

DIXIEME COLLOQUE SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 20 au 24 MAI 1985

AMÉLIORATION EN CODAGE D'IMAGES D'UNE MÉTHODE A DEUX COMPOSANTES

J.F. BERTRAND*, D. BARBA**, J. RODRIGUEZ*

*Laboratoire d'Automatique - I.N.S.A. - 20, Avenue des Buttes de Coësmes - 35043 RENNES CEDEX
**C.C.E.T.T. - Rue du Clos Courtel B.P. 59 - 35510 CESSON SEVIGNE

RESUME

Ce papier concerne l'amélioration des performances d'un système de codage d'images par une méthode à deux composantes. Le principe d'un tel système repose sur la séparation du signal d'image monochrome en deux composantes, la composante par plateau et la composante de texture, liées chacune à l'une des deux caractéristiques principales d'une image : les contours et les textures. La méthode de codage repose sur une segmentation de l'image en zones homogènes fermées à partir d'une procédure de détection, de suivi et de fermeture des contours. Nous décrivons un certain nombre d'améliorations à la méthode précédemment proposée, dans le sens tant d'une amélioration de la qualité que d'une réduction de débit. Ces améliorations portant d'une part sur la description de la géométrie des régions : les contours peuvent être représentés, soit par segments de droite, soit codés directement, le choix s'effectuant selon un nouveau critère. D'autre part, chaque région est découpée en sous-régions pour lesquelles les variations spatiales de la première composante (composante par plateau) sont linéaires : représentation par un plan à luminance constante où à variations linéaires avec les coordonnées. Un accroissement très sensible de la qualité est ainsi obtenu avec, généralement, un accroissement du taux de compression.

SUMMARY

This paper deals with the problem of increasing the performances of a two-component coding system. That system is based upon the decomposition of the grey-level signal into two additive components : the plateau luminance signal and the texture signal. They take into account the properties of the two fundamental characteristics of the image signal : edges and texture. Coding relies upon segmentation of the image field into connected homogeneous regions extracted by edge-detection and contour-following techniques.

In order to increase both the visual quality and the compression rate, we described here some modifications in the coding system. The first one is connected with the description of the boundaries of the regions : contours can now be described either by piecewise linear segments or directly coded by Freeman code, depending of a new criterion. In the second one, every region is cut up into sub-regions in which variations of the plateau component are linear with the spatial coordinates.

We obtained increased performances both in compression rate and, perhaps more sensitive, in visual quality.



INTRODUCTION

Le but premier de tout système de codage d'image à compression de l'information est de diminuer la quantité d'éléments (binaire) nécessaires à la représentation de l'information contenue dans l'image à coder. Cette réduction peut s'opérer en exploitant les redondances de type statistique, liées aux corrélations spatiales (spatio-temporelles pour les séquences d'images), et les redondances de type psychovisuel, liées aux informations non pertinentes. Diverses interprétations possibles d'une bonne exploitation des redondances ont conduit au développement de différentes familles de schémas de codage telles que les codages prédictifs (MICD), les codages par blocs avec transformation (DCT Haar-Hadamard, K-L, ...) ou non (codage B.T.C.), les codages hybrides, les codages statistiques. Dans leur principe, la plupart de ces systèmes de codage s'appuient sur un modèle plus ou moins stationnaire du signal d'image. Cependant, comme cette propriété n'est généralement pas du tout vérifiée, des techniques d'adaptation plus ou moins importantes ont été élaborées de façon à améliorer leurs performances (détection et suivi de contour en MICD, modification du nombre et de la loi de codage des coefficients en codage par transformation, ...).

Parallèlement, mais plus récemment, une analyse fine du mécanisme de la formation des images a conduit certains chercheurs à modéliser le signal d'image monochrome sous la forme de deux composantes additives : la première, discontinue, représente les variations d'intensité liées à la présence d'objets distincts dans la scène, l'autre, continue, correspond à la texture. Chaque composante est codée séparément et les deux composantes sont recombinaées à la réception.

A partir de ce modèle, des formes plus ou moins simplifiées avec modèle unidimensionnel ont conduit aux premiers systèmes de codage par plages et au système de codage par composantes de SCHREIBER [1] avec séparation des deux composantes par filtrage linéaire. Des systèmes plus élaborés avec traitements unidimensionnels et extrapolation de ligne à ligne ont ensuite été proposés [2],[3]. Cependant, les performances restent limitées par le fait que les prétraitements privilégient les structures horizontales ou verticales. Une extension à deux dimensions a été étudiée par KOCHER et KUNT [4] en utilisant une méthode de segmentation d'images par croissance de régions. Pour le codage, ils classifient les points d'images en deux catégories : les points contour et les points texture. Les points contour délimitent par leur connexité des

régions (pavés) à l'intérieur desquelles l'évolution du niveau de gris est donnée par les points texture. Le signal de texture est à son tour séparé en deux signaux dont l'un représente l'évolution de la valeur moyenne locale du niveau de gris à l'intérieur des pavés, l'autre représente la granularité. Nous avons proposé récemment [5] un système de codage analogue, basé sur un modèle à deux composantes et reposant également sur une segmentation de l'image en zones homogènes fermées. Les différences essentielles portent à la fois sur la méthode de segmentation, basée sur la détection des contours, et sur la définition et le codage des deux composantes. L'objet de ce papier est de décrire un certain nombre d'améliorations que nous avons apporté à la méthode que nous proposons, études s'inscrivant dans une démarche d'optimisation des performances du codage, dans le sens tant d'une amélioration de la qualité que d'une réduction de débit.

Avant de présenter ces améliorations, il nous faut rappeler les caractéristiques essentielles du schéma de codage que nous avons proposé.

II - CODAGE A DEUX COMPOSANTES

En chaque point d'image repéré par ses coordonnées spatiales (m : indice de ligne, n : indice de colonne), le signal de luminance $I(m,n)$ se décompose en deux composantes additives :

$$I(m,n) = I_p(m,n) + I_T(m,n) \quad (1)$$

où : - la composante par plateau I_p supporte l'information de valeur moyenne locale constante ou à variations spatiales simples (bilinéaire) et de saut de valeur moyenne locale sur les zones de contour.

- la composante de texture I_T supporte les variations du signal d'image autour de la composante par plateau. Ces variations peuvent être aussi bien à caractère plutôt ordonné (texture de type structuré) qu'à caractère plutôt désordonné (texture de type aléatoire).

Des considérations sur le processus de formation des images amènent à considérer celles-ci comme constituées d'un ensemble de régions connexes à l'intérieur desquelles la distribution spatiale du signal de luminance est homogène. Les frontières des régions correspondent aux ruptures dans le signal d'image (contours). La représentation du signal d'image doit donc comporter la description de la géométrie des régions homogènes ainsi que celles des signaux I_p et I_T à l'intérieur de celles-ci. L'ensemble du schéma de codage est représenté sur la figure 1-a. Il est composé de 3 sous-ensembles correspondant successivement à la segmentation du plan



AMÉLIORATION EN CODAGE D'IMAGES D'UNE MÉTHODE A DEUX COMPOSANTES

de l'image par détection des contours, le codage de la composante par plateau, le décodage de celle-ci et le codage de la composante de texture.

Segmentation en régions homogènes

Le principe utilisé est la détection et le suivi des contours. L'ensemble du processus de segmentation utilise successivement la détection des points de contour par opérateurs de convolution de type gradient orienté, seuillage et amincissement des contours, puis formation de chaînes de contours par suivi et fermeture des contours. Cette dernière phase, assez délicate, permet d'obtenir des régions complètement fermées. Enfin un processus de transformation de représentation permet d'obtenir une représentation structurée et hiérarchisée de la topologie des régions sous forme de 3 tables (tables des régions, des arcs, des noeuds) remplaçant celle de l'image des points contour et mieux adaptée à la représentation des connaissances nécessaires au codage des deux composantes.

Codage de la composante par plateau

Le choix que nous avons fait d'un décodage simple, nous a amené à réduire les formes complexes obtenues dans la phase de segmentation à des formes simples de type polygones, d'où l'approximation des contours par segments de droite. De plus, pour simplifier au maximum le décodage, les polygones (souvent complexes) sont eux-mêmes décomposés en sous-polygones convexes (facettes) avec production pour chacun d'eux d'un bloc d'information (descripteur) contenant une description de la géométrie de la facette et des niveaux de gris moyen aux sommets de celle-ci. La composante par plateau est considérée être à variations linéaires suivant les côtés des facettes (arêtes) et suivant la direction horizontale (balayage TV), ce qui conduit à une méthode de décodage très simple, analogue à celle utilisée en synthèse d'images.

Codage de la composante de texture

La seconde composante est obtenue au codeur par différence entre l'image originale et la composante par plateau décodée. Cette dernière composante comporte donc le véritable signal de texture auquel s'ajoute un signal résiduel \hat{Y}_p dû aux erreurs de codage de la composante par plateau. Elle est codée par un système de codage par blocs sans transformation de type B.T.C., mais dont les formes des blocs correspondent aux polygones convexes contenus dans les descripteurs de la première composante. Dans le décodeur (figure 1-b), on

synthétise la composante par plateau suivant le même schéma que celui du codeur et le signal de texture décodé lui est ajouté pour former l'image complète restituée.

Les figures 2-a à 2-e permettent de montrer les résultats obtenus avec ce schéma de codage. La figure 2-b est l'image des contours obtenus par segmentation de l'image fig. 2-a. La figure 2-c montre l'approximation polygonale et le découpage en polygones convexes. Les figures 2-d et 2-e visualisent la composante par plateau synthétisée et l'image complète restituée au décodeur ainsi que les débits moyens associés. On peut remarquer un certain nombre de défauts qui proviennent entre autres pour certains, de l'approximation systématique des contours par segments de droite et du découpage des polygones en polygones convexes. L'approximation linéaire de la composante par plateau dans les facettes polygonales amène également quelques défauts dans certaines régions. Au vu des résultats obtenus dans le codage d'un jeu d'images de nature assez différente, un certain nombre de remarques peuvent être faites concernant d'une part la qualité, d'autre part le coût de la représentation de l'information associée. On peut citer plus particulièrement les deux remarques suivantes :

- L'approximation de la géométrie des régions sous la forme de polygones conduit dans certains cas à une approximation soit trop imprécise de la position réelle des contours, soit très coûteuse en informations nécessaires à sa représentation.
- Le découpage des polygones en polygones convexes, nécessité par le choix fait d'un décodage très simple, introduit parfois une structure parasite dans la composante par plateau qui n'existe pas dans l'image à coder. Ce défaut est difficilement compensable par le codage de la composante de texture car le signal d'erreur de codage de la première composante et le signal de texture ont des propriétés alors très différentes.

Dans la partie suivante nous présentons donc des modifications au schéma de codage précédemment décrit qui permettent d'améliorer très sensiblement les défauts précités.

III - AMÉLIORATION DU CODAGE DE LA COMPOSANTE PAR PLATEAU

1 - Description des contours

Le choix de la description d'un contour a été revu. Le premier algorithme réalisé donnait une représentation de tous les contours par segments de droite. Les points essentiels, sont les points dont les coordonnées sont nécessairement données explicitement dans la description d'un contour élémentaire appelé arc. Le choix de ces



points était basé sur une étude trop locale du contour et était, dans certains cas, critiquable. Cette phase d'extraction des points essentiels n'est plus seulement locale mais également globale.

Un calcul d'efficacité du codage décide de la représentation de l'arc soit sous forme d'un code de FREEMAN (codage sans approximation), soit de façon approchée par segments de droite. Les critères sont non seulement spatiaux (erreur quadratique moyenne) mais également basés sur les valeurs du signal de luminance de part et d'autre du contour. L'approximation d'un arc se déroule en deux phases.

a) On définit l'angle $\theta(i, j)$ que fait un segment de droite orienté reliant n'importe quel couple de points (P_i, P_j) de l'arc avec un axe de référence. Pour chaque point P_i de l'arc, on définit deux angles moyens $\theta_1(i)$ et $\theta_2(i)$, de part et d'autre de ce point (figure 3) :

$$\theta_1(i) = \frac{1}{N_1} \sum_{k=i_1}^{i-1} \theta(k, i) \quad (N_1 = i - i_1)$$

$$\theta_2(i) = \frac{1}{N_2} \sum_{k=i+1}^{i_2} \theta(i, k) \quad (N_2 = i_2 - i)$$

L'angle $\alpha(i)$ est la différence entre ces deux angles

$$\alpha(i) = \theta_1(i) - \theta_2(i)$$

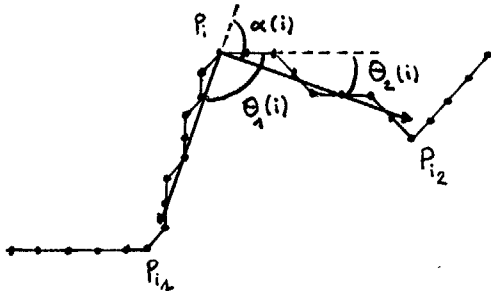


Figure 3 - Exemple d'arc et angles associés

Pour calculer les angles moyens, on s'éloigne progressivement du point P_i . A chaque nouveau point P_k , on calcule l'angle moyen provisoire et on le compare avec les angles au delà du point P_k . Si les différences entre ces angles sont supérieures à des seuils fixés, on arrête l'accumulation des angles et on définit ainsi les points P_{i1} et P_{i2} . Selon la valeur de l'angle α et des nombres N_1 et N_2 , une classe est affectée au point i . La première classe contient les points qui sont considérés comme essentiels, c'est-à-dire nécessairement conservés comme sommets (extrémités de segments) ; la seconde regroupe les points jugés comme sommets éventuels alors que la troisième rassemble les points dont l'angle α est sensiblement nul.

b) L'arc est divisé en sous-arcs dont les extrémités sont les points essentiels. Chaque sous-arc comporte éventuellement des sommets appartenant au second groupe ;

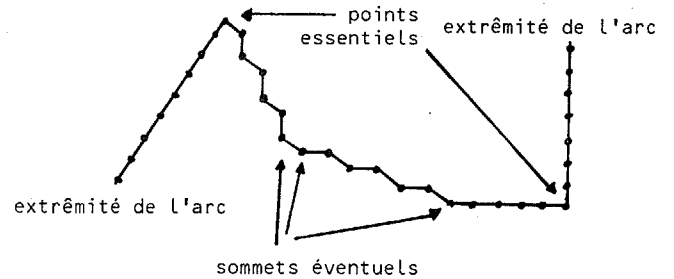


Figure 4 - Découpage d'un arc

Les sous-arcs sont traités séparément. Pour chacun d'eux, on trace une droite reliant ces deux extrémités.

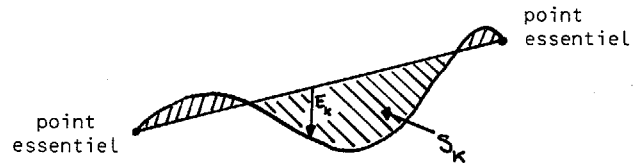


Figure 5 - Calculs d'erreurs

On calcule les écarts E_k créés par l'approximation linéaire du sous-arc et l'aire des surfaces S_k , qu'ils engendrent (voir figure 5). Les écarts sont modulés par le contraste local $C(k)$ du point P_k normalisé à 1 :

$$E'_k = E_k (\beta C(k) + (1 - \beta))$$

où β est un coefficient de pondération appartenant au domaine $[0, 1]$. On définit deux seuils fixés a priori μ_1 et μ_2 où :

le seuil μ_1 représente l'écart maximal E'_k toléré et le seuil μ_2 représente la moyenne maximale des écarts tolérée sur la surface S_k . Si les seuils sont dépassés, on découpe le sous-arc en introduisant un des sommets intermédiaires appartenant à la deuxième classe. On compare les C_N^1 solutions produites par l'introduction d'un des N sommets. Si la meilleure solution (erreurs minimales) ne satisfait pas les critères (seuils dépassés) on introduit alors deux sommets intermédiaires. De façon itérative, on étudie les C_N^k possibilités de découpage du sous-arc avec les k sommets choisis parmi N , jusqu'à ce que les critères d'erreurs soient satisfaits. On compare le coût du codage de l'arc original (codage de FREEMAN) et le coût du codage de l'arc représenté par des segments de droite, les coordonnées de chaque sommet étant codées de façon différentielle :

- Coût du codage de FREEMAN = NB POINT * 3 bits
où NB POINT est le Nombre de points de l'arc
- Coût du codage par segments de droite ≈
NB SOMMET * 12 bits
où NB SOMMET est le Nombre de sommets

2 - Description du signal de valeur moyenne

La méthode précédemment mise en oeuvre consiste à représenter les contours par un ensemble de polygones convexes (figure 2-c). En estimant la luminance sur chacun des sommets, on peut reconstruire la composante par plateau à l'intérieur du polygone en faisant varier linéairement la luminance sur les arêtes entre deux sommets, puis linéairement sur les lignes entre deux arêtes. Cette méthode employée en synthèse d'images pour générer des facettes planes n'est pas bien adaptée à notre problème. Après segmentation, nous pouvons obtenir de très grandes régions dont la luminance varie lentement mais dans de grandes proportions et souvent de façon assez complexe et qui ne correspond pas à l'hypothèse arbitraire d'évolution linéaire horizontale ou le long des arêtes de la valeur moyenne locale du signal de luminance. Ceci entraîne un défaut pouvant être très gênant (figure 2-d) encore accentué par le découpage en polygones convexes. Dans le cas de l'utilisation maintenant envisagée d'un décodeur à fonctions plus complexes, nous pouvons corriger ce défaut. Le découpage en polygones convexes a été abandonné ce qui, de plus, améliore le taux de compression. Un algorithme simple permet d'explorer chaque région de l'image quel qu'elle soit sa forme. Toutefois il est nécessaire de marquer chaque point exploré, ce qui oblige à disposer au décodeur d'une mémoire image d'au moins 2 bits par point.

La nouvelle méthode de représentation de la composante par plateau consiste en un découpage de chaque région en sous-régions dans lesquelles la valeur moyenne locale du signal de luminance peut être convenablement représentée par un plan. Chaque plan est défini par trois paramètres qui sont les luminances de trois points intérieurs à la région espacés et disposés de façon à définir la luminance de chaque point de la région avec le plus de précision possible. Un calcul simple permet de déterminer les coordonnées de ces trois points de façon identique au codeur et au décodeur. La réduction de l'information nécessaire à la description d'un plateau est compensée par l'obligation d'augmenter le nombre de régions de l'image de façon parfois comparable à celle de la décomposition de polygones en polygones convexes mais le gain procuré se situe surtout

dans l'amélioration de la qualité visuelle de l'image décodée.

On cherche à diviser une région en sous-régions de points connexes dans lesquelles pour tous les points le signal de luminance moyen peut être donnée par l'équation d'un plan (P) :

$$L = a X + b Y + C$$

Le plan (P) choisi pour représenter les variations du signal de luminance d'une sous-région est celui qui minimise l'erreur quadratique moyenne. La première étape (étape préliminaire) est un lissage du signal de luminance à l'intérieur de la région à l'aide d'un opérateur médian à support 2-D dont tous les éléments sont toujours intérieurs à la région considérée. La formation d'une sous-région débute par la recherche d'une base de points connexes dont le signal de luminance vérifie l'équation d'un plan avec une erreur quadratique moyenne inférieure à un seuil fixé. Dans l'étape n^oφ, le nombre de points de la sous-région est N_φ. Les étapes suivantes sont des extensions successives de la sous-région. A chaque étape k, un nombre M_k de points de la région parmi ceux qui bordent la frontière de la sous-région en cours de construction est inclus selon les critères suivants :

- Soit L_i la luminance du point i testé. En utilisant les valeurs du signal de luminance des points voisins qui ne sont pas encore inclus, on estime, à l'aide d'un opérateur moyen, une pente horizontale a_i et une pente verticale b_i. On définit :

$$\Delta L = |L - L_i|, \quad \Delta a = |a - a_i|, \quad \Delta b = |b - b_i|.$$

Deux seuils η₁ et η₂ (η₁ < η₂) sont fixés pour ΔL et un seuil η₃ pour Δa et Δb.

si ΔL < η₁ , le point est inclus

si ΔL > η₂ , le point est exclu

sinon, si Δa < η₃ et Δb < η₃ , le point est inclus
sinon, il est exclu.

On obtient donc, à chaque étape k, un nombre de points

$$N_k = N_{k-1} + M_k.$$

Les coefficients a, b, c du plan P sont réactualisés à chaque étape k₁ où l'accumulation :

$$\left(\sum_{k=k_0+1}^{k_1} M_k \right)$$

des points inclus à la sous-région depuis la précédente réactualisation à l'étape k₀ dépasse un certain pourcentage du nombre N_{k₀}.



Quand plus aucune extension n'est possible, un contour est créé sur la frontière de la sous-région qui vient d'être créée.

IV - RESULTATS ET CONCLUSION

Nous présentons sur les figures 2a à 2-h les résultats obtenus sur l'image Ballon avec l'ancien schéma de codage, puis ceux obtenus avec les modifications que nous venons de préciser. Avec la première méthode, l'image de la composante par plateau (fig. 2-d) présente deux types de défauts. L'un lié à l'approximation linéaire des arcs, assez "large" pour assurer une réduction convenable du débit, et qui entraîne une mauvaise restitution des formes arrondies (ballon). L'autre lié au découpage en polygones convexes, qui introduit des effets de trainée à l'intérieur de certaines régions. Ces erreurs provoquent de plus une très mauvaise représentation des textures comme on le voit sur l'image codée (fig. 2-e). Les trois dernières images présentent les résultats obtenus avec la méthode modifiée de codage de la composante par plateau. Le découpage des régions en sous-régions (fig. 2-f) double, en moyenne, leur nombre. La différence (fig. 2-h) entre l'image d'origine et la nouvelle représentation de la composante par plateau (fig. 2-g) montre que celle-ci a été correctement codée. A l'intérieur des régions, seul le signal de texture reste à coder. Les erreurs se situent surtout dans la représentation de la luminance sur les contours. Ceci n'est pas très gênant car la précision demandée au codage est nettement moindre sur les contours en raison de l'effet de masquage qu'ils engendrent. Sur l'image ballon présentée, l'information nécessaire à la nouvelle représentation de la composante par plateau reste identique à l'ancienne et correspond à 0.48 eb/point. Il faut noter que ces valeurs sont des limites maximales dans le sens

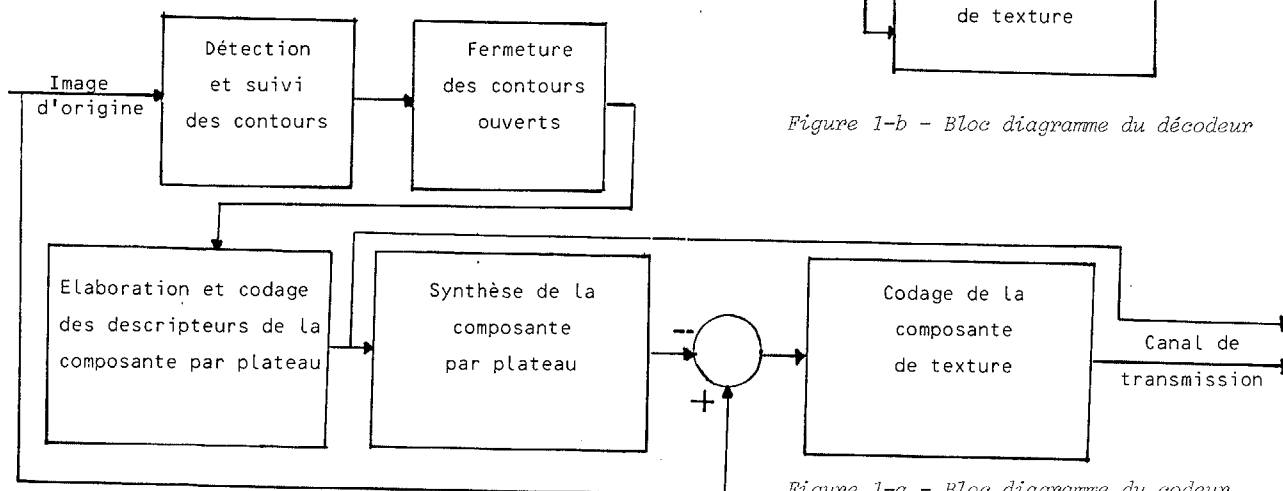


Figure 1-a - Bloc diagramme du codeur

qu'elles ont été calculées sans aucune optimisation du codage des paramètres de la représentation. Un gain substantiel doit être encore réalisé à ce niveau. Des gains supplémentaires doivent être obtenus d'une part dans le codage des très petites régions introduites parfois par le processus de segmentation de l'image, d'autre part dans la représentation et le codage de la texture.

BIBLIOGRAPHIE

- [1] W.F. SCHREIBER, C.F. KNAPP, N.D. KAY : "Synthetic highs : an experiment TV band width reduction system", J.S.Motion Pic. Tel. Eng., Vol. 68, Aug 1959, pp. 525-537.
- [2] J.K. YAN, D.J. SAKRISON : "Encoding of images based on a two-component source-model". IEEE Trans. on Com., Vol. 25, n°11, pp. 1315-1322.
- [3] L. FEITS, L. CHARIGLIONE : "Image coding by means of a two-component source-coding scheme", CSELT R.T., Vol. 8, n°2, pp. 83-92.
- [4] M. KOCHER : "Codage d'images à haute compression basé sur un modèle contour-texture". Thèse de Doctorat E.P.F.L., Lausanne, 1983.
- [5] J. RODRIGUEZ, D. BARBA : "Codage d'images par une méthode à deux composantes". 1er Colloque Image GRETSI-CESTA, Biarritz, Mai 1984.

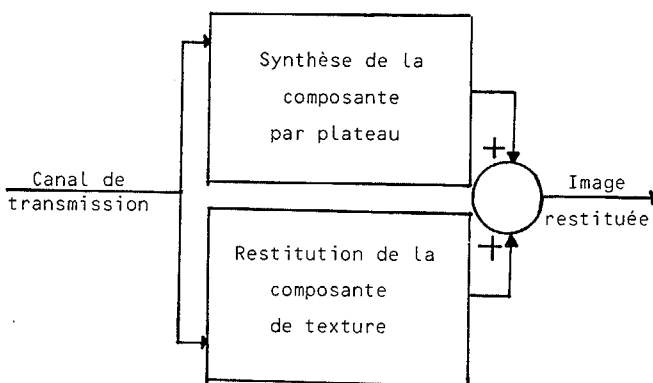


Figure 1-b - Bloc diagramme du décodeur

AMÉLIORATION EN CODAGE D'IMAGES D'UNE MÉTHODE A DEUX COMPOSANTES

Figure 2-a -
Image
d'origine

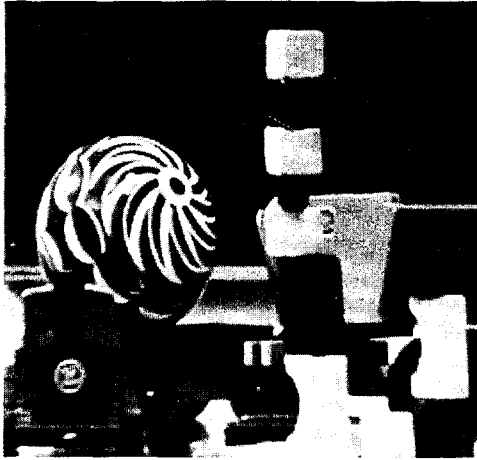


Figure 2-b -
Image
des contours
fermés

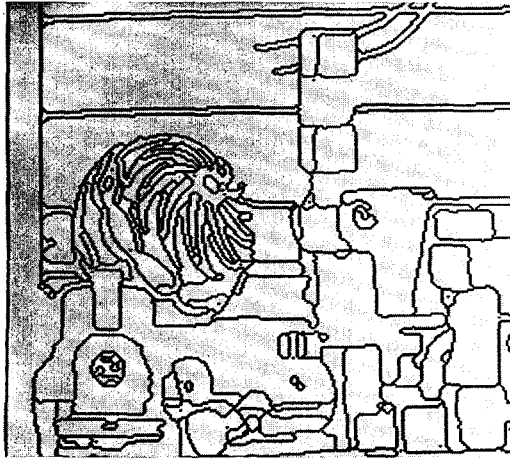


Figure 2-c -
Image
des polygones
convexes

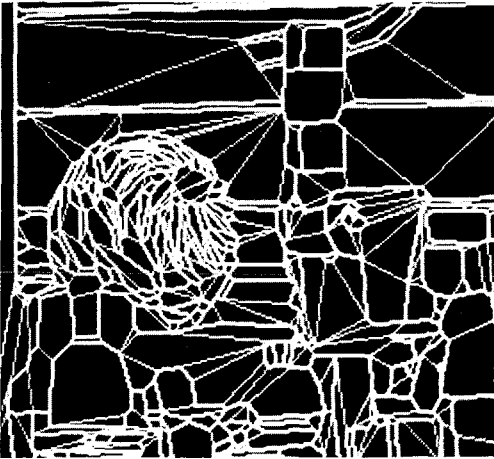


Figure 2-d -
Image
des plateaux
D=0.48eb/point

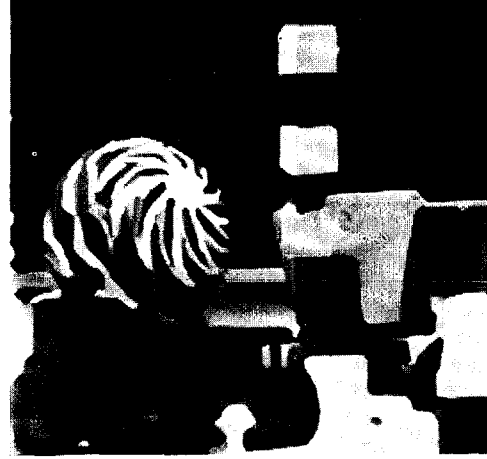


Figure 2-e -
Image
restituée
D=1.5eb/point

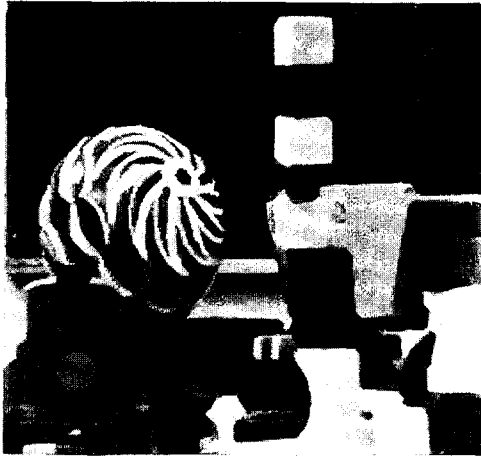


Figure 2-f -
Nouvelle image
des contours
fermés

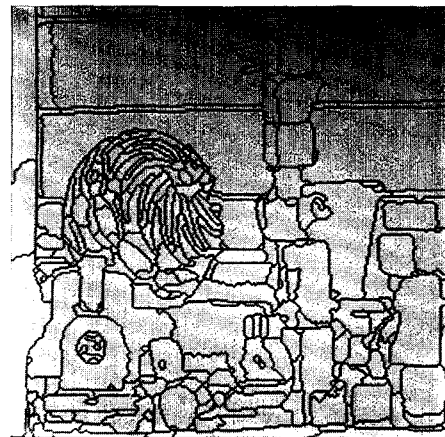


Figure 2-g -
Nouvelle image
des plateaux
D=0.48eb/point

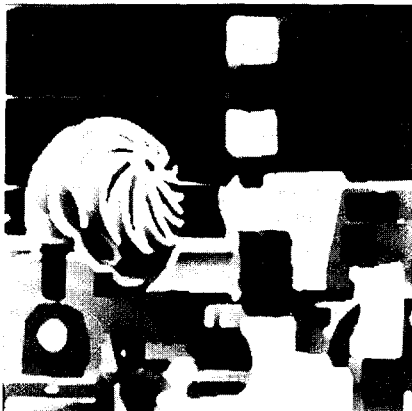


Figure 2-h -
Nouvelle image
de texture
avant codage

