

CARTE PYRAMIDE D'IMAGE DEMIE OCTAVE EN TEMPS RÉEL VIDÉO

R. KOUDSIEH, P.Y. COULON et G. BOUVIER

Laboratoire de Traitement d'Images et Reconnaissance de Formes
INPG, 46 Av. Félix Viallet, 38031 Grenoble Cédex, France

RÉSUMÉ

ABSTRACT

Cet article décrit la réalisation d'une carte électronique pour la construction en temps réel vidéo d'une pyramide demie octave d'image stéréoscopique.

L'algorithme adopté utilise un noyau de convolution binomial. Nous présentons les différents algorithmes utilisés dans la littérature; les rapports signal sur bruit et les fréquences de coupure sont comparées à ce que nous obtenons avec notre méthode.

La carte électronique réalisée est au format double Europe et interfacée par un bus VME standard. Le temps de construction de la pyramide est inférieur à 40 ms pour des images stéréoscopiques de 2256x256 pixels représentées sur 6 bits. Le rapport signal sur bruit dû au repliement de spectre est supérieur ou égal à 57 dB.

We describe, in this paper, the implementation of an electronic board to build a semi octave pyramid of a stereoscopic image in video real time.

The adopted algorithm uses a binomial convolution kernel. We present the most well known image octave pyramid construction algorithms found in the literature; noise to signal ratios and cut-off frequencies are compared with our results.

The realised electronic board is in double Europe format interfaced with a VME standard bus. Construction of such a pyramid requires less than 40 ms for stereoscopic images of 256x256 pixels coded with 6 bits. Noise to signal ratio due to spectral overlapping is greater than 57 dB.

1. INTRODUCTION

Dans de nombreuses applications de traitement d'images, il n'est pas nécessaire de connaître les détails dans l'image traitée. Il est alors préférable de travailler sur des images de résolution plus faible pour obtenir un gain de temps. La représentation multirésolution (ou multiéchelle) permet de choisir le facteur de résolution voulu, d'où son intérêt dans des structures algorithmiques du type focalisation sur une région d'intérêt ou "coarse to fine analysis" [1].

L'obtention d'une pyramide pose évidemment le problème du temps de construction. Les opérations de convolution peuvent être assez longues suivant le type de filtrage employé; le but que nous nous sommes fixé est de mettre en oeuvre une carte électronique qui fournirait en temps réel vidéo non pas l'image issue de la caméra, mais la pyramide correspondant à cette image.

Pour être performante cette carte doit utiliser un algorithme de construction simple et rapide sans bien sûr négliger la qualité (rapport signal sur bruit) de la pyramide obtenue.

Ce travail entre dans le cadre du projet Esprit BRA 3038 " Vision As Process " (VAP).

2. PYRAMIDE D'IMAGE PASSE-BAS

C'est une suite d'images, chaque image est déduite de l'image précédente par un filtrage et un sous échantillonnage. Le principe est le suivant: étant donnée une image dont la pulsation maximale est de π , une séquence de convolution de cette image par un filtre passe-bas permet de réaliser d'une manière contrôlée une réduction de la fréquence maximale contenue dans cette image. Lorsque la fréquence maximale a été réduite dans un rapport de 2 (pulsation égale à $\pi/2$), on peut procéder à un sous échantillonnage tout en restant en accord avec le théorème de Schannon. Donc une séquence de filtrage suivie par un sous échantillonnage permet d'obtenir à partir d'une image I_0 de taille $N \times N$ (bande passante $[0.. \pi]$), une suite d'images I_1 : $N/2 \times N/2$ (bande passante $[0.. \pi/2]$), I_2 : $N/4 \times N/4$ (bande passante $[0.. \pi/4]$), I_3 : $N/8 \times N/8$ (bande passante $[0.. \pi/8]$) ..., à chaque fois on divise la bande passante par un facteur deux. La représentation fréquentielle dans le cas idéal est montrée sur la figure (1) ci-dessous:

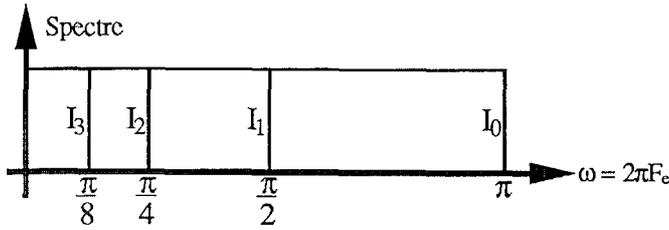


Fig (1) Représentation fréquentielle de la pyramide.
Fe: fréquence d'échantillonnage.

3.CONSTRUCTION DE LA PYRAMIDE

3.1. Méthode classique:

Le schéma usuel utilisé dans les algorithmes de production d'une pyramide d'image est représenté par l'itération de l'opérateur de base décrit dans la figure (2) suivante:

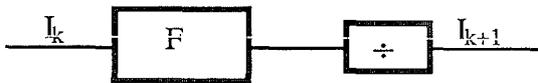


Fig (2) Opérateur de construction de la pyramide.
(F: filtrage; ÷: sous échantillonnage).

Le choix du filtre est un problème délicat, le filtre idéal qui divise la bande passante de l'image par 2 n'existe pas physiquement, donc il faut le remplacer par un filtre qui se rapproche au mieux du filtre idéal; ce type de filtre est très difficile à réaliser car ses coefficients sont réels. Tout en respectant les lois de l'échantillonnage nous avons choisi d'utiliser un filtre moins performant mais plus facile à implanter électroniquement. Le filtre binomial possède des propriétés intéressantes: coefficients entiers, séparable (deux convolutions en une dimension à la place d'une convolution en deux dimensions). Le filtre binomial, en une dimension est de la forme $g = [1 \ 2 \ 1]$. Ce filtre se rapproche d'un gaussien de variance $\sigma^2 = 0.5$.

Le spectre de ce filtre est de la forme $(1+\cos\omega)/2$ avec une amplitude du lobe secondaire égale à 1/2 ce qui correspond à un gain de -6 dB, donc un phénomène d'aliasing important lors du sous échantillonnage, pour le diminuer, on prend comme filtre g^{*n} (g convolué n fois), ce qui donne un gain de -6n dB. Par exemple pour obtenir un gain de -24 dB, il faut utiliser un noyau de dimension 9x9 (n=4). La figure (3) montre la réponse spectrale de ce filtre pour n = 1, 2, 3, 4.

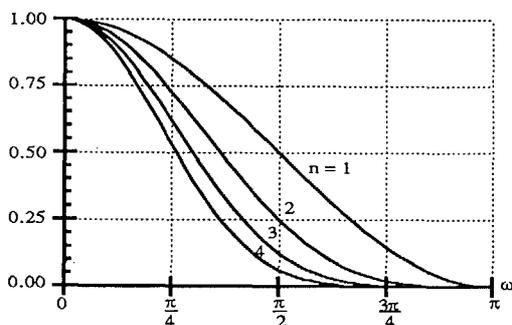


Fig (3)

Dans la littérature, Clark et Lawrence [2] utilisent un noyau de 25x25, pour obtenir un gain de -33 dB.

Un autre exemple, Crowley et Stern [3] utilisent des noyaux gaussiens de dimensions 7x7 et 13x13 pour assurer un gain de -54 dB.

3.2. Méthode de Burt:

L'algorithme de Burt [4,5] est donné figure (4), où le sous échantillonnage est réalisé avant le filtrage.

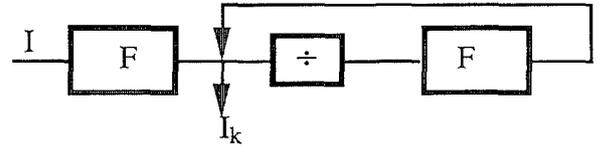


Fig (4) Algorithme de Burt

Le masque de convolution pour une dimension est toujours $[1 \ 2 \ 1]$. Le sous échantillonnage suivi par un filtrage F est équivalent à un filtre G suivi par un sous échantillonnage où le spectre de G est donné par: $(1+\cos\omega)/2$. La réponse spectrale résultante est égale au produit $(1+\cos\omega).(1+\cos2\omega)/4$. Lors de la première itération le terme $(1+\cos\omega)$ est dû au filtrage initial, ce qui est illustré sur la figure (5) ci-dessous:

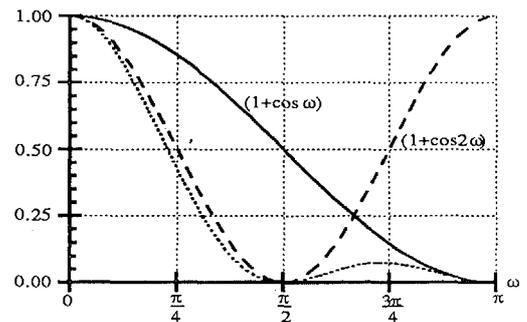


Fig (5)

Comme le montre la figure (5) il y a un repliement de spectre dû au sous échantillonnage. Dans ce cas là, si on calcule le rapport signal sur bruit on trouve $S/B = -22.6$ dB. Pour augmenter ce rapport, on prend $F = g^{*n}$ où n est le nombre de convolution du noyau $g = [1 \ 2 \ 1]$ effectué sur l'image. Pour n = 2 on obtient un rapport S/B égal à -45 dB, le noyau de Burt a alors la forme suivante:

$$F(x,y) = \frac{1}{16^2} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}^{*2}$$

En calculant le rapport signal sur bruit d'une telle pyramide, ainsi que la fréquence de coupure correspondante, déterminée par la variance du filtre binomial σ^2 , pour chaque niveau de pyramide on obtient les résultats suivants (table I):

Niveau	0	1	2	3
S/B(dB)	-	45	51	52
σ^2	1	5	21	85

Table I

3.3. Pyramide demie octave à variance juste:

L'algorithme décrit précédemment ne fournit pas une pyramide d'octave juste; en effet la suite des variances obtenues n'est pas de la forme 2^{2n} . D'autre part il peut être intéressant de disposer de niveaux intermédiaires de pyramide, c'est à dire d'avoir une résolution non plus en octave mais en demie octave. Les bandes passantes des images obtenues seront alors de

$[0..π] [0..π/\sqrt{2}] [0..π/2] [0..π/2\sqrt{2}] [0..π/4]$ etc, correspondant à un filtrage gaussien de variances $\sigma^2 = 1, 2, 4, 8, 16, 32, \dots$ On retrouve alors une pyramide d'octave haute fréquences $\sigma^2 = 1, 4, 16, \dots$ entrelacée avec une autre pyramide d'octave décalée en fréquence et de variances $\sigma^2 = 2, 8, 32, \dots$

La structure de cet algorithme, proposé par A. Chehikian et J.L. Crowley, est illustrée sur la figure (6):

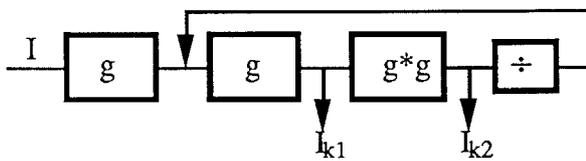


Fig (6) Algorithme d'une pyramide demie octave

La figure (7) montre la réponse en fréquence du noyau utilisé et la table II montre le rapport signal sur bruit et la fréquence de coupure pour chaque niveau de la pyramide.

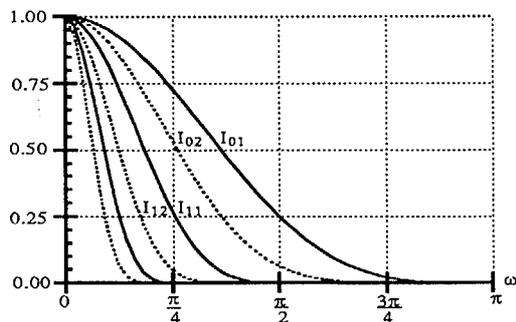


Fig (7) Réponses en fréquence pyramide haute fréquence (en continu). Pyramide basse fréquence (en pointillés)

Niveau	0	1	2	3
S/B(dB)	-	58	82	57
σ^2	1	2	4	8
			16	32
				54
				128

Table II

On observe qu'on a toujours un rapport signal sur bruit supérieur ou égal à 57 dB, ce qui est meilleur à comparer avec les 52 dB pour Burt. En ce qui concerne la variance, il y a moins de perte d'information dans cet algorithme que celui donné par Burt; car on a pour tous les niveaux de la pyramide:

$$\frac{F_{\max}}{\sqrt{\sigma_{\text{Burt}}^2}} < \frac{F_{\max}}{\sqrt{\sigma_{\text{Ltif}}^2}} \text{ avec } F_{\max} \text{ la fréquence maximale de l'image.}$$

4. REALISATION

La carte est au format double Europe VME (figure 8), elle permet de faire :

- L'acquisition des images stéréos.
- La construction de la pyramide proprement dite.
- Stockage des images produites dans les mémoires.

Elle est structurée par les divers sous-ensembles suivants:

- Deux caméras: l'une fonctionnant en maître et l'autre en esclave, synchronisée par la première.
- Deux convertisseurs analogique-numérique ADC1 et ADC2.
- Quatre convolveurs
- Mémoires d'images.
- Un générateur d'adresses
- Une interface au standard VME.

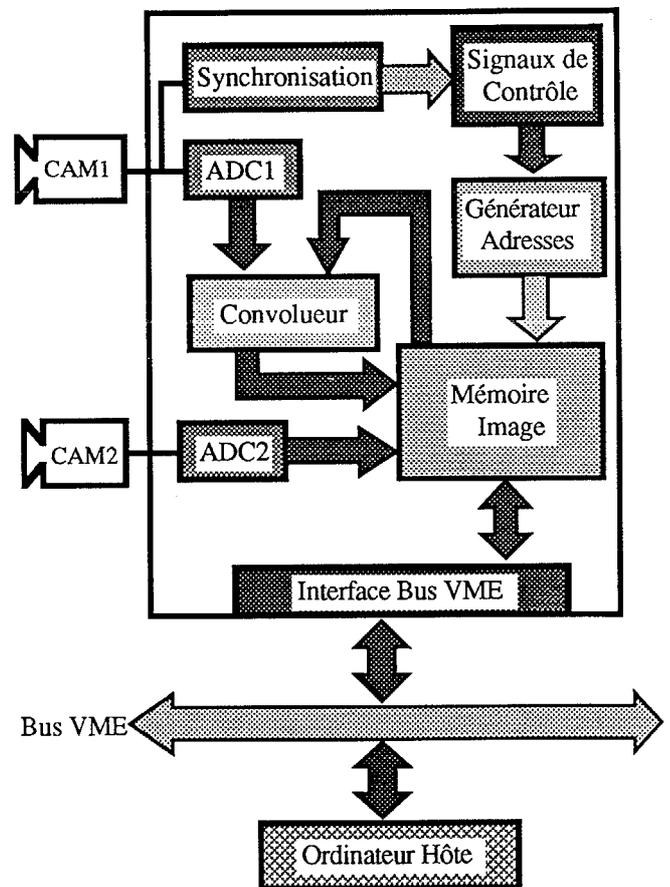


Fig (8) Schéma synoptique de la carte

*Les convertisseurs analogique-numérique numérisent le signal vidéo sur 6 bits.

*Les convolveurs sont des circuits intégrés spécifiques développés dans le laboratoire pour cette application. Le noyau de convolution choisi conformément aux spécifications du filtre binomial est:

$$g = [1 \ 2 \ 1]^* \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$



*Les mémoires image permettent de stocker les différents niveaux de la pyramide (il y a au total quatre pyramides, deux pyramides pour chaque caméra). Ces mémoires sont accessibles soit par l'opérateur, soit par le bus VME.

* Vu du bus VME par l'ordinateur hôte, la carte se présente comme un module esclave. L'accès au bus VME est interdit pendant la construction de la pyramide.

5. PERFORMANCES

On utilise des images stéréoscopiques de 256x256 pixels codées sur 6 bits. Dans un premier temps, et en parallèle avec l'acquisition, on calcule le premier niveau de la pyramide de la première caméra et on stocke l'image issue de la deuxième caméra; une fois cette étape terminée, on calcule les différents niveaux de la première pyramide puis de la deuxième et cela dans un temps inférieur à 40 ms (acquisition incluse) c'est à dire avant l'arrivée de la deuxième image, ce qui nous permet de travailler en temps réel vidéo. Voir le chronogramme figure (9) ci-dessous:

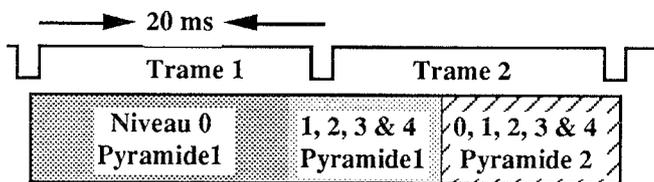


Fig (9)

La figure (10) nous montre un résultat de construction d'une pyramide demie octave.



Fig (10a) Pyramide demie octave hautes fréquences

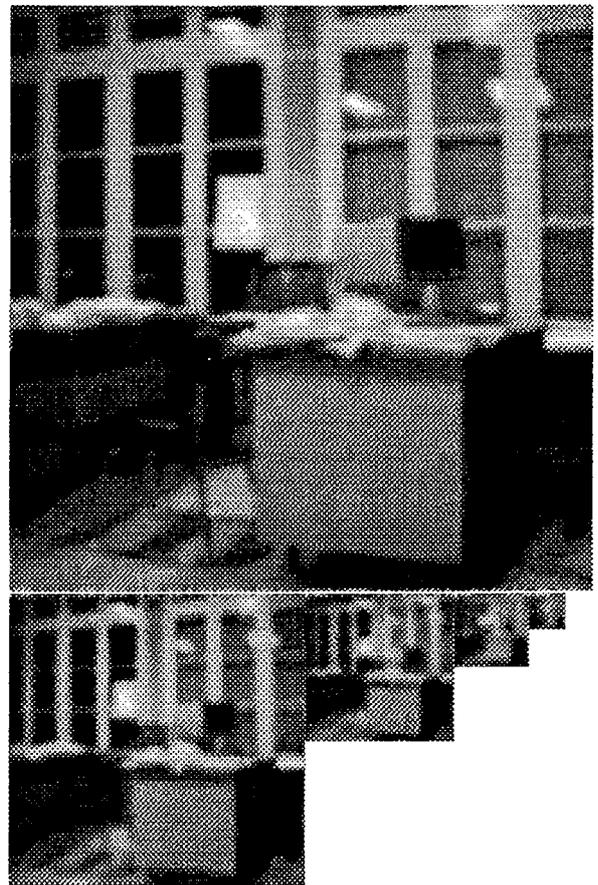


Fig (10b) Pyramide demie octave basses fréquences

6. CONCLUSION

L'algorithme utilisé (demie octave) possède des propriétés intéressantes à savoir: rapport signal sur bruit très élevé, toujours supérieur à 57 dB, et permet la construction de la pyramide en temps réel vidéo.

La carte réalisée, au format double Europe, est compatible au bus VME. Elle constitue l'élément de base de la machine de vision développée dans le cadre du projet Esprit VAP. En effet c'est à partir des données disponibles sous forme de pyramide que travailleront les autres cartes constituant la machine (détection d'évènement, extraction d'indices visuels, ...).

7. REFERENCES

- [1] Charles R. Dyer, " Multiscale image understanding ", in Parallel Computer Vision, L. Uhr (ed), New York: Academic Press, (1987).
- [2] Clark J. J., Lawrence P. D. " A systolic parallel processor for the rapid computation of multiresolution edge images.
- [3] Crowley J. L., Stern R. M., " Fast computation of the difference of low-pass transform ", IEEE trans. on P.A.M.I., 6, 212-222, (1984).
- [4] Burt P. J., " Fast filtre transform for image processing ", Computer Graphics and Image Processing, 16, 20-51, (1981).
- [5] Burt P. J., " The Laplacien as a Compact Image Code", IEEE trans. on P.A.M.I., 4, 532-540, (1983).