

Pyramide irrégulière : une représentation pour la vision exploratoire

Irregular Pyramid: A Hierarchical Structure for Exploratory Vision

par Antoine MANZANERA, Jean-Michel JOLION

*Reconnaissance de Formes et Vision –
Institut National des Sciences Appliquées – Bât. 403
F-69221 Villeurbanne cedex*

Résumé

Nous introduisons dans cet article une nouvelle classe de structures hiérarchiques irrégulières où la nature du sous-échantillonnage est une fonction de la position spatiale. Cette structure s'apparente à la structure rétinienne (notion de fovéa et de périphérie). Ce nouveau mécanisme permet de conserver un échantillonnage précis dans la zone de focus tout en résumant le restant de la scène. L'image résultat étant de taille réduite, les traitements qui lui sont appliqués sont plus rapides. Nous présentons des exemples et une validation de cette approche sur des images dynamiques dans le contexte de la vision exploratoire.

Mots clés : Pyramide, multirésolution, focus d'attention, mouvement.

Abstract

We present a new class of hierarchical structures known as the irregular pyramid. This pyramid is characterized by a non regular subsampling defined as a function of the spatial location of the points. This hierarchical structure is made as consistent as possible with the human visual mapping. This novel structure allows both fine sampling in the focus area and coarse sampling elsewhere in the scene so resulting in smaller images. Examples are shown as well as an exploratory vision application in motion detection.

Key words : Pyramid, multiresolution, motion.

1. Introduction

La vision artificielle nécessite, tout particulièrement en environnement dynamique, l'extraction rapide d'informations dans une (des) image(s). Par souci de précision (de mesure ou de localisation), l'échantillonnage doit être le plus fin possible et donc l'image de grande taille. Les capteurs classiques actuels, *i.e.* caméra CCD, fournissent des images de 512×512 pixels, ce qui est à la fois beaucoup en terme de volume de données à traiter, et très peu pour une faible incertitude de mesure. De plus, le système incluant le module de vision a rarement besoin d'une telle précision sur toute la scène observée et, plus simplement, il n'a généralement pas le temps de traiter un tel volume de données. Il apparaît donc utile, peut-être indispensable, de considérer l'image acquise à des résolutions de plus en plus réduites, de façon à effectuer le traitement voulu de manière grossière (et rapide), pour l'affiner, si besoin est, en ré-augmentant la résolution (principe de la vision exploratoire).

Une image réduite contient généralement une information résumée uniformément sur toute la scène. Ce n'est pas ce qui est requis lorsque l'on procède à une exploration d'une scène. En effet, dans cette approche, on définit un mode d'observation en zones. La zone fovéale correspond au focus d'attention donc à la partie de la scène où l'on désire conserver le plus d'information (échantillonnage le plus fin possible). La zone périphérique peut, elle, être résumée. Elle correspond à une partie de la scène beaucoup plus vaste mais qui n'a que peu d'importance à cet instant et peut donc être fortement résumée (sans pour autant être totalement éliminée comme dans les techniques dites de fenêtrage). Dans le cas de la vision humaine, cette zone est très importante pour la détection d'objets en mouvement entrant dans le champ visuel. L'objectif n'est pas de permettre une mesure ou localisation précise, mais plutôt de permettre la génération de signaux de type alarme, influant sur le mécanisme de contrôle de la position de la zone fovéale. Cette approche est donc différente de celles proposées dans [1,2,3,4]. En effet, dans l'algorithme proposé par Yeshurun et Schwartz, un objet est analysé par accumulations d'évidences dans un système à multiples vues à des

résolutions différentes. Il n'y a donc pas utilisation simultanée des deux types de vision, fovéale et périphérique. Les trois autres modèles réalisent le concept de vision fovéale par le biais d'une gestion de fenêtres. A un instant donné, seule une partie (petite en général) de la scène est prise en compte. Ce fonctionnement en aveugle présente un gros risque de perte d'informations ou de non-détection d'évènements importants dans le cas des scènes dynamiques.

Il serait donc souhaitable de posséder un capteur à échantillonnage irrégulier. Cependant, ce type de capteur est encore trop peu disponible malgré quelques réalisations telle celle proposée par Bolduc et Levine [5]. Nous proposons une autre approche consistant à transformer la sortie d'un capteur classique grâce à l'emploi d'une structure pyramidale irrégulière. Cet outil très simple dans sa construction et son utilisation est intégré dans un environnement de programmation dédié aux structures pyramidales. Il permettra l'étude de l'influence de la structure image sur les traitements classiques (détection de contours, suivi d'objets, mesure, reconstruction tri-dimensionnelle, ...) utilisés en vision par ordinateur.

2. Le modèle pyramidal

La pyramide est une structure hiérarchique qui supporte le mécanisme de multirésolution dont une description détaillée peut être trouvée dans [6,7]. Sous sa forme la plus canonique, elle est constituée d'une suite d'images dont la taille décroît géométriquement d'un facteur 2 dans chaque dimension. Le niveau 0 est l'image initiale. Si la taille de celle-ci est de la forme 4^N alors l'ultime niveau, ou apex, est constitué d'un unique pixel.

Il existe de nombreuses manières de construire une représentation multirésolution d'une image donnée. Dans le mécanisme le plus simple, chaque cellule, ou pixel, d'un niveau p , $p > 0$, (cellule père) provient d'une combinaison linéaire de quatre cellules (fils) du niveau $p - 1$ (cf Figure 1).

On définit aussi des pyramides avec d'autres types de liens : 9 fils-1 à 4 père(s) avec des niveaux de taille impaire; 16 fils-4 pères où chaque cellule du niveau père est une combinaison linéaire de 16 cellules du niveau fils (il en résulte un recouvrement qui fait que chaque cellule contribue à 4 cellules du niveau supérieur) [6].

Toutes ces structures sont néanmoins régulières; les liens sont uniformes dans toute l'image et le masque de pondération est unique (sauf peut-être sur les bords ce qui induit des comportements de moins en moins conformes à la théorie au fur et à mesure que l'on réduit la taille de l'image).

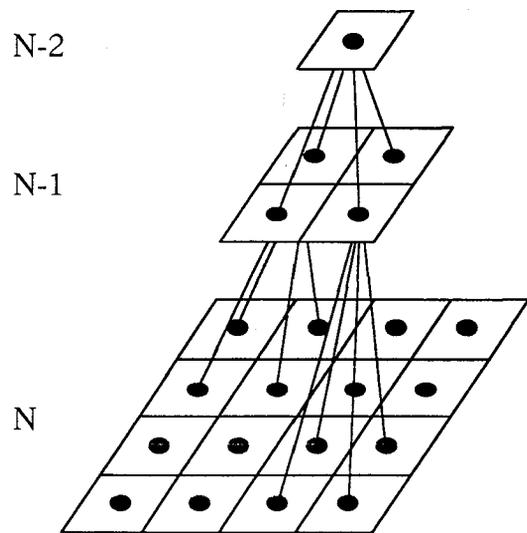


Figure 1. - La pyramide quaternaire 1 père → 4 fils.

3. Pyramide irrégulière

Les pyramides dites irrégulières sont le plus souvent liées à des structures de graphes telles que les pyramides stochastiques et adaptatives introduites dans [8,9,10,11]. Le sous-échantillonnage est une fonction du contenu du signal, et ne peut donc être prédéterminé. On peut aussi regrouper dans cette classe les pyramides utilisant des pavages particuliers telle que la pyramide proposée par Kropatsch pour l'analyse de courbes [12]. Nous introduisons dans cet article une autre catégorie de pyramides irrégulières où le type de lien est une fonction de la position spatiale, les tailles successives des différents niveaux de la pyramide restant similaires à celles d'une pyramide classique (*i.e.* pyramide quaternaire). En nous inspirant de la distribution des neurones de transmission directe de la rétine, nous définissons un système de liens non uniforme entre un niveau p et un niveau $p+k$, en plaçant un nombre d'autant plus important de liens que l'on est éloigné du centre de l'image.

Dans un premier temps, considérons un signal discret I_N mono-dimensionnel connu en 2^N valeurs. On souhaite établir une relation entre ce signal et le signal I_{N-k} connu en 2^{N-k} valeurs. Pour cela, il faut définir, pour tout x de I_{N-k} , le nombre et le détail de positions des y de I_N intervenant dans le calcul de $I_{N-k}(x)$, *i.e.* les fils de x . Nous noterons $\#fils(x)$ le nombre d'éléments de I_N en relation avec x . Nous avons vu que dans une structure pyramidale régulière, $\#fils(x)$ est une constante quel que soit x . Dans le cas d'une structure hiérarchique irrégulière, $\#fils(x)$ est déterminé grâce à d'autres contraintes :

Contrainte C_1 : définition de $\#fils(x)$

- $\#fils(x)$ est entier et borné : $\#fils(x) \in [1, F]$

- $\#fils(x)$ est une fonction croissante de $|x - fovéa|$ où *fovéa* désigne le centre du domaine spatial recouvert par le signal I_{N-k} .

Contrainte C₂ : Comportement de $\#fils(x)$

- $\#\{x : \#fils(x) = i\} = 2L \geq 2$ pour tout $i \in [1, F]$ (on ne privilégie aucun nombre de fils, et toutes les configurations, i.e. 1, ..., F, existent)

La parité se justifie par un souci de symétrie du comportement de part et d'autre de la fovéa.

On peut très simplement montrer que ces premières contraintes induisent (cf annexe 1) :

$$FL = 2^{N-k-1}, \quad F = 2^f \quad \text{et} \quad L = 2^h$$

La Figure 2 présente un exemple pour deux niveaux consécutifs ($k = 1$). Deux systèmes de liens réguliers ((1 père, 2 fils) et (2 pères, 4 fils)) y sont également présentés pour comparaison.

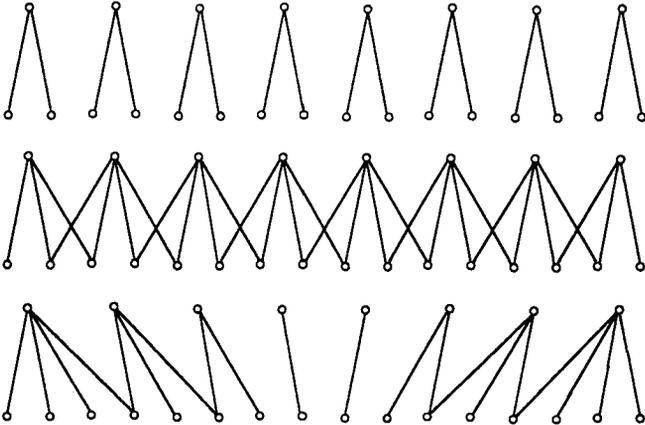


Figure 2. – Les différents types de liens en dimension 1.

Le nombre total de liens entre I_N et I_{N-k} est égal à : $C(k) = (F + 1)FL$. Pour aller plus avant dans la détermination de la répartition de ces liens, nous utilisons une troisième contrainte, classique dans les structures pyramidales et caractérisant la non perte d'information entre deux niveaux :

Contrainte C₃ :

Tout point de I_N a au moins un père dans le niveau I_{N-k} . (Notons ici que le concept de père peut en effet être utilisé car il est défini de manière très générale entre deux niveaux consécutifs, quelles que soient leurs tailles respectives.)

De cette contrainte, on déduit aisément (cf annexe 2) :

- $C(k) \geq 2^N$
- $f \geq k + 1$ et $h \leq N - 2(k + 1)$
- Il y a $2^N \left(\frac{F+1}{2^{k+1}}\right)$ liens supplémentaires par rapport à une structure sans réduction.

Cette dernière remarque induit la nécessité de gérer les recouvrements, lorsqu'un élément de I_N intervient dans le calcul de plus d'une valeur de I_{N-k} . L'avantage principal du recouvrement est qu'il permet une meilleure conservation de la continuité du signal

initial. C'est pourquoi l'on choisit le plus souvent comme candidat au recouvrement les éléments frontières entre deux zones consécutives donnant lieu à deux positions voisines y dans le signal I_{N-k} . Nous avons choisi un type particulier de recouvrement symbolisé par la contrainte suivante :

Contrainte C₄ :

Soit x un élément de I_{N-k} . Si $\#fils(x) \geq m$ alors x partage au plus un fils commun avec chacun de ses voisins sinon, il n'y a pas de recouvrement autorisé. De plus, les fils de x constituent une zone connexe de I_N .

(La valeur seuil, m , permet de n'autoriser le recouvrement qu'à partir d'une taille significative du champ récepteur, zone recouverte par les fils d'un élément donné de I_{N-k} .)

On en déduit les résultats suivants (cf annexe 3) :

- Tout élément de I_N a au plus deux pères au niveau I_{N-k} .

$$F = 2^{k+1}$$

$$L = 2^{N-2(k+1)}$$

$$C(k) = 2^N + 2^{N-k-1}$$

$$m = 1 + 2^k$$

$$k \leq \frac{N}{2} - 1$$

On constate que le système de liens permettant le passage de I_N à I_{N-k} est complètement spécifié. Un exemple, pour $k = 1$, est proposé à la Figure 2.

Ce modèle permet de définir de manière explicite, pour une cellule de position x sur le segment :

– la formule donnant $\#fils(x)$:

$$\#fils(x) = 1 + \frac{\varphi(x)}{L}$$

avec

$$\varphi(x) = |x - fovéa| + \begin{cases} -1 & \text{si } x < fovéa \\ 0 & \text{sinon} \end{cases}$$

– la fonction donnant la localisation du fils le plus à gauche (cf annexe 4) : $loc(x)$

si $x \geq fovéa$

alors

$$loc(x) = \begin{cases} z(x) & \text{si } \#fils(x) < m \\ z(x) - (\varphi(x) - 1) + (m - 1)L & \text{sinon} \end{cases}$$

sinon $loc(x) = 2^N - loc(2^{N-k} - x - 1) - \#fils(x)$

avec

$$z(x) = 2^{N-1} + \#fils(x) \left[\frac{L}{2} (\#fils(x) - 1) + (\varphi(x) \bmod L) \right]$$

et $fovéa = 2^{N-k-1}$.

Les valeurs de $\#fils(x)$ et $loc(x)$ peuvent être calculées *a priori* et n'induisent donc pas de sur-coût lors de l'utilisation de cette pyramide.

Nous pouvons maintenant généraliser simplement ce modèle en deux dimensions. L'ensemble des fils d'une cellule (x, y) est un rectangle de longueur $\#fils(y)$ et de largeur $\#fils(x)$.

Son nombre de fils est donc égal à $\#fils(x) * \#fils(y)$. La cellule en haut à gauche du rectangle des fils a pour coordonnées $(loc(x), loc(y))$.

Ce passage extrêmement simple à la dimension supérieure présente l'avantage de respecter notre restriction à des recouvrements simples. Ici, une cellule a 1, 2 ou 4 pères; les régions rectangles définies dans le niveau fils comme des ensembles de « frères » ne se chevauchent que sur l'épaisseur d'une cellule au plus.

La Figure 3, qui définit le partitionnement du niveau père en régions de même nombre de fils, évoque une image de distance. Il serait donc intéressant de déterminer la fonction (ainsi que la distance discrète sous-jacente) qui définit le nombre de fils en fonction de l'éloignement au centre. On constate aisément que l'on ne peut pas exprimer ces nombres en fonction d'une distance dérivée de la distance de la 4 ou de la 8-connexité si ce n'est par une fonction globalement croissante (cf Figure 4).

Par rapport à l'implantation, il faut noter que l'ensemble des éléments du modèle ne fait intervenir que des nombres entiers, aucune approximation n'est nécessaire lors de l'implantation effective (par opposition au travail présenté dans [13]). De plus, il n'y a aucun effet de bord, *i.e.* l'ensemble des fils de tout élément y de I_{N-k} est complètement inclus dans le niveau I_N . Enfin, dans notre cas, cette implantation a été facilitée par l'emploi de l'environnement de simulation pyramidale développé au sein de notre laboratoire [14,15].

Le mécanisme de construction de la hiérarchie de résolutions irrégulières nécessite la définition d'autant de masques de pondération qu'il y a de tailles de champs récepteur. Ces masques permettent de différencier les comportements des cellules d'un même niveau en tenant compte de leur appartenance à la fovéa ou

16	12	8	4	4	8	12	16
12	9	6	3	3	6	9	12
8	6	4	2	2	4	6	8
4	3	2	1		2	3	4
4	3	2	1		2	3	4
8	6	4	2	2	4	6	8
12	9	6	3	3	6	9	12
16	12	8	4	4	8	12	16

Figure 3. – Carte définissant une partition du niveau père selon le nombre de fils.

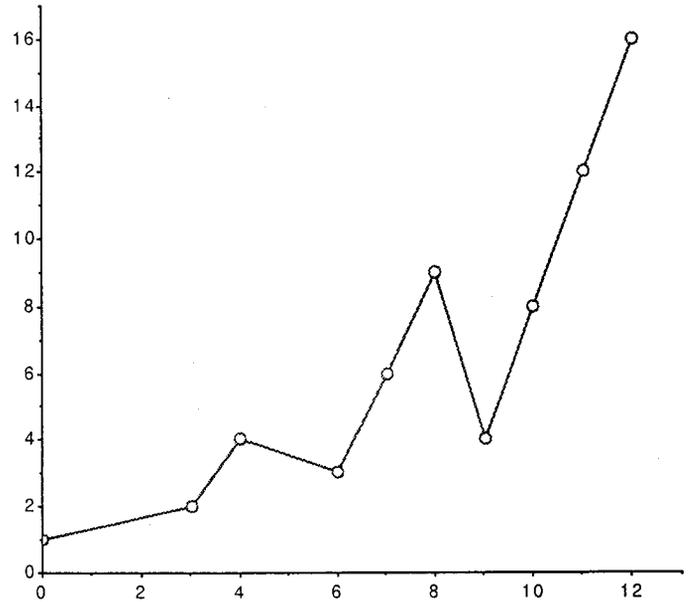


Figure 4. – Graphe du nombre de liens en fonction de la distance 3-4 du chanfrein à la fovéa.

à la périphérie. Notons que de par le mécanisme de construction, ces masques sont séparables.

On peut tout à fait construire une pyramide de type Gaussienne similaire à l'opérateur introduit par Burt [16]. Dans chaque champ récepteur, la valeur $I_{N-k}(y)$ est obtenue par :

$$I_{N-k}(y) = \sum_{x \text{ fils de } y} G(x, y) I_N(y)$$

où $G(x, y)$ désigne les valeurs d'une fonction gaussienne, centrée sur la projection de y dans le niveau N , réduite, et dont l'écart type est également une fonction croissante de y de la forme :

$$\sigma(y) = \sigma(1 + \alpha|y|)$$

Cette structure s'apparente à la structure rétinienne (notion de fovéa et de périphérie). La variation de σ correspond à la variation des caractéristiques de la rétine en fonction de l'excentricité en accord avec le modèle proposé dans [17]. Ce nouveau mécanisme permet de conserver une précision maximale (en regard de l'échantillonnage initial) dans la zone de focus (correspondance 1-1) tout en résumant le restant de la scène. La Figure 5 présente un exemple de réduction irrégulière.

L'image résultat étant de taille réduite, les traitements qui lui sont appliqués sont plus rapides. Cet outil est donc tout à fait approprié dans un contexte d'analyse de scènes par exploration. Cette approche nécessite l'existence d'un système de contrôle du focus, ce qui assure que la fovéa soit implicitement au centre du champ visuel, comme nous en avons fait l'hypothèse dans cette étude. Ce contrôle peut être assuré par un mécanisme logiciel [18] ou matériel [19].



Figure 5. – Réduction irrégulière sur l'image de Léna, les 3 premiers niveaux.
La zone fovéale se situe sur l'œil droit de Léna.

4. Application au mouvement

Pour confirmer le comportement de cette structure irrégulière, nous faisons appel à une caractéristique de la perception du mouvement chez l'homme. En effet, il est connu que la vision fovéale et la vision périphérique ont des sensibilités différentes vis-à-vis d'objets en mouvement dans la scène observée. La vision fovéale est très rapidement saturée par une quantité de mouvement se situant dans son champ récepteur. Cette saturation, *i.e.* hyper-sensibilité, rend cette zone peu adaptée aux mouvements

rapides. Par contre, la vision périphérique, de part son intégration plus importante, est moins sensible et est donc plus adaptée à ces mêmes mouvements (et par conséquent moins adaptée aux mouvements lents). Nous avons reproduit cette propriété du champ visuel sur une séquence d'images réelles montrant trois personnages en mouvement dans le champ de la caméra (cf Figure 6). L'interprétation de cette figure est la suivante :

Première ligne : trois images consécutives de la séquence (128×128). Les deux personnages centraux se déplacent plus lentement que le troisième personnage.

Deuxième ligne : les deux images de différences (valeur absolue point à point, amplifiée par un gain de 4). Les zones de mouvements sont clairement visibles.

Troisième ligne : chaque image de différences est réduite d'un facteur 4 correspondant au deuxième niveau de la structure pyramidale (puis agrandie à la taille des images initiales pour des raisons d'affichage). Cette réduction de résolution est faite par deux mécanismes différents. Les deux images de droite correspondent à une réduction régulière associée à un masque Gaussien (pour plus de détails sur l'évaluation du mouvement utilisant la pyramide gaussienne, voir [7, p. 164-167]). Les deux images de gauche correspondent à la réduction irrégulière proposée dans cet article. Il faut noter que la valeur associée à chaque pixel de ces images est liée à la quantité de mouvement présente dans le champ récepteur du pixel considéré. Ces images sont très utiles pour la détection d'objets en mouvement (image facilement seuillable et de petite taille).

Quatrième ligne : chaque image représente la variation de quantité de mouvement entre les deux couples d'instant (1,2) et (2,3), pour les deux types de liens (irrégulier à gauche et régulier à droite), *i.e.* différence entre les deux images de la ligne précédente. Ces images traduisent la sensibilité à la variation de quantité de mouvement. On constate que la structure de liens régulière fait apparaître deux zones au comportement similaire. Par contre, dans le cas irrégulier, la zone fovéale subit une variation beaucoup plus importante de la quantité de mouvement qui témoigne de la plus grande sensibilité de cette zone. De plus, dans la zone périphérique, la variation, *i.e.* sensibilité, est moins importante dans le cas irrégulier du fait de son champ récepteur plus grand.

5. Discussion

Nous avons proposé dans cet article une nouvelle classe de structures hiérarchiques irrégulières conservant les propriétés de simplicité des pyramides quaternaires (réduction régulière de la taille des images) tout en permettant la mise en oeuvre d'un mécanisme multi-zones (fovéale et périphérique). La structure proposée reste cependant très « régulière » de par le modèle simplifié utilisé et afin d'en faciliter l'implémentation sur machine. Cette structure n'est donc bien sûr pas une piste vers un modèle

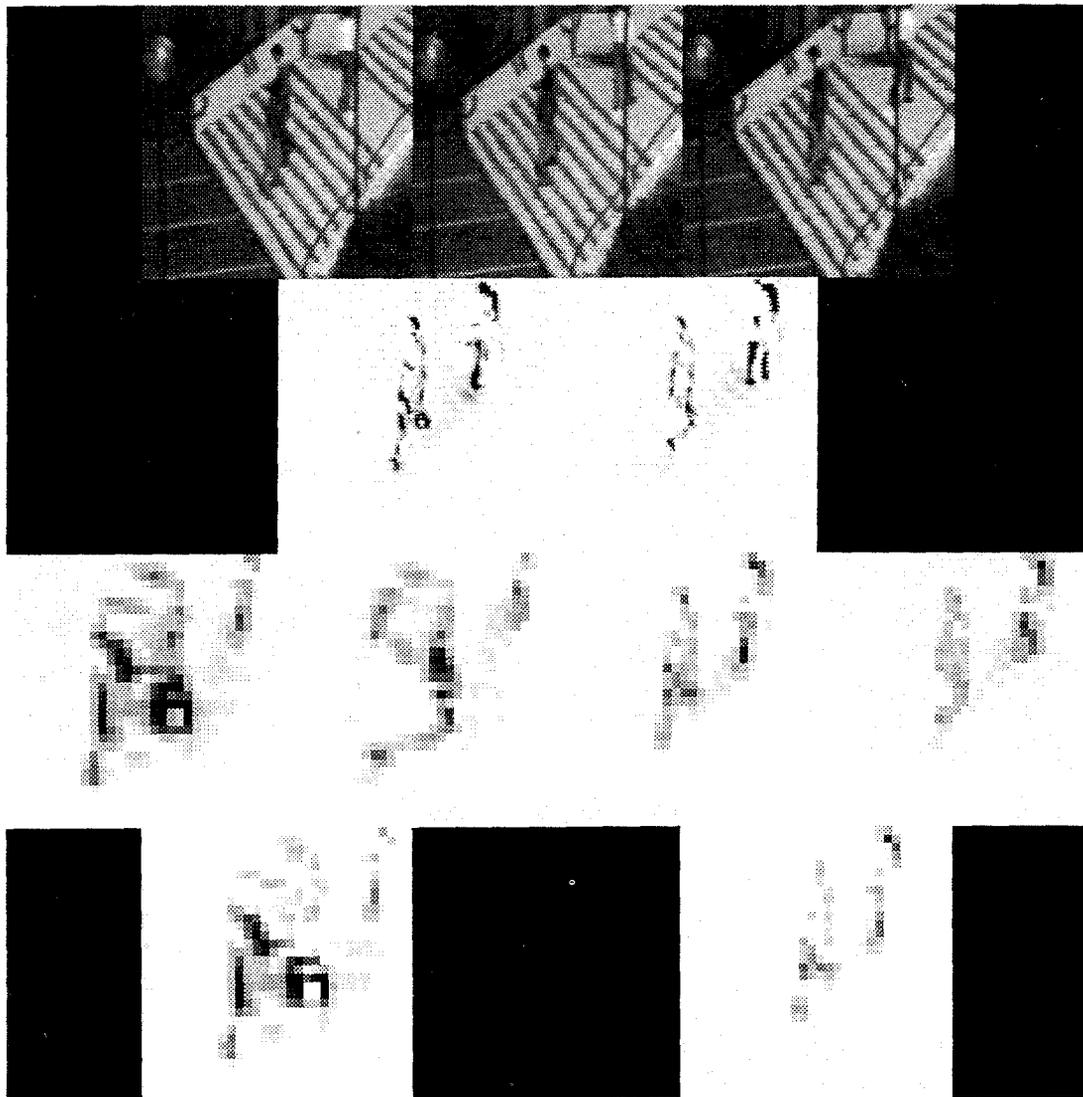


Figure 6. – Application à l'étude du mouvement
(légende : voir texte).

du capteur œil! (Ce n'est en aucun cas notre objectif.) Cependant, elle permet de simuler certains comportements globaux de la vision humaine. Des fonctionnalités plus précises sont obtenues avec des modèles moins rigides comme le montrent les essais de Ahumada où la position des récepteurs est apprise en fonction de l'objectif à atteindre [20]. D'autres exemples relatifs à la pyramide irrégulière, et tout particulièrement une application à la restauration d'images, sont présentés dans [21].

Cette approche doit être vue comme une ébauche dans un travail plus général visant à étudier le comportement des fonctions usuelles utilisées en vision par ordinateur en regard d'une modélisation de tous les acteurs intervenant dans le système incluant un module de vision [22]. Cette démarche est actuellement mise en œuvre dans notre équipe dans le contexte de l'étude du mouvement.

6. Annexes

6.1. EXPLOITATION DE LA CONTRAINTE C_2

D'après C_1 , $\#fils(x) \in [1, F]$. Plus exactement, $\#fils(x)$ prend toutes les valeurs de 1 à F . Chaque valeur est prise $2L$ fois d'après C_2 . Donc, le nombre d'éléments de I_{N-k} est $2LF$. Comme de plus, les images I_m ont pour taille 2^m , on en déduit

$$FL = 2^{N-k-1}$$

Comme F et L sont des entiers, ils sont forcément de la forme $F = 2^f$ et $L = 2^h$ avec comme nouvelle relation :

$$f + h = N - k - 1$$

6.2. EXPLOITATION DE LA CONTRAINTE C_3

Le nombre de liens entre les deux niveaux I_N et I_{N-k} est noté $C(k)$. I_{N-k} peut se décomposer en F régions R_i définies par :

$$R_i = \{x \in I_{N-k} : \#fils(x) = i\}$$

On obtient donc

$$C(k) = \sum_{i=1}^F Card(R_i) i = 2L \sum_{i=1}^F i = F(F+1)L$$

Soit, en tenant compte des relations déduites de C_2 :

$$C(k) = (F+1)2^{N-k-1}$$

La contrainte C_3 impose qu'il y ait au moins un lien vers chaque élément de I_N de taille 2^N donc

$$C(k) \geq 2^N \Rightarrow 2^f + 1 \geq 2^{k+1}$$

donc $f \geq k+1$

et comme $f+h = N-k-1$, on obtient $h \leq N-2(k+1)$.

6.3. EXPLOITATION DE LA CONTRAINTE C_4

f étant un entier, l'inégalité $2^f + 1 \geq 2^{k+1}$ (cf annexe 2) ne peut se réduire à une égalité qui induirait un et un seul lien pour chaque élément de I_N . Il y a donc forcément des éléments multiples (*i.e.* éléments ayant plus d'un père). En fait, $FL(F+1) - 2^N$ est le nombre de liens « en trop » par rapport à une structure sans réduction.

La contrainte C_4 est schématisée par la figure 2. Il existe un fils commun pour tout élément de chaque région R_i , pour $i \geq m$. Le nombre de fils commun est donc donné par : nombre de régions R_i de taille supérieure ou égale à m fois la taille d'une région R_i , soit

$$(F-m+1)2L$$

Ce qui conduit à

$$\begin{aligned} (F-m+1)2L &= FL(F+1) - 2^N \\ 2FL - 2L(m-1) &= FL + F^2L - 2^N \\ FL(1-F) + 2^N &= 2L(m-1) \\ m &= 1 + \frac{2^{N-1}}{L} + \frac{F}{2}(1-F) \end{aligned}$$

F et L étant des puissances de deux, on obtient

$$m = 1 + 2^{N-1-h} + 2^{f-1}(1-2^f)$$

Comme $h = N-k+1-f$, on a

$$m = 1 + 2^f [2^k + \frac{1}{2} - 2^{f-1}]$$

On a vu précédemment (cf annexe 2) que $f \geq k+1$. Posons $f = k + \alpha$ (avec $\alpha \geq 1$). On obtient alors

$$\begin{aligned} m &= 1 + 2^{k+\alpha} [2^k + \frac{1}{2} - 2^{k+\alpha-1}] \\ m &= 1 + 2^{k+\alpha} [\frac{1}{2} - 2^k(2^{\alpha-1} - 1)] \end{aligned}$$

La variable m est strictement positive, ce qui conduit à

$$2^k(2^{\alpha-1} - 1) \leq \frac{1}{2}$$

Comme $k \geq 1$, cette relation induit que la seule valeur possible pour α est $\alpha = 1$. On peut donc, tenant compte des relations déjà établies, déduire

$$\begin{aligned} f = k+1 &\Rightarrow F = 2^{k+1} \\ h = N - 2(k+1) &\Rightarrow L = 2^{N-2(k+1)} \\ m &= 1 + 2^k \\ C(k) &= (2^{k+1} + 1)2^{N-k-1} = 2^N + 2^{N-k-1} \end{aligned}$$

La taille de chaque région R_i est strictement positive, ce qui induit $L \geq 1$ (*i.e.* $l \geq 0$) donc

$$k \leq \frac{N}{2} - 1$$

6.4. DÉTERMINATION DE $loc(x)$

Considérons un élément x de I_{N-k} se situant à droite de la fovéa et avant le début des recouvrements. Son fils le plus à gauche a pour position $loc(x)$ telle que $loc(x) \geq 2^{N-1}$. $x \in R_{\#fils(x)}$, il faut donc tenir compte de toutes les régions R_i telles que $i < \#fils(x)$, ainsi que des autres éléments de la région $R_{\#fils(x)}$. Ceci conduit à l'expression

$$z(x) = 2^{N-1} + L \sum_{i=1}^{\#fils(x)-1} i + (\varphi(x) \bmod L) \#fils(x)$$

Soit

$$z(x) = 2^{N-1} + \#fils(x) \left[\frac{L}{2} (\#fils(x) - 1) + (\varphi(x) \bmod L) \right]$$

Cependant, cette relation ne donne $loc(x)$ que pour les régions sans recouvrement (*i.e.* $\#fils(x) < m$). Dans les autres cas, il faut retrancher les éléments multiples. On obtient alors

$$loc(x) = \begin{cases} z(x) & \text{si } \#fils(x) < m \\ z(x) - (\varphi(x) - 1 - (m-1)L) & \text{sinon} \end{cases}$$

Le cas des éléments situés à gauche de la fovéa est résolu grâce à la symétrie du système de coordonnées. On obtient dans ce cas

$$loc(x) = 2^N - loc(2^{N-k} - x - 1) - \#fils(x)$$

Le dernier terme de cette expression ($-\# \text{ fils}(x)$) est retranché de telle sorte que $\text{loc}(x)$ désigne toujours la position du fils de plus à gauche.

BIBLIOGRAPHIE

- [1] Y. Yeshurun, E.L. Schwartz, « Shape Description with a Space-Variant Sensor : Algorithms for Scan-Path, Fusion and Convergence over Multiple Scans », *IEEE Trans. Pattern Anal. Machine Intel.*, 11 1989, p. 1217-1221.
- [2] A. Califan, R. Kjeldsen, R.M. Bolle, « Data and Model Driven Foveation », *Proceeding of the 10th Int. Conf. on Pattern Recognition*, 1990, vol. A, p. 1-7.
- [3] S.M. Culhane, J.K. Tsotsos, « An Attentional Prototype for Early Vision », *2nd European Conf. on Computer Vision*, Lecture Notes in Computer Vision, G. Sandini (ed.), Springer Verlag, 588, 1992, p. 551-560.
- [4] J.K. Tsotsos, « Towards a Computational Model of Visual Attention », *Linking Psychophysics, Neurophysiology, and Computational Vision*, T. Pappas (ed.), MIT Press, Bradford books, 1994.
- [5] M. Bolduc, M.D. Levine, « A Foveated Retina System for Robotic Vision », *Vision Interface'94*, 1994, p. 1-8.
- [6] J.M. Jolion, « Analyse d'Images : Le Modèle Pyramidal », *Traitement du Signal*, 7, n° 1, 1990, p. 5-17.
- [7] J.M. Jolion, A. Rosenfeld, *A Pyramid Framework for Early Vision*, Kluwer, 1993.
- [8] S. Peleg, O. Federbush, R.A. Hummel, « Custom-Made Pyramids », *Parallel Computer Vision*, L. Uhr. (ed.), Academic Press, Orlando, FL, 1986, 125-146.
- [9] P. Meer, « Stochastic Image Pyramids », *Computer Vision, Graphics, and Image Processing*, 45, 1989, p. 269-294.
- [10] A. Montanvert, P. Meer, A. Rosenfeld, « Hierarchical Image Analysis Using Irregular Tessellations », *IEEE Trans. Pattern Anal. Machine Intel.*, 13, 1991, p. 307-316.
- [11] J.M. Jolion, A. Montanvert, « The Adaptive Pyramid : A Framework for 2D Image Analysis », *CVGIP : Image Understanding*, 55, n° 3, 1992, p. 339-348.
- [12] W.G. Kropatsch, « A Pyramid that Grows by Powers of 2 », *Pattern Recognition Letters*, 6, 1985, p. 179-284.
- [13] W. Bruger, « Image Warpings, Fixations, and Structural Recognition », *Proc. of the 16th ÖAGM meeting on Pattern Recognition*, 1992, p. 199-213.
- [14] S. Bataouche, « Définition d'une Organisation Logicielle Pyramidale », *Rapport de Recherche RR 22-90, LISPI*, Université C. Bernard, Lyon 1, 1990.
- [15] S. Bataouche, « Raisonnement Hiérarchique en Vision par Ordinateur », *thèse de doctorat*, Université C. Bernard, Lyon 1, 1993.
- [16] P.J. Burt, « Fast Filter Transforms for Image Processing », *Computer Graphics and Image Processing*, 16, 1981, p. 20-51.
- [17] H.R. Wilson, J.R. Bergen, « A Four Mechanism Model for Threshold Spatial Vision », *Vision Research*, 19, 1979, p. 19-32.
- [18] J.K. Tsotsos, « An Inhibitory Beam for Attentional Selection », *Spatial Vision for Humans and Robot*, L. Harris et M. Jenkin (eds), Cambridge University Press, 1993.
- [19] D.H. Ballard, C.M. Brown, « Principles of Animate Vision », *CVGIP : Image Understanding*, 56, n° 1, 1992, p. 3-21.
- [20] A.J. Ahumada, « Learning Receptor Positions », *Computational Models of Visual Processing*, M.S. Landy et J.A. Movshon (eds) MIT Press, Bradford books, 1991, p. 23-34.
- [21] A. Manzanera, J.M. Jolion, « Tessellations Hiérarchiques Irrégulières », *Colloque Géométrie Discrète en Imagerie*, 1993, p. 98-107.
- [22] J.M. Jolion, « Computer Vision Methodologies », *CVGIP : Image Understanding*, 59, n° 1, 1994, p. 53-71.

Manuscrit reçu le 5 Mai 1994.

LES AUTEURS

Antoine MANZANERA



Antoine Manzanera a obtenu le DEA d'Informatique Fondamentale de l'ENS de Lyon en 1993 au sein de l'équipe LIGIA du LISPI dans le domaine de la vision artificielle. Il effectue depuis son service militaire et poursuivra par la suite en thèse de Doctorat ses travaux sur les structures discrètes en vision par ordinateur.

Jean-Michel JOLION



Jean-Michel Jolion est ingénieur de l'INSA de Lyon (option informatique, 1984). Thèse de Doctorat en Informatique (analyse d'images en microscopie électronique, 1987). Séjour post-doctoral au Computer Vision Laboratory (Université du Maryland, 1987-1988). Professeur à l'INSA de Lyon dans le laboratoire Reconnaissance de Formes et Vision. Activités de recherche dans le domaine de la perception visuelle en liaison avec la vision humaine (en particulier la multirésolution et l'architecture pyramidale) et sur la méthodologie de conception des systèmes de vision artificielle. Co-auteur du livre *A Pyramidal Framework for Early Vision*, avec le Professeur Rosenfeld.