement ni post-traitement additionnel. Cependant, elle fournit une alternative intéressante aux corrélations par blocs, en ne privilégiant aucune échelle particulière. Plus encore que ces dernières, elle peut s'intégrer à une chaîne complète de traitements, incluant en particulier des post-traitements destinés à « régulariser » la solution ou des filtres liés à la restitution stéréoscopique [Per98]. L'objectif du présent article étant de se focaliser sur la fonction de corrélation localisée, nous avons délibérément choisi de limiter au maximum les post-traitements.

## 2. rappels mathématiques

Bien que les images à traiter soient par nature des signaux discrets (et même de taille finie), nous limitons notre présentation à des fonctions d'une ou de deux variables continues, dans le but d'alléger les notations. L'analyse par ondelettes permet de donner une justification rigoureuse à cette approche, les échantillons de l'image pouvant toujours être vus comme des valeurs ponctuelles

d'une version lissée de la fonction (voir [Mal89] par exemple). Dans toute la suite, les signaux considérés sont donc des fonctions de  $\mathcal{L}^2(\mathbb{R})$  ou de  $\mathcal{L}^2(\mathbb{R}^2)$ . Pour simplifier de nouveau, nous ne les considérons que comme des fonctions d'une variable, notée  $t$ . L'extension aux signaux de deux variables (telles que les images) se fait aisément en considérant une variable vectorielle. De plus, nous considérerons ces signaux comme des fonctions à valeurs réelles, même si, pour donner l'expression dans le cas de fonctions à valeurs complexes, nous gardons les conjugués complexes.

## 2.1. corrélation, transformation de Fourier

### 2.1.1. la corrélation

L'estimation d'un paramètre de décalage entre deux signaux est généralement effectuée au moyen de méthodes de corrélation. Étant données deux fonctions  $f(t)$  et  $g(t) = f(t + \tau) + b(t)$ , la seconde étant une copie translatée (de  $\tau$ ) et bruitée (par un bruit noté ici  $b(t)$ ) de la première, l'estimation de  $\tau$  peut se faire en minimisant par rapport à  $\delta$  une norme bien choisie de la différence entre  $g(t)$  et  $f(t + \delta)$  :

$$\min_{\delta} \|g - f(\cdot + \delta)\|^2 \quad (1)$$

On choisit généralement la norme de  $\mathcal{L}^2(\mathbb{R})$  :  $\|f\|^2 = \int |f(t)|^2 dt$ , auquel cas le problème de minimisation ci-dessus se ramène à maximiser la quantité

$$h(\delta) = (f \star g)(\delta) = \int f(t)g(t + \delta) dt. \quad (2)$$

Le symbole  $\star$  représente la corrélation des fonctions. De plus, si le bruit est blanc, l'estimation ainsi obtenue est optimale.

Il est bien connu que la transformation de Fourier  $\mathcal{F} : f \in \mathcal{L}^2(\mathbb{R}) \mapsto \mathcal{F}[f] = \hat{f} \in \mathcal{L}^2(\mathbb{R})$ , définie par

$$\hat{f}(\omega) = \mathcal{F}[f](\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t)e^{-j\omega t} dt, \quad (3)$$

transforme la corrélation en produit simple :

$$f \star g = \sqrt{2\pi} \mathcal{F}^{-1} \left( \mathcal{F}[f] \cdot \overline{\mathcal{F}[g]} \right). \quad (4)$$

### 2.1.2. applications à la mise en correspondance

La mise en correspondance d'images stéréoscopiques est bien entendu plus complexe que le problème évoqué ci-dessus, car le décalage  $\delta$  varie d'un point à l'autre. Il est donc nécessaire d'effectuer des corrélations de façon plus locale. Plusieurs algorithmes ont été proposés utilisant la corrélation pour effectuer des appariements stéréoscopiques [Mar82, Nis87]. Leurs approches

se classent dans la catégorie des algorithmes dits *coarse-to-fine*, où l'information utilisée est issue de filtrages des images initiales à des fréquences de plus en plus hautes. Elles se différencient fondamentalement de certaines approches ultérieures supposant un traitement simultané des différentes fréquences comme celle proposée par Fleet [Fle94] où le produit de corrélation utilisé est la corrélation de phase (la TF de la fonction de corrélation est divisée par son module), équivalente à un « vote dans lequel chaque fréquence vote de manière sinusoïdale et où la localisation des pics est déterminée par des différences de phase [entre les signaux] » [Fle94].

Dans l'analyse des images de corrélation, le problème principal reste celui de la localisation. Si une image de corrélation donne une mesure quantitative de la similitude entre un signal et les traduits d'un second signal, elle ne donne aucune indication sur les régions qui ont motivé une forte valeur de la fonction de corrélation (s'il existe des méthodes de *rétro-projection de la fonction de corrélation*, elles sont assez lentes et peu précises).

La solution utilisée couramment pour localiser les régions fortement corrélées entre elles consiste à découper l'image en im-  
a-  
gettes, puis à effectuer une corrélation (linéaire, de phase,...) de chacune des im-  
a-  
gettes avec la seconde image. Utilisant ces dis-  
a-  
parités *a priori*, la méthode est alors récursivement appliquée à des im-  
a-  
gettes plus petites. Le résultat est néanmoins sensible à la taille des im-  
a-  
gettes utilisées, effet qui est corrigé par un post-  
a-  
traitement approprié.

## 2.2. représentation en ondelettes

Les décompositions en ondelettes ou *temps-échelle* proposent un compromis entre les représentations temporelles et fréquentielles. Leur caractéristique fondamentale est d'adapter la taille de la fenêtre temporelle d'analyse à la période caractéristique des signaux analysés en décomposant le signal sur une famille de fonctions obtenues par dilatation et translation<sup>1</sup> d'une fonction de base appelée ondelette-mère. En notant  $\psi$  l'ondelette-mère, les fonctions de la famille sont :

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi \left( \frac{t-b}{a} \right) \quad b \in \mathbb{R}, a \neq 0, \quad (5)$$

où  $a$  est un indice d'échelle, correspondant à la taille des détails analysés, et  $b$  un indice de localisation.

Les coefficients du signal sont alors les nombres

$$\mathcal{O}[f](a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{+\infty} f(t) \bar{\psi}_{a,b}(t) dt. \quad (6)$$

1. En dimension 1; pour les dimensions supérieures, d'autres transformations interviennent qui ne sont pas l'objet du propos développé ici et pour lesquelles nous renvoyons à [Mal89] par exemple.

## Une fonction de corrélation localisée

Si la fonction  $\psi \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$  est telle que la constante  $C_\psi = \int \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega$  vérifie  $0 < C_\psi < +\infty$ , alors le signal peut être reconstruit à partir de ses coefficients par la formule

$$f(t) = \frac{1}{C_\psi} \int \int_{\mathbb{R}^2} \mathcal{O}[f](a, b) \psi_{a,b}(t) \frac{da db}{a^2}. \quad (7)$$

Un corollaire important de cette formule est l'existence d'un analogue de la formule de Parseval-Plancherel : pour toutes fonctions  $f, g \in \mathcal{L}^2(\mathbb{R})$ , on a

$$\int f(t) \overline{g(t)} dt = \frac{1}{C_\psi} \int \mathcal{O}[f](a, b) \overline{\mathcal{O}[g](a, b)} \frac{da db}{a}. \quad (8)$$

Il existe en fait une infinité de formules de reconstruction; en particulier, si la fonction  $\psi$  est telle que la constante  $k_\psi = \int \frac{\hat{\psi}(\omega)}{\omega} d\omega$  vérifie  $0 < |k_\psi| < +\infty$ , nous obtenons alors

$$f(t) = \frac{1}{k_\psi} \int_0^{+\infty} \mathcal{O}[f](a, t) \frac{da}{a^{3/2}}. \quad (9)$$

C'est cette dernière version que nous utiliserons par la suite. Nous supposons que l'inégalité  $0 < |k_\psi| < +\infty$  est vérifiée et normalisons l'ondelette  $\psi(t)$  en posant  $k_\psi = 1$ .

### Remarque

Nous n'utiliserons jamais explicitement les formules d'inversion de la transformation en ondelettes que nous venons d'expliquer. Les inégalités  $0 < |k_\psi| < +\infty$  (ou  $0 < C_\psi < +\infty$  dans l'autre cas) sont toutefois importantes, dans la mesure où elles imposent que l'ondelette  $\psi(t)$  soit une fonction d'intégrale nulle. Par conséquent, les coefficients d'ondelettes  $\mathcal{O}[f](a, b)$  sont insensibles à la valeur moyenne de la fonction  $f(t)$  : pour toute constante  $C$ , les fonctions  $f(t)$  et  $f(t) + C$  ont mêmes coefficients d'ondelettes. Nous verrons plus loin en quoi ce point concourt à la robustesse de la méthode que nous présentons.

L'ondelette-mère est généralement choisie de manière à être relativement localisée à la fois dans le temps et dans l'espace des fréquences. Les fonctions couramment utilisées, notamment pour des modélisations psychophysiques et en traitement d'images, sont les dérivées de gaussienne (en particulier le laplacien de gaussienne, LOG), des différences de gaussiennes (DOGs) ainsi que les fonctions de Gabor.

## 3. localiser la corrélation

### 3.1. définition

L'ensemble des coefficients  $(\mathcal{O}[f](a, b))_{a>0}$  de la décomposition en ondelettes de  $f$  en  $b$  représente un vecteur de description

du signal en  $b$ . La décomposition en ondelettes nous fournit alors un champ de vecteurs : chaque vecteur peut être interprété comme une variable aléatoire, relativement décorrélée des autres variables. Le problème de la mise en correspondance se ramène à la détermination, pour chaque vecteur d'une image, de la variable aléatoire de l'autre image la plus proche pour une statistique donnée. Le produit scalaire est une statistique adaptée à des données vectorielles.

Soient alors  $f$  et  $g$  deux fonctions de  $\mathcal{L}^2(\mathbb{R})$ . Une conséquence de la formule de Parseval-Plancherel généralisée (8) est que l'on peut écrire la fonction de corrélation  $f \star g$  sous la forme

$$(f \star g)(\delta) = \int \mathcal{O}[f](a, b) \overline{\mathcal{O}[g](a, b + \delta)} \frac{da db}{a}. \quad (10)$$

Ceci suggère de « localiser » la fonction de corrélation en supprimant l'intégration par rapport à  $b$  : nous définissons la *fonction de corrélation localisée*  $f \star^{\text{loc}} g$  de  $f(t)$  et  $g(t)$  par

$$s(b, b') = (f \star^{\text{loc}} g)(b, b') = \int \mathcal{O}[f](a, b) \overline{\mathcal{O}[g](a, b')} \frac{da}{a}, \quad (11)$$

où nous retrouvons bien le produit scalaire évoqué plus haut.

Si la fonction  $s$  est localisée, dans le sens où elle mesure l'affinité des signaux  $f$  au point  $b$  et  $g$  au point  $b'$ , elle n'est pas pour autant purement locale, au sens où elle ne prendrait en compte que les caractéristiques ponctuelles de l'image. En effet, de la définition même de la représentation en ondelettes, les différents coefficients  $\mathcal{O}[f](a, b)$  décrivent des régions de l'image de taille proportionnelle à  $a$ . Ainsi, à travers l'intégrale décrite par (11), nous prenons en compte des informations concernant les caractéristiques des voisinages de  $b$  et  $b'$  à des tailles différentes.

Notons aussi que ce choix ne privilégie *a priori* aucune échelle dans le signal : toutes les échelles sont prises en compte également. Ceci n'est pas le cas des approches de corrélations par blocs, dans lesquelles l'échelle des blocs est privilégiée (effet qui est partiellement corrigé en considérant des blocs de tailles différentes, dans un algorithme hiérarchique).

### 3.2. noyau de corrélation

A partir des définitions intégrales de  $\mathcal{O}[f](a, b)$  et  $\mathcal{O}[g](a, b')$ ,  $s(b, b')$  peut se réécrire sous la forme

$$s(b, b') = \int f(x) \bar{g}(y) K_{b,b'}(x, y) dx dy, \quad (12)$$

où le noyau de corrélation  $K_{b,b'}$  est défini par

$$K_{b,b'}(x, y) = \mathcal{K}(x - b, y - b'), \quad \text{où} \quad (13)$$

$$\mathcal{K}(x, y) = \int \bar{\psi}\left(\frac{x}{a}\right) \psi\left(\frac{y}{a}\right) \frac{da}{a^2}. \quad (14)$$

Notons une différence fondamentale entre le noyau défini en (14) et celui de la corrélation usuelle :

– dans le cas de la corrélation usuelle, le noyau est une fonction de la différence  $b' - b$  puisqu'il s'agit de la distribution de Dirac en  $b' - b$ ,  $\delta_{(b'-b)}(y - x)$ ;

– dans le cas présent, le noyau de corrélation est une fonction de  $b$  et  $b'$  et non de leur différence. C'est précisément ce qui permet à cette fonction de corrélation d'être localisée, c'est-à-dire de mesurer la similarité de  $f$  et  $g$  autour de  $b$  et  $b'$  respectivement, et non de mesurer les similarités de  $f$  et de la translatée de  $g$  par la quantité  $b' - b$ .

Quelques exemples de noyaux de corrélation sont donnés plus loin (§ 3.4) qui permettent de mieux comprendre la spécificité du noyau utilisé dans la corrélation localisée.

### 3.3. adaptabilité du noyau en échelle

La définition du noyau (14), de même que la définition de la fonction de corrélation (11), suggèrent de limiter la borne d'intégration après avoir défini l'échelle adaptée à la structure du signal au point considéré : en effet, si l'intégrale sur les échelles  $a$  s'effectue de 0 à  $+\infty$ , nous prenons implicitement en compte des régions de taille plus petites que les structures locales du signal, qui peuvent être assimilées à des bruits. Limiter la borne d'intégration inférieure à une valeur  $A > 0$  est équivalent à un débruitage des signaux à mettre en correspondance. De plus, dans le cas de signaux échantillonnés comme ceux que nous traitons, nous sommes limités dans la gamme des fréquences traitées du fait de l'échantillonnage : chercher à intégrer jusqu'à l'échelle  $a = 0$  reviendrait à introduire du repliement de spectre sur les images.

Pour prendre en compte la translation d'un objet de taille donnée  $A$ , il convient de limiter l'intégration à l'échelle  $A$ . Nous obtenons alors un noyau présentant une grande possibilité d'adaptation au signal puisqu'il est spécifique du point considéré, du retard et de l'échelle locale :

$$K_{A,b,b'}(x, y) = \int_A^{+\infty} \bar{\psi} \left( \frac{x - b}{a} \right) \psi \left( \frac{y - b'}{a} \right) \frac{da}{a^2}. \quad (15)$$

La détermination en chaque point de l'échelle adaptée  $A(b)$  s'effectue à partir des coefficients  $\mathcal{O}[f](a, b)$ . Nous considérons que la taille des objets occupant localement l'image correspond à l'échelle possédant le plus grand coefficient d'ondelettes en module (i.e. à la bande de fréquences de plus haute énergie). Cette définition correspond qualitativement à ce que nous pouvons entendre par *contenu local des signaux traités*. L'échelle adaptée  $A(b)$  est alors l'échelle telle que l'énergie de la décomposition en ondelettes réduite  $\tilde{\mathcal{O}}$  soit une fraction de l'énergie du signal initial, par exemple 80% de l'énergie initiale.

Nous noterons  $\tilde{\mathcal{O}}[f](a, b)$  la décomposition en ondelettes réduite de  $f$  définie par

$$\tilde{\mathcal{O}}[f](a, b) = \begin{cases} \mathcal{O}[f](a, b) & \text{si } a \geq A(b), \\ 0 & \text{si } a < A(b). \end{cases} \quad (16)$$

Un exemple de détermination des échelles locales est donné par la figure 3 : il est obtenu à partir de l'image gauche de la paire stéréoscopique présentée sur la figure 1.

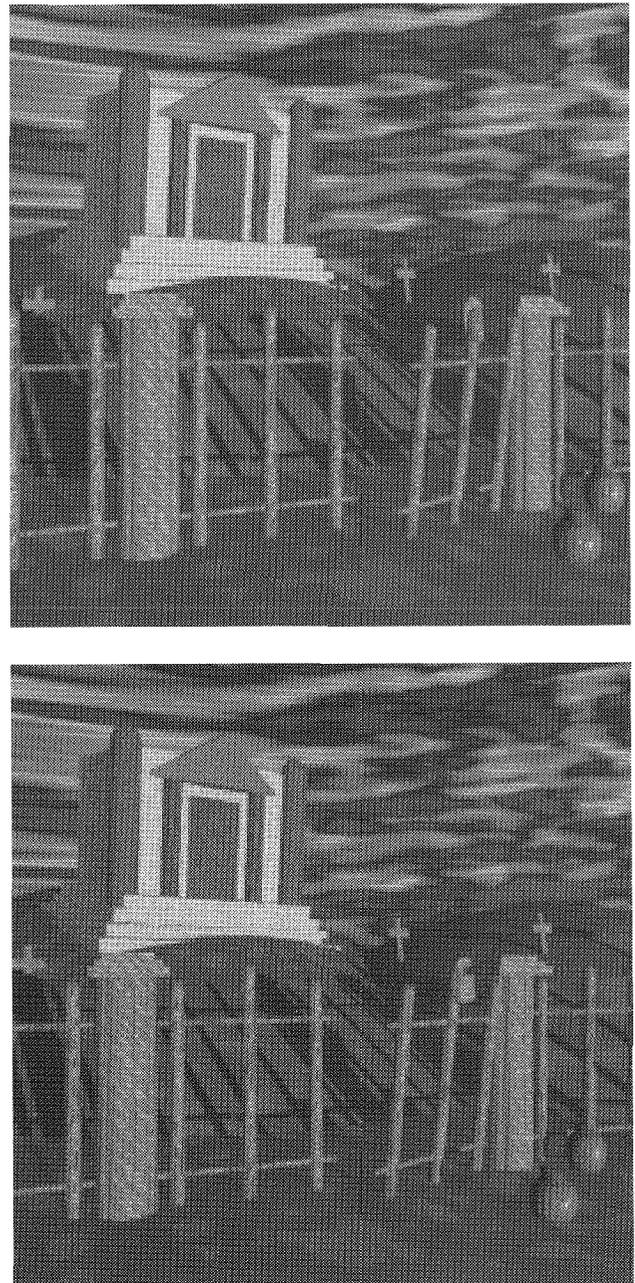


Figure 1. – Images stéréoscopiques utilisées pour tester l'algorithme de corrélation localisée.

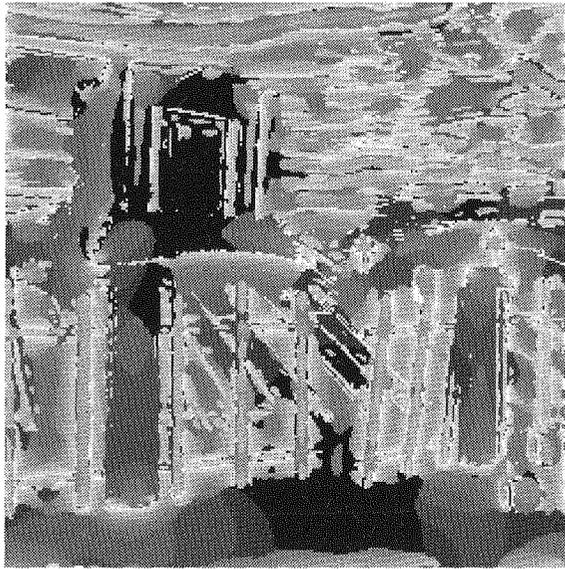


Figure 2. – Détermination des échelles locales par détection du coefficient de décomposition maximal; en noir, les échelles locales les plus frustes, en blanc, les plus fines.

### 3.4. exemples de noyau de corrélation

Les noyaux de corrélation que nous considérons par la suite sont obtenus à partir d'ondelettes de type « dérivée  $n$ -ième d'une fonction de lissage  $\phi$  », car l'ordre de l'ondelette (qui contrôle la localisation en fréquence) peut être facilement déterminé.

Par ailleurs, pour des raisons de clarté, les expressions ne sont données ici que pour les échelles adaptées  $A = 0$ . Deux types de fonctions sont considérés : les dérivées de Gaussienne et les fonctions de Cauchy. Les noyaux obtenus pour ces deux types d'ondelettes sont représentés sur les figures 3(a) et 3(b) respectivement.

#### 3.4.1. ondelette mère « dérivée de Gaussienne »

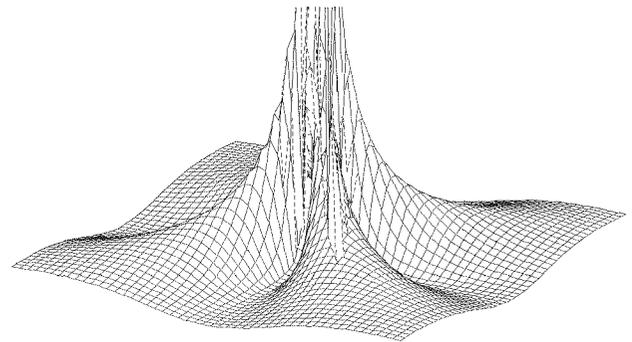
L'ondelette mère et sa TF sont de la forme

$$\begin{aligned} \psi(x) &= \frac{d^n}{dx^n} e^{-x^2/2}, \quad n \in \mathbb{N} \\ \hat{\psi}(\omega) &= (j\omega)^n e^{-\omega^2/2}. \end{aligned}$$

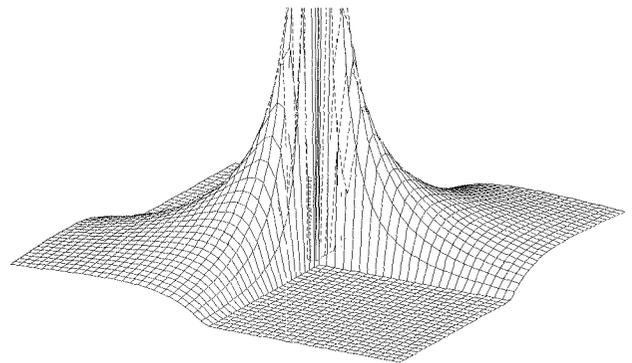
Le noyau  $\mathcal{K}(x, y)$  est défini dans le cas  $n = 2$  par

$$\mathcal{K}^2(x, y) = \frac{C x^2 y^2}{(x^2 + y^2)^{5/2}}, \quad (17)$$

avec  $C = 3\sqrt{\pi/2}$ . Il est représenté sur la figure 3a.



(a) noyau obtenu avec une ondelette LOG



(b) noyau obtenu avec une fonction de Cauchy

Figure 3. – Noyaux obtenus pour des ondelettes de type laplacien de gaussienne ou fonction de Cauchy : les paramètres  $x$  et  $y$  évoluent de  $-4$  à  $4$ .

#### 3.4.2. ondelette mère « fonction de Cauchy »

L'ondelette mère et sa TF sont de la forme

$$\begin{aligned} \psi(x) &= \frac{d^n}{dx^n} \left( \frac{1}{x + j} \right), \quad n \in \mathbb{N} \\ \hat{\psi}(\omega) &= (j\omega)^n e^{-\omega} H(\omega), \end{aligned}$$

où  $H$  est la fonction de Heaviside.

Le noyau  $\mathcal{K}(x, y)$  est alors défini dans le cas  $n = 1$  par

$$\begin{aligned} \mathcal{K}^1(x, y) &= \frac{xy}{(x + y)^3} \\ &\quad \left[ 2\pi \mathbb{1}^+(x, y) + j \left( \log \frac{y^2}{x^2} - y^2 + x^2 \right) \right], \quad (18) \end{aligned}$$

avec  $\mathbb{1}^+(x, y) = (\text{Signe}(x) + \text{Signe}(y))/2$ . Il est représenté sur la figure 3b.

#### 3.4.3. remarques

Nous pouvons remarquer, particulièrement dans le cas de l'ondelette de Cauchy, en plus de la concentration à l'origine  $x = y = 0$ , une concentration du noyau autour de la droite  $y = x$ . Rappelons que la fonction de corrélation usuelle est définie par le noyau  $\delta(x - y)$ , c'est-à-dire par un noyau concentré sur la

droite  $y = x$ . La fonction de corrélation développée ici est effectivement une définition locale de la corrélation au sens usuel : le noyau obtenu est une variante concentrée autour de 0 du noyau initial.

Le fait que le noyau se répande autour de la droite  $y = x$  montre les distorsions autorisées dans la comparaison des régions. En effet, les valeurs du noyau le long d'une droite  $y = \alpha x$  sont les pondérations dont vont être affectées les valeurs  $f(t - b)\bar{g}(\alpha(t - b'))$  dans le calcul du produit de corrélation localisée, correspondant à une comparaison du signal  $f$  autour de  $b$  et du signal  $g$  dilaté d'un facteur  $\alpha$  autour de  $b'$ . Dans le cas de la corrélation usuelle, le noyau est nul en dehors de  $y = x$ , c'est-à-dire qu'une contrainte implicite de rigidité est prise en compte<sup>2</sup>.

Dans le cas de l'ondelette de Cauchy, le noyau  $\mathcal{K}$  est nul pour les valeurs  $\alpha < 0$ , ce qui peut s'interpréter comme un respect de l'ordre chronologique (dans le cas unidimensionnel) ou de l'organisation de l'espace (dans le cas des images).

Pour le laplacien de gaussienne, le noyau est indifférent au signe de  $x$  et  $y$ , ce qui se traduit par une tolérance à la symétrie.

## 4. pré-traitements : normalisation et contraste

### 4.1. utilité et choix d'une normalisation

La fonction de corrélation est très sensible à la valeur moyenne des signaux. En effet, telle qu'elle est définie, elle apparaît comme une fonction bilinéaire des deux signaux  $f$  et  $g$ , c'est-à-dire que la taille du pic de corrélation est une grandeur multiplicative de leur puissance. Ainsi, il se peut que, entre les structures locales de  $g$  en  $t_1$  et  $t_2$ , celle en  $t_1$  soit plus proche de la structure locale de  $f$  en  $t_0$ , mais que l'intensité de  $g$  en  $t_2$  soit suffisamment élevée pour compenser la faiblesse de similarité des structures. Dans le cas de la corrélation usuelle, plusieurs solutions ont été proposées dont par exemple la corrélation de phase [Fle94] ou des normalisations locales. Nous proposons d'utiliser un pré-traitement en contraste présentant des similarités avec le système visuel humain, le contraste multiéchelle [DD91,DDMT93].

Par ailleurs, il est utile de conserver l'interprétation de la fonction de corrélation localisée comme résultat d'un problème de minimisation. Or, minimiser

$$\int |\mathcal{O}[f](a, b + \delta) - \mathcal{O}[g](a, b)|^2 \frac{da}{a} .$$

2. Deverny et Faugeras proposent de la contourner en utilisant une fonction de corrélation «enrichie», dépendant des paramètres de localisation et de distorsion [DF94].

par rapport à  $\delta$  n'est pas équivalent à maximiser  $s(b, b')$ , car  $\int |\mathcal{O}[f](a, b)|^2 da/a$  dépend de  $b$  en général. Ceci suggère de normaliser au préalable les signaux en imposant

$$\forall b, \int |\mathcal{O}_n[f](a, b)|^2 \frac{da}{a} = 1 ,$$

où  $\mathcal{O}_n(b, a)$  sont les coefficients normalisés. La corrélation localisée  $s(b, b')$  s'interprète alors comme le cosinus de l'angle entre les vecteurs  $|\mathcal{O}[f](a, b)$  et  $|\mathcal{O}[f](a, b')$ . Ses valeurs sont toujours comprises dans l'intervalle  $[-1, +1]$  et peuvent être interprétées comme des coefficients de confiance sur l'appariement de  $b$  et  $b'$ . Puisque nos vecteurs sont maintenant tous de même norme, la maximisation du produit scalaire est équivalente à la minimisation de la différence des vecteurs en norme quadratique,

$$\int |\mathcal{O}_n[f](a, b) - \mathcal{O}_n[g](a, b')|^2 \frac{da}{a} . \quad (19)$$

#### Remarque

Une alternative consisterait à minimiser directement la quantité  $\int |\mathcal{O}[f](a, b) - \mathcal{O}[g](a, b')|^2 \frac{da}{a}$ , (voir par exemple [JonMal91], qui préfèrent toutefois la norme  $\mathcal{L}^1(\mathbb{R})$  à la norme  $\mathcal{L}^2(\mathbb{R})$ ). Ceci présente l'avantage d'éviter l'étape de normalisation, mais ne conduit plus à une fonction de corrélation localisée simple et contrôlable.

### 4.2. contraste multirésolution

Soit une fonction  $f \in \mathcal{L}^1(\mathbb{R}) \cap \mathcal{L}^2(\mathbb{R})$  positive. Soit  $\mathcal{P}[f](a, b)$  l'approximation de  $f$  par la famille de fonctions  $\psi_{\alpha, b}$ ,  $\alpha \geq a$ , définie par (voir l'équation (9)) :

$$\mathcal{P}[f](a, b) = \int_a^{+\infty} \mathcal{O}[f](\alpha, b) \frac{d\alpha}{\alpha^{3/2}} . \quad (20)$$

$\mathcal{P}[f](a, b)$  peut être interprété comme le filtrage de  $f$  par la fonction de lissage  $\rho_{a, b}$  définie par

$$\rho_{a, b}(t) = \int_a^{+\infty} \psi_{\alpha, b}(t) \frac{d\alpha}{\alpha^2} , \quad (21)$$

et on a  $\lim_{a \rightarrow 0} \mathcal{P}[f](a, t) = f(t)$ . Alors le contraste du signal  $f$  à l'échelle  $a$  en  $b$  est défini comme le rapport du détail  $\mathcal{O}[f](a, b)$  et de l'approximation  $\mathcal{P}[f](a, b)$  :

$$\mathcal{C}[f](a, b) = \frac{\mathcal{O}[f](a, b)}{\mathcal{P}[f](a, b)} = -a^{3/2} \frac{\partial \log \mathcal{P}[f](a, b)}{\partial a} . \quad (22)$$

Duval-Destin *et al.* [DDMT93] montrent alors que cette décomposition multirésolution non-linéaire possède une formule de reconstruction :

$$f(t) = \mathcal{P}[f](a, t) \left( \int_0^a \mathcal{C}[f](\alpha, t) \frac{d\alpha}{\alpha^{3/2}} \right) . \quad (23)$$

**Remarque**

Cette équation est particulièrement intéressante si on la compare à (20), dans la mesure où elle montre que les coefficients de contraste  $\mathcal{C}[f](a, b)$  peuvent être interprétés comme des coefficients d'ondelettes du logarithme de l'image (un énoncé précis peut être trouvé dans [DDMT93]). Ceci, combiné à la remarque faite en fin de section 2.2, montre que les coefficients de contraste ne changent pas si on multiplie le signal par une constante. Pour le problème qui nous concerne, ceci a pour effet d'améliorer la robustesse aux changements d'illumination.

Le contraste multirésolution extrait les transitions spatiales de l'intensité en fonction de leur importance morphologique, parce qu'il affranchit (partiellement) l'image de la luminance ambiante. D'un point de vue physiologique, cette définition du contraste correspond à une généralisation de la fraction de Weber et permet de prédire les seuils de contraste du système visuel pour des stimulations complexes [DD91].

**4.3. utilité du contraste et de la normalisation quadratique**

L'introduction du contraste de préférence à la luminance permet d'homogénéiser entre eux les coefficients d'ondelettes correspondant aux différentes bandes de fréquences. Par ailleurs, alors que les coefficients de luminance sont liés à une description des régions de l'image, l'information privilégiée en utilisant les coefficients de contraste est l'information de contour, *a priori* plus robuste, dans le sens où elle est moins sensible aux différences de gain de luminance des caméras (cf. remarque en section 4.2). Enfin, les coefficients de contraste présentent l'avantage sur les coefficients de luminance d'être bornés [DD91] (ce qui ne garantit pas toutefois que les valeurs de la fonction de corrélation le soient).

La normalisation quadratique résout le problème de la finitude de la fonction de corrélation localisée. En pratique, ce problème est aussi résolu par la compacité du support des signaux étudiés : l'intégrale sur les échelles est à bornes finies.

L'algorithme d'appariement exposé précédemment utilise en entrée les coefficients de la décomposition en ondelettes des images gauche et droite. Les coefficients obtenus après les prétraitements suggérés (contraste et normalisation quadratique) peuvent être interprétés comme les coefficients d'ondelettes d'images ayant subi un filtrage non-linéaire avant leur mise en correspondance : ce filtrage a pour effet principal de réhausser les contrastes dans l'image initiale.

La figure 4 montre l'effet de ce filtrage appliqué à l'image gauche de la paire stéréoscopique représentée sur la figure 1. Les structures de l'image sont les mêmes, grâce à la localisation de la décomposition en ondelettes, et désormais, le contraste est équilibré en chaque point de l'image. En particulier, les contours dus aux ombres sont désormais aussi marqués que ceux provenant des arêtes des objets présents dans la scène et la texture de la colline dans l'ombre des poteaux devient perceptible.

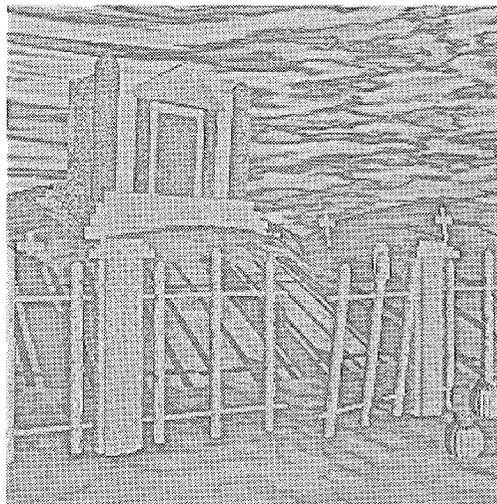


Figure 4. – L'utilisation du contraste multiéchelle et de la norme quadratique des coefficients interprétée comme un filtrage non-linéaire réhaussant le contraste des images.

**4.4. taux de confiance des appariements**

Grâce à la normalisation quadratique, il est possible d'interpréter la valeur des produits de corrélation localisée comme un taux de confiance dans l'appariement. Jusqu'à présent, nous avons utilisé la corrélation localisée comme une mesure de similarité des signaux  $f$  autour de  $b$  et  $g$  autour de  $b'$ . Or, après normalisation, tous ces produits sont compris dans l'intervalle  $[-1, 1]$ . En étudiant l'histogramme des valeurs prises par le produit de corrélation, il est possible de déterminer un seuil de confiance des appariements  $\sigma_c$ , en-deçà duquel l'appariement, jugé non fiable, est rejeté.

La figure 5 donne un exemple d'histogramme obtenu à partir des valeurs de la fonction de corrélation localisée sur toutes les paires

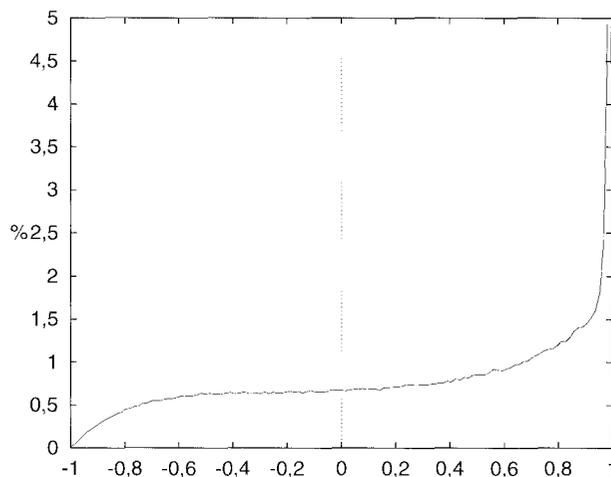


Figure 5. – Histogramme des valeurs du produit de corrélation localisée obtenu sur les lignes d'une image-test : un seuil de confiance dans les appariements peut être placé à  $\sigma_c = 0,9$ .

de pixels de la figure 1. D'autres histogrammes ont été calculés avec des images différentes. Tous présentent la particularité d'un pic très prononcé en 1, précédé d'un long plateau. La valeur  $\sigma_c$  retenue peut être la transition entre le plateau et le pic, soit dans notre cas  $\sigma_c \simeq 0,9$ . Dans ce cas précis, 15% des appariements testés sont acceptés, contre 85% rejetés.

## 5. applications potentielles dans le cadre de la réalité virtuelle

### 5.1. réalité virtuelle et perception stéréoscopique

Replaçons maintenant la fonction de corrélation développée précédemment dans le cadre que nous nous sommes fixés au départ, à savoir le contrôle de la profondeur de champ avant la restitution d'images stéréoscopiques à l'utilisateur. Usuellement, deux approches sont proposées :

- une approche fondée sur la disparité, où les images sont traitées de manière à limiter la gamme de disparités présentes dans la paire stéréoscopique [par exemple];

- une approche fréquentielle, visant à limiter le contenu en fréquences spatiales par introduction de flou dans les images [BBS<sup>+</sup>97].

Une approche mixte est celle proposée par Perrin [Per98], limitant le contenu en fréquences spatiales en fonction de la disparité présente localement dans l'image. Cette approche est fondée sur une décomposition multiéchelle de l'image, sur la connaissance de la disparité en tout point et sur une étude ergonomique [Wop95]. Cette dernière permet de définir une *fonction de confort stéréoscopique* [PFRP98], donnant la plus haute fréquence spatiale admissible à disparité donnée. Elle nécessite donc de connaître *localement* le contenu en fréquences spatiales et la disparité.

Une méthode fondée sur la fonction de corrélation localisée semble donc tout à fait appropriée à ce besoin, et ce d'autant plus si nous considérons le problème du temps réel, c'est-à-dire de l'optimisation du temps de calcul. En effet, la décomposition multiéchelle nécessaire au traitement ergonomique (et au demeurant, parallélisable) est déjà effectuée pour la corrélation localisée : dans ce sens, aucun surcoût de calcul n'est introduit. De plus, la fonction de corrélation localisée est, par définition, locale et renseigne donc sur la disparité en un point. La méthode de détermination de la profondeur dans une paire stéréoscopique par la fonction de corrélation localisée se présente donc comme

une étape relativement naturelle pour un traitement ergonomique fondé sur un critère fréquence-disparité.

Notons enfin que l'algorithme d'appariement par corrélation localisée est massivement parallélisable. En effet, après avoir effectué la décomposition multiéchelle, l'appariement est réalisé de manière indépendante point par point : nous pouvons donc espérer que le temps de calcul nécessaire à la mise en correspondance soit compatible avec des applications « temps réel », utilisant une perception de l'environnement par caméras stéréoscopiques pour la navigation autonome de robots.

### 5.2. la corrélation localisée pour l'appariement stéréoscopique

La décomposition en ondelettes continue exposée aux paragraphes précédents peut être discrétisée de plusieurs manières. En choisissant de conserver au maximum l'invariance par translation par une transformation en ondelettes discrétisée sur une grille de la forme  $(a_0^j, kb_0)$ ,  $j, k \in \mathbb{Z}$  c'est-à-dire uniforme, la définition de la fonction de corrélation localisée donnée à l'équation (11) devient

$$\begin{aligned} s(b, b') &= (f \star^{\text{loc}} g)(b, b') \\ &= \sum_j \mathcal{O}[f](2^{j/2}, b) \overline{\mathcal{O}[g](2^{j/2}, b')}. \end{aligned} \quad (24)$$

En considérant la décomposition en ondelettes des signaux à apparier comme une représentation vectorielle de ces signaux, l'appariement est réalisé entre les vecteurs les plus proches au sens de la colinéarité : l'expression (24) fait intervenir le produit scalaire de ces vecteurs, qui sera d'autant plus important que les vecteurs seront semblables. L'effet de l'amplitude des vecteurs est quant à lui gommé par la normalisation quadratique. Notons que dans ce cas, le noyau de corrélation est une fonction de quatre variables (ou deux variables vectorielles).

Plusieurs formes d'ondelettes-mères peuvent être envisagées, menant à différentes options pour la maximisation du produit scalaire [Per98] :

- les dérivées de Gaussienne : utilisées régulièrement en vision, le passage de signaux unidimensionnels à des images se fait sans problème en utilisant des gaussiennes radiales, et les résultats explicités précédemment s'étendent sans peine;

- les fonctions de Cauchy : définies pour des signaux unidimensionnels, le noyau associé présente de bonne qualité de sélectivité, mais l'extension aux images est plus compliquée; citons Antoine *et al.* [AMV97] qui proposent une classe d'ondelettes bidimensionnelles directionnelles à partir des fonctions de Cauchy;

- les fonctions de Gabor,  $\psi(x, y) = e^{-jkx} e^{-(x^2+y^2)/2}$  : à sélectivité directionnelle également, elles permettent de mettre l'accent sur les structures verticales de l'image, structures d'autant plus importantes que nous estimons des décalages essentiellement horizontaux.

## Une fonction de corrélation localisée

Un algorithme simple d'appariement par la fonction de corrélation localisée peut maintenant se résumer par

1. acquisition des signaux  $\Rightarrow f, g$ ;
2. décomposition en contraste multiéchelle  
 $\Rightarrow \mathcal{C}[f](a, b), \mathcal{C}[g](a, b)$ ;
3. adaptation des échelles  
 $\Rightarrow \tilde{\mathcal{C}}[f](a, b), \tilde{\mathcal{C}}[g](a, b)$ ,  
où  $\tilde{\mathcal{C}}$  est défini de manière analogue à  $\tilde{\mathcal{O}}$ ;
4. normalisation quadratique  
 $\Rightarrow \mathcal{N}[f](a, b) = \frac{\tilde{\mathcal{C}}[f](a, b)}{\|\tilde{\mathcal{C}}[f](b)\|}, \mathcal{N}[g](a, b') = \frac{\tilde{\mathcal{C}}[g](a, b')}{\|\tilde{\mathcal{C}}[g](b')\|}$ ;
5. corrélation localisée  
 $\Rightarrow s(b, b') = \int \mathcal{N}[f](a, b) \cdot \overline{\mathcal{N}[g](a, b')} \frac{da}{a}$ ;
6. calcul de la disparité en  $b$  si  $s(b, b') \geq \sigma_c$  : pour un certain  $b'$   
 $d(b) = \arg \max_{b'} s(b, b') - b$ ;  
sinon :  
aucune estimation n'est possible.

La dernière étape ci-dessus établit que le point correspondant à  $b$  est celui permettant de maximiser la fonction de corrélation localisée. Pour simplifier nous avons décrit ici un algorithme de recherche simple. Il va de soi que cette étape peut être optimisée, par exemple en imposant des contraintes sur la vitesse des variations de la disparité.

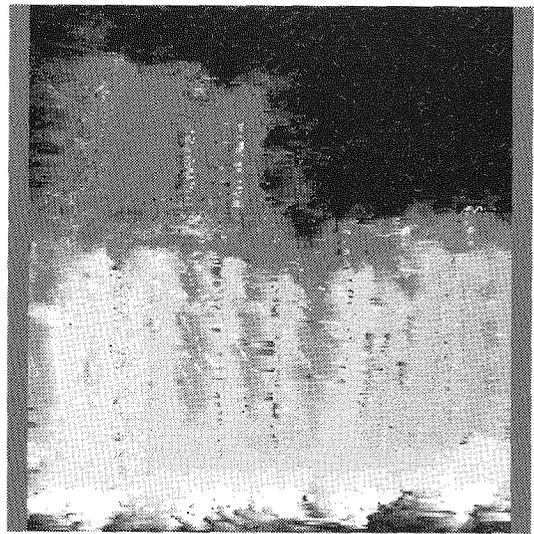
### 5.3. un exemple d'appariement stéréoscopique

L'algorithme a été testé sur les images présentées sur la figure 1, de taille  $256 \times 256$  pixels. L'image droite est obtenue par décalage horizontal de la position de la caméra par rapport à l'image gauche : les deux axes optiques étant parallèles, les disparités entre les points des images gauche et droite sont purement horizontales, information prise en compte dans l'appariement qui se ramène alors à un problème à une dimension. L'appariement est effectué dans un seul sens : nous cherchons dans l'image droite un correspondant à chaque point de l'image gauche. Une mise en correspondance complète nécessiterait d'effectuer l'appariement dans l'autre sens et de fusionner les informations. Les décompositions multirésolution ont été effectuées en choisissant une ondelette-mère de type DOGs. Après l'appariement, un filtre médian ( $3 \times 5$  pixels) est appliqué pour limiter le bruit dans la carte des disparités.

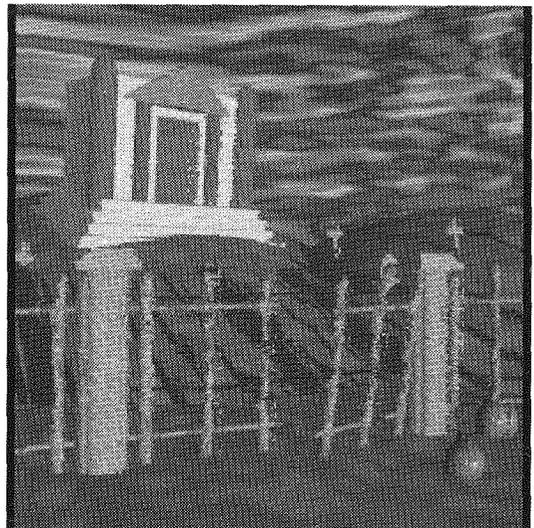
Les résultats sont estimés selon deux critères visuels :

- la carte de disparité des points (figure 6(a)), représentative de leur profondeur (dans le cas d'axes optiques parallèles, la profondeur est inversement proportionnelle à la disparité);
- la reconstruction de l'image droite à partir de la luminance des points de l'image gauche et de leur disparité (figure 6(b)).

La reconstruction donne de bons résultats : en dehors de certains défauts dus aux occlusions (les bords du bâtiment et des poteaux...), l'image droite et sa reconstruction sont quasiment identiques. La carte des parallaxes est relativement bruitée mais nous y retrouvons 1) le plan du sol (dégradé de gris à l'avant) sur lequel se découpent en plus clair les poteaux, 2) le bâtiment principal et les collines. En la comparant à la carte des parallaxes exactes (que nous connaissons sur ces images de synthèse), nous obtenons que l'erreur moyenne sur l'ensemble de l'image est de 1,00 pixel (soit de l'ordre de la résolution de l'image) et que 78% des points ont une erreur inférieure à un pixel. Notons en particulier que, grâce à la localisation de la fonction de corrélation, la colline visible entre les poteaux (et plus éloignée) est effective-



(a) Carte des parallaxes estimées, d'autant plus claires qu'elles sont importantes (c'est-à-dire que les objets sont proches des caméras)



(b) Reconstruction de l'image droite à partir de l'image gauche et de la carte des disparités

Figure 6. – Résultats de l'algorithme d'appariement par fonction de corrélation localisée.

ment estimée à une profondeur supérieure à celle des poteaux. Le temps de calcul nécessaire à l'appariement de ces images, en recherchant le maximum de la fonction de corrélation localisée sur une fenêtre de taille  $1 \times 33$  pixels, sur une station de travail Silicon Graphics O2, est de l'ordre de 2 minutes (dont plus de la moitié consacrée aux décompositions en contraste). A titre de comparaison, nous avons implémenté un algorithme hiérarchique à corrélation linéaire, utilisant quatre échelles différentes (64, 32, 16 et 8 pixels). Les résultats et le temps de calcul sont comparables (74% d'appariements corrects pour 1min40 de calcul).

Nous avons également testé l'algorithme sur des stéréogrammes de points aléatoires. Ces stéréogrammes présentent la particularité de ne contenir aucune information de profondeur monoculaire : l'information de profondeur est uniquement contenue dans la disparité des points du stéréogramme et ne peut donc être obtenue que par fusion binoculaire [Jul71]. Avec le stéréogramme de points aléatoires classique (un carré de taille moitié de l'image en hauteur et en largeur se découpe dans l'image), l'erreur moyenne obtenue est cette fois de 0,34 pixel, avec 94% des disparités estimées à moins d'un pixel d'erreur.

## 6. conclusions

Nous avons proposé la définition d'une fonction de corrélation localisée de deux signaux en utilisant leur décomposition en ondelettes. Cette corrélation est construite, non pas en intégrant sur la variable d'espace comme dans la corrélation usuelle, mais sur la variable d'échelle.

Nous avons alors montré que cette opération est équivalente à définir un noyau de corrélation présentant des propriétés importantes de localisation dans l'espace produit des deux espaces de départ des signaux. Plus particulièrement, le noyau obtenu mesure les affinités de deux signaux dans des fenêtres de temps (ou d'espace) bien localisées. La largeur de ces fenêtres peut de plus être adaptée en déterminant un paramètre physique local d'échelle qui fixe la borne d'intégration inférieure de notre fonction de corrélation.

Enfin, nous avons proposé des pré-traitements des signaux à mettre en correspondance pour améliorer les performances de notre algorithme. Ces pré-traitements utilisent d'une part, une décomposition en contraste multi-résolutions des signaux; d'autre part une normalisation quadratique de la décomposition multi-échelle qui permet de mesurer le taux de confiance dans les appariements et équivaut à un ajustement du contraste avant l'appariement.

Nous pouvions craindre qu'un algorithme d'appariement stéréoscopique dense n'utilisant ni contrainte globale [OLL96], ni information de haut niveau soit peu robuste dans les appariements. Les résultats obtenus sur des images de synthèse montrent qu'il résout plutôt bien le problème de l'appariement, mais qu'il doit être associé à des informations complémentaires pour limiter le bruit.

Mais l'avantage principal de cet algorithme est de pouvoir s'insérer avec un coût de calcul minimal dans des chaînes de traitement de l'information visuelle fondées sur une approche multi-échelle, que ce soit pour prendre en compte le fonctionnement du système visuel humain ou pour une compression des images à base d'ondelettes. L'exemple de l'ergonomie stéréoscopique a été donné dans le cadre de cet article. Une autre application potentielle de cet algorithme pourrait être de déterminer la structure de l'environnement à partir du *flux optique*, c'est-à-dire le champs des vitesses des points de l'image [Per98].

## BIBLIOGRAPHIE

- [AMV97] J. P. Antoine, R. Murenszi, V. Vandergheynst. «The Cauchy wavelets : a new class of 2D, directionnal wavelets». Technical Report UCL-ILPT-97-01, Louvain la Neuve, Belgique, 1997.
- [BBS+97] W. Blohm, I.P. Beldie, K. Schenke, K. Fazel, S. Pastoor. « Stereoscopic image representation with synthetic depth of field ». *Journal of the Society for Information Display*, 1997.
- [CR68] F. W. C. Campbell, J. Robson. « Application of Fourier analysis to the visibility of gratings ». *Journal of physiology (London)*, 197 :551-556, 1968.
- [CSS61] L.K. Cormack, S. B. Stevenson, C.M. Schor. «Interocular correlation, luminance contrast and cyclopean processing». *Vision Research*, 31 (12) : 2195-2207, 1991.
- [DD91] M. Duval-Destin. *Analyse spatiale et spatiotemporelle de la stimulation visuelle à l'aide de la transformée en ondelettes*. Doctorat de physique théorique, Faculté des Sciences de Luminy, 1991.
- [DDMT93] M. Duval-Destin, M.A. Muschietti, B. Torrèsani. «Continuous wavelet decompositions, multiresolution and contrast analysis». *SIAM Journal of Mathematical Analysis*, 24(3) : 739-755, 1993.
- [DF94] F. Devernay, O. Faugeras. «Computing differential properties of 3D shapes from stereoscopic images without 3D models». Technical Report 2304, INRIA (Robotique, image et vision), 1994.
- [ECR66] C. Enroth-cugel, J. D. Robson. «The contrast sensitivity of retinal ganglion cells of the cat». *Journal of Physiology*, 187 :517-522, 1966.
- [Fle94] D.J. Fleet. «Disparity from local weighted phase-correlation ». In *IEEE International Conference on Systems, Man and Cybernetics*, volume 1, pages 48-54, 1994.
- [Gri81] W.E.L. Grimson. In *From images to surface : a computational study of the human early vision system*. M.I.T. Press, Cambridge, 1981.
- [Hel88] U.V. Helava. «Object-space least- squares correlation». *PE&RS*, 54(6) : 711-714, 1988.
- [JM91] D.G. Jones, J. Malik. «A computational framework for determining stereo correspondence from a set of linear spatial filters». Technical Report UCB/CSD 91/655, Computer Science Division, Berkeley, University of California, 1991.
- [Jul71] B. Julesz. In *Foundations of the cyclopean perception*. University of Chicago Press, Chicago, 1971.
- [LNN94] K. O. Ludwing, H. Neumann, B. Neumann. «Local stereoscopic depth estimation». *Image and Vision Computing*, 12(1) :16-35, 1994.
- [Mal89] S.G. Mallat. «A Theory for multiresolution signal decomposition : the wavelet representation». *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7) :674-693, 1989.
- [Mar82] D. MARR. In *Vision : a computational investigation into the human representation and processing of visual information*. Freeman, New York, 1982.

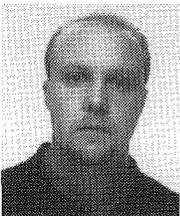
## Une fonction de corrélation localisée

- [MGA96] H.A. Mallot, S. Gillner, P.A. Arndt. «Is correspondence search in human stereo vision a coarse-to-fine process». *Biological Cybernetics*, 74:95-106, 1996.
- [MSJ96] J.S. MCVEIGH, M.W. SIEGEL, A.G. JORDAN. «Algorithm for automated eye strain reduction in real stereoscopic images and sequences». In *Proceedings of the SPIE*, volume 2657, pages 307-316, 1996.
- [Nis87] H.K. Nishihara. In M.A. Fichler, O. Firschein, éditeurs, *Readings in Computer Vision*. Kauffman, Los Altos, 1987.
- [OLL96] M.H. OUALI, H. LANGE, C. LAURGEAU. «An energy minimization approach to dense stereovision». In *Proceedings of the IEEE International Conference on Image Processing*, volume 2, pages 841-846, Lausanne (Suisse), septembre 1996.
- [Per98] J. Perrin. *Profondeur et binocularité : algorithmie, étude psychophysique et application à l'ergonomie des interfaces stéréoscopiques*. Doctorat de Informatique temps réel, Robotique et automatique, Ecole des Mines de Paris, 1998.
- [PFRP98] J. Perrin, P. Fuchs, C. Roumes, F. Perret. «Improvement of stereoscopic comfort through control of the disparity and of the spatial frequency content». IN *Processing of the SPIE, Aerosense '98*, Orlanda (FL), 1998.
- [Rou95] C. Roumes. *Contenu fréquentiel de l'image et vision binoculaire : étude psychophysique chez l'homme*. Doctorat de neurosciences, Université Paris 6, 1995.
- [Sma95] H.S. Smallman. «Fine-to-coarse scale disambiguation in stereopsis». *Vision Research*, 35(8) :1047-1060, 1995.
- [Wöp95] M. Wöpking. «Viewing comfort with stereoscopic pictures : an experimental study on the subjective effects of disparity magnitude and depth of focus». *Journal of the SID*, 3 :101-103, 1995. »

Manuscrit reçu le 8 juin 1998.

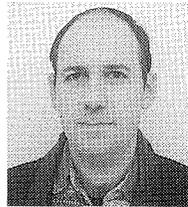
### LES AUTEURS

#### J. PERRIN



Diplômé de l'Ecole Polytechnique, après un doctorat en Robotique à l'Ecole des Mines de Paris en collaboration avec l'Institut de Médecine Aérospatiale du Service de Santé des Armées, Jérôme Perrin est actuellement chef du département « Evaluation de l'électronique à bord des véhicules terrestres » à l'Etablissement Technique d'Angers (DGA/ETAS).

#### B. TORRÉSANI



Bruno Torrèsani est professeur à l'Université de Provence, Marseille, et membre du *Laboratoire d'Analyse, Topologie et Probabilités* de Marseille. Ses domaines d'intérêt sont majoritairement liés aux méthodes mathématiques du traitement du signal. Il est l'auteur des livres « Introduction à l'Analyse Continue par Ondelettes » (InterEditions/Editions du CNRS) et de « Practical Time-Frequency Analysis » (Academic Press, avec R. Carmona et W.L. Hwang).

#### P. FUCHS



Philippe Fuchs, Professeur agrégé au Centre de Robotique de l'Ecole des Mines de Paris, est le responsable de l'équipe « Réalité Virtuelle et Réalité Augmentée », orientée vers les applications industrielles. Il est l'auteur du livre « Les interfaces de la Réalité Virtuelle ».