

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION

Adaptive multimode coding scheme for broadcast television sequences .

PHILIPPE ROBERT CLAUDE LABIT

IRISA/CENTRE INRIA DE RENNES: CAMPUS DE BEAULIEU, AV. DU GÉN. LECLERC 35042 RENNES CEDEX FRANCE

RESUME

Afin de réduire le débit binaire d'informations à transmettre pour le codage des séquences d'images de Télévision Haute Qualité ($f_s=13.5\text{Mhz}$), cette étude présente un schéma de codage MICDA, à débit variable, et à stratégies de prédiction multiples.

Les prédicteurs sont choisis intraimages, ou inter-images avec ou sans compensation de mouvement. Ces trois types de prédicteurs conduisent :

- . à une optimisation locale en chaque pixel de l'erreur de prédiction qui n'est transmise que pour les points non compensables en mouvement;
- . à l'envoi d'informations d'adresses des différentes zones de prédiction qui représentent bien souvent une part importante du débit binaire total de transmission.

Des techniques optimisées d'adressage ont déjà été proposées dans le domaine du codage par plages, du codage d'images Télécopie..etc. Celles-ci méritent d'être étendues dans le cadre général d'un codeur multimodes.

Pour ce faire, cette étude propose deux axes de recherche pour la réduction du débit d'adresses. La première optimisation porte sur le nombre total d'adresses qui peut être réduit par des algorithmes de segmentation d'images et de regroupements de paquets. La deuxième a été menée grâce à une étude comparative sur les algorithmes d'adressage différentiel monodimensionnel ou bidimensionnel. La notion de prédiction d'adresses permet de réduire efficacement la taille des adresses différentielles à coder.

L'ensemble de ces techniques de prédiction et d'adressage a été testé sur deux séquences d'images (à caractéristiques différentes) et autorise pour chacune d'elles des gains significatifs en compression de l'information à transmettre.

MOTS CLES : Codage, séquences, prédiction intra/inter, compensation de mouvement, adressage différentiel.

SUMMARY

This paper presents an ADPCM multimode coder scheme reducing the transmission bit rate of high quality ($f_s=13.5\text{Mhz}$) television picture sequences.

Set of predictors have been used in both intra- and interframe domains. Motion compensation has been performed to improve the coding of moving areas. This multimode structure authorizes :

- . a local optimization, at each pixel, of the prediction error which is only transmitted for motion-uncompensated pixels.
- . an addressing bit rate which is often quite important into the total bit rate and has to be locally optimized.

Some techniques have been already used for block coding, telecopy... and could be extended to any multimode coding scheme.

We present two studies reducing the addressing bit rate. The first step is to reduce the total number of addresses. It can be performed using segmentation, filtering and growing algorithms applied to the image clusters. Then some differential (1-dimensional or 2-dimensional) addressing coding can be helpful to optimally code these boundaries.

All of these coding algorithms have been tested using sequences with different spatial and temporal characteristics and enable quite significant decrease in terms of information amount to be transmitted.

KEYWORDS :Image sequence coding, intraframe, interframe prediction, motion compensation, differential addressing



CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION
Adaptive multimode coding scheme for broadcast television sequences .
PHILIPPE ROBERT CLAUDE LABIT

I . INTRODUCTION

Les études que nous avons menées ([1], [2], [3]) dans le domaine du codage de séquences d'images de télévision visent de manière prospective à cerner, analyser et utiliser les redondances quelles que soient leurs natures (spatiales, temporelles, spatio-temporelles) présentes dans une séquence d'images. Pour ce faire, il convient de considérer une séquence d'images comme un signal spatiotemporel $I(x,y,t)$ où les corrélations sont à rechercher simultanément dans les trois dimensions:

- i) l'utilisation des corrélations spatiales (en x et y) autorise des prédictions intraimages bidimensionnelles efficaces; cette possibilité doit être utilisée dans les zones spatialement homogènes;
- ii) l'analyse des corrélations temporelles permet de segmenter l'image en zones fixes et zones en mouvement. Deux types d'algorithmique à complexité différente peuvent être introduits à ce niveau: une simple détection de mouvement permet de supprimer la transmission des informations d'image contenues dans les zones fixes; pour les zones en mouvement, une étape supplémentaire — celle d'une estimation quantitative des mouvements des points "changés" temporellement — convient d'être introduite afin de prédire par compensation de mouvement les points d'images des zones en mouvement.

Les algorithmes ont été testés sur deux séquences d'images à caractéristiques spatiales et temporelles différentes (Fig 1 et 2)

II APPROCHE ET DEFINITION DE LA STRUCTURE DU CODEUR

II.1 : Approche adoptée

La nature prospective de l'étude présentée ici nous a imposé la volonté d'extraire et d'utiliser au maximum

les redondances présentes dans l'image et nous a conduit à opter pour un système de codage à débit variable où aucune information redondante du signal d'image n'est transmise. Le caractère variable du débit est créé par :

- i) un envoi non systématique d'informations d'erreurs; une erreur de prédiction n'est transmise que pour les points non compensables en mouvement;
- ii) un codage statistique (code de Huffman par ex.) des informations transmises.

Dans ce cadre algorithmique, deux volumes principaux d'informations à transmettre sont à minimiser afin de réduire le débit total de transmission :

- 1) la minimisation du débit d'erreurs consiste en la recherche de prédicteurs optimisés. De nombreuses recherches portent sur ce sujet ([4], [5], [6]). Nous présentons ici un schéma de prédiction adaptatif multimodes à prédicteurs tantôt intraimages, tantôt interimages avec ou sans compensation de mouvement.

- 2) un codage non systématique induit l'introduction d'adresses de zones d'image à stratégies de prédiction distinctes. Ce volume d'informations d'adressage représente bien souvent une part importante du débit total binaire (jusqu'à 60% de ce débit [7]). Peu d'études en codage multimodes ont cherché à minimiser ce débit d'adresses par des techniques prédictives de codage d'adresses. De telles techniques optimisées d'adressage ont été proposées dans le domaine du codage par plages, du codage d'images Télécopie...Celles-ci méritent d'être étendues aux structures multimodes. De telles techniques d'adressage suppose une protection renforcée des codes-canal choisis vis-à-vis des erreurs de transmission.

Figure 1 : 4 images consécutives de la séquence "JUDITH" de visiophone ($f_{\text{éch}}=2\text{Mhz}$, mouvements jusqu'à 2pixels/fr)

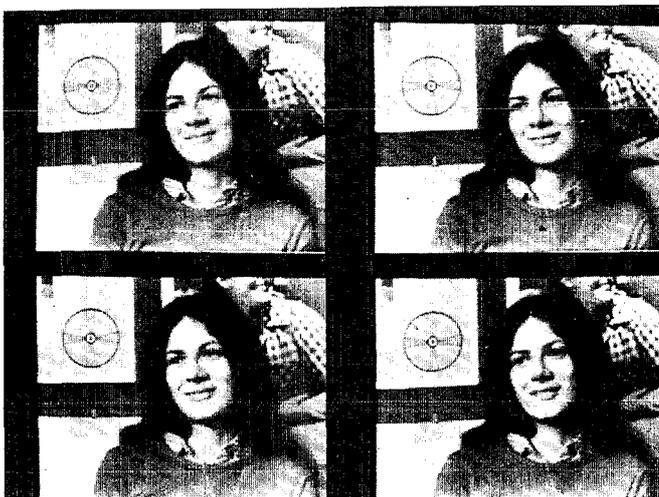
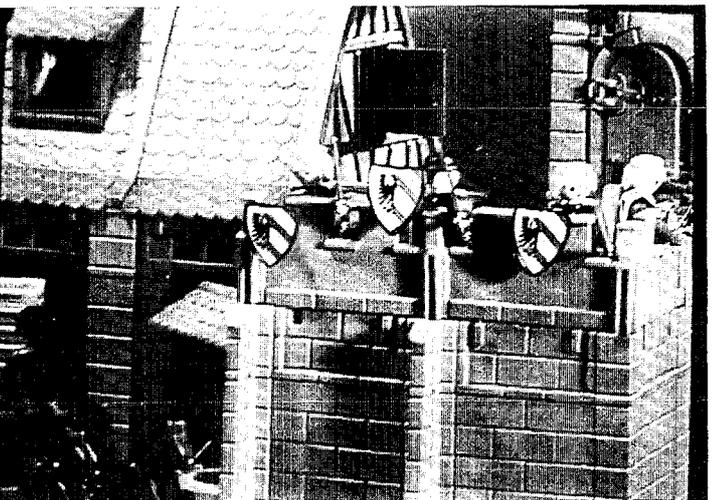


Figure 2 : Image source de la séquence "CHATEAU" ; ($f_{\text{éch}}=12\text{Mhz}$, mouvements jusqu'à 6 pixels/frame)





CODAGE ADAPTATIF MULTIMODES POUR LES SEQUENCES D'IMAGES DE TELEVISION
Adaptive multimode coding scheme for broadcast television sequences .

PHILIPPE ROBERT CLAUDE LABIT

III . SEGMENTATION ET CLASSIFICATION DE L'IMAGE

III . 1 : Algorithme de segmentation hiérarchique.

* La première méthode de segmentation proposée est effectuée à partir des deux données temporelles suivantes disponibles en chaque pixel :

- la différence interimage ("Frame difference = "FD")

$$FD(x,t) = I(x,t) - \hat{I}(x,t-1)$$

- la différence interimage avec compensation de mouvement ("Displaced Frame Difference = "DFD")

$$DFD(x,t) = I(x,t) - \hat{I}(x-D(x,t), t-1)$$

* L'algorithme de segmentation est décrite en Figure 5. Il consiste à détecter tout d'abord les 3 points fixes, puis les points compensés et enfin les points restants, critiques pour la prédiction, qui sont principalement situés sur les zones contours. Les résultats de cette segmentation sont visualisés en Annex Figure A . Le choix des seuils de segmentation λ_1 et λ_2 doit être effectué en respect de critères objectifs (réduction du bruit temporel, segmentation non redondante) mais également de critères subjectifs (dégradation visuelle maximale autorisée, haute qualité après reconstruction...) Le seuil λ_1 discrimine les zones fixes des zones en mouvements mais aussi les "vrais" changements temporels du bruit temporel de source ou de quantification. Une valeur élevée de λ_1 permet une élimination de celui-ci mais induit des effets dynamiques "de vitre sale" fortement nuisibles. Le seuil λ_2 est à considérer comme la valeur maximale de l'erreur temporelle autorisée dans les zones à mouvement compensé.

III . 2 : Segmentation avec filtrage spatial.

III.2.a : Certains points d'images, au vu des deux paramètres temporels FD et DFD, peuvent être déclarés à la

fois fixes et compensés en mouvement. Cette remarque préliminaire amène tout naturellement l'introduction d'une classe supplémentaire dans le processus de segmentation de l'image. L'ajout de cette nouvelle classe "temporaire" permettra de tenir compte de l'ambiguïté inhérent à certains pixels et de partitionner l'image en quatre classes disjointes dont les définitions (avec notations de la Fig.5 sont :

$$\begin{aligned} \text{classe } T'_4 &= T_1 \cap T_2 \\ \text{classe } T'_1 &= T_1 - T_2 \\ \text{classe } T'_2 &= T_2 - T_1 \\ \text{classe } T'_3 &= T_3 \end{aligned}$$

III.2.b : Filtrage spatial sur la classe T'_4 . Plusieurs méthodes de filtrage sur les pixels de type T'_4 ((3)) ont été testées. La plus efficace semble être un schéma d'hystérésis spatial ou politique conservatrice d'affectation de type entre le point précédent et le point courant :

si X_n (point à classer) $\in T'_4$
alors si X_{n-1} (point précédent) $\in T'_1$ alors $X_n \in T'_1$
 si $X_{n-1} \in T'_2$ alors $X_n \in T'_2$
 si $X_{n-1} \in T'_3$ alors $X_n \in T'_2$

III.2.c : Résultats et améliorations. Les résultats de segmentation sont visualisés en Annexe. Figure B montre la partition en 4 classes disjointes avant filtrage sur une des séquences. Figures C et D montrent la partition finale en 3 classes après filtrage par hystérésis pour les deux types de séquences utilisées.

De nombreuses améliorations en termes de segmentation optimale pour une qualité d'image accrue et une minimisation du nombre d'adresses de paquets à transmettre. D'une part des algorithmes d'hystérésis à mémoire permettent de parfaire le filtrage. D'autre part il est nécessaire d'introduire des règles de nettoyage supprimant les paquets de trop faible taille. Ces nettoyages peuvent être effec-

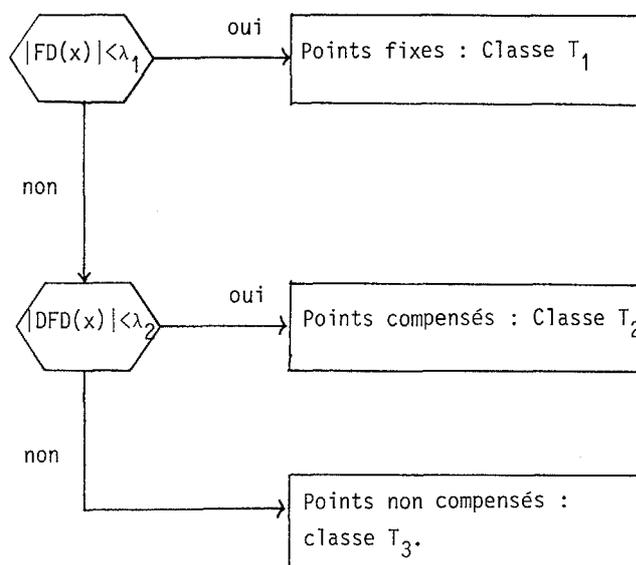


FIGURE 5

Algorithme de segmentation hiérarchique en 3 classes

CODAGE ADAPTATIF MULTIMODES POUR DES SÉQUENCES D'IMAGES DE TÉLÉVISION
 Adaptive multimode coding scheme for broadcast television sequences .
 PHILIPPE ROBERT CLAUDE LABIT

tués sans dégradations pour l'image reconstruite (remplacement d'adresses par des erreurs de prédiction à transmettre) ou avec dégradations (pontage de paquets) mais non perceptibles par l'oeil.

III . 3 : Résultats et conclusions.

Nous présentons en Figure 6 les résultats des différents algorithmes sur les deux séquences avec pour chacune d'entre elles, deux jeux de seuils λ_1 et λ_2 . Il apparait nettement que les techniques de filtrage et nettoyage permettent de fortes réductions du nombre d'adresses à transmettre et une meilleure répartition des points dans les trois classes initiales en privilégiant celle des points compensés en mouvement (minimisation des effets de "vitre sale" dynamiques).

SEQUENCES	SEUILS	ALGORITHME DE SEGMENTATION	NOMBRE D'ADRESSES PAR LIGNE	POURCENTAGE DES POINTS
"JUDITH"	λ_1 λ_2	Segmentation hiérarchique	19.5	88 9.5 2.5
		Segmentation avec filtrage	8.7	51.2 46.3 2.5
	4 7	Segmentation avec filtrage et nettoyage	5.5	51 46 3
		Segmentation hiérarchique	27	24 12 4
		Segmentation avec filtrage	14	50.5 45.5 4
		Segmentation avec filtrage et nettoyage	8	50 45 5
"CHATEAU"	7 12	Segmentation hiérarchique	190	47 36 17
		Segmentation avec filtrage	103.6	13 70 17
	5 9	Segmentation avec filtrage et nettoyage	52.5	11 68 21
		Segmentation hiérarchique	229	37 40 23
		Segmentation avec filtrage	144	13 64 23
		Segmentation avec filtrage et nettoyage	68	10 61 29

FIGURE 6 : Segmentation de l'image

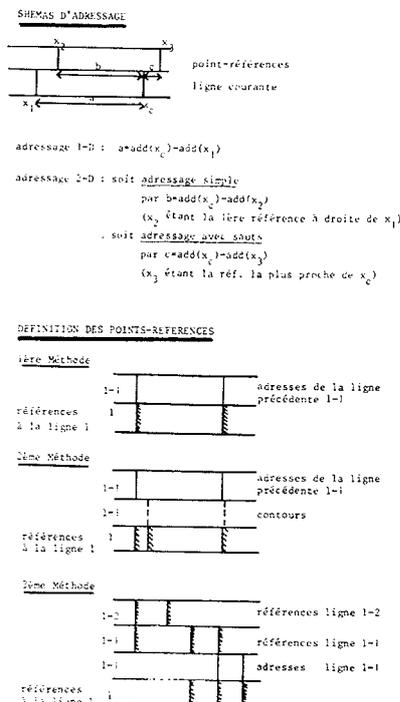


FIGURE 7 : Adressage bidimensionnel

IV ALGORITHMES D'ADRESSAGE :

Pour réduire le débit binaire résultant du codage des adresses des zones précédemment segmentées, il convient également de prendre en compte les corrélations existantes entre adresses à coder par un adressage différentiel monodimensionnel (sur une ligne d'image) ou bidimensionnel (sur plusieurs lignes) . Certaines techniques d'adressage de ce type ont déjà été développées dans d'autres domaines (télécopie, fac-similé..) et méritent d'être rappelées ([8]) :

IV.1 : Adressage différentiel bidimensionnel

Il s'agit d'adresser les zones différenciellement sur la ligne courante et également en comparaison avec des références de zones de la ligne précédente:

..La méthode EDIC vise à détecter trois types d'événements: i) une apparition d'adresse pour laquelle aucune adresse ne correspond à la ligne précédente

ii) une disparition d'adresse par rapport à une adresse existante à la ligne précédente

iii) un déplacement borné d'adresses entre les deux lignes. Chaque type sera caractérisé et codé différemment.

..La méthode RAC choisit en chaque adresse à coder $add(x_c)$ la stratégie de codage minimale en distance entre l'adressage différentiel 1-D et 2-D. Un indicateur de permutation de mode d'adressage est alors transmis.

..La méthode READ combinant les deux méthodes précédentes apparait comme un système efficace d'adressage pour le codage des documents.

Nous avons repris et généralisé ces méthodes d'adressage bidimensionnel dans le cadre d'un schéma de codage multimodes.

IV.2 : Application au codeur multimodes :

Avec les notations de la Figure 7, nous distinguerons l'adressage différentiel 1-D qui transmet

$$a_1 = add(x_c) - add(x_1), \quad x_1 \text{ étant la dernière adresse transmise}$$

et l'adressage différentiel 2-D qui transmet

$$a_2 = add(x_c) - add(réf), \quad réf \text{ étant l'adresse-pixel de référence par rapport auquel l'adresse courante } add(x_c) \text{ va être calculée.}$$

De nombreuses simulations [3] ont été faites pour tester les diverses définitions possibles de points-références. Trois schémas principaux (à complexité croissante) peuvent être retenus:

i) les point-références d'adresses pour la ligne courante sont les adresses de paquets de la ligne précédente

ii) A ces références peuvent être adjointes les adresses des points contours détectés à la ligne précédente. Cette amélioration est introduite pour tenir compte de



CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION
Adaptive multimode coding scheme for broadcast television sequences .
PHILIPPE ROBERT CLAUDE LABIT

la forte corrélation existante parfois entre les points de transition de zones et les points contours (voir la Fig. A en Annexe). La prise en compte de l'orientation des contours dans cette projection n'apporte qu'un faible gain eu égard à la complexité algorithmique supplémentaire

iii) Une définition récursive de ces points-références par extension de la mémoire des points de références à la ligne ante-pénultième est décrite en Fig.7. Cet algorithme facilement implantable permet des adressages efficaces même en présence de trous dans les corrélations ligne-à-ligne des adresses de zones.

L'adressage différentiel bidimensionnel pourra être défini selon deux types :

1) l'adressage par $b = \text{add}(x_c) - \text{add}(x_2)$, simple en réa-lisation, est sous-optimale du point de vue distance référence-adresse courante.

2) l'adressage avec sauts de références permet de minimiser les adresses différentielles mais ajoute des informations annexes (nombre de sauts de références) à transmettre

	TYPE DE TRANSITION	MODE D'ADRESSAGE	ENTROPIE	REPARTITION EN Z	
				1-D	2-D
JUDITH	transition $T_1 \rightarrow T_2$ ou $T_2 \rightarrow T_1$	monodimensionnel	5,7	100	
		bidimensionnel simple	5,0	48 52	
		bidimensionnel avec sauts	5,0		
	transition $X \rightarrow T_3$	monodimensionnel	5,4	100	
		bidimensionnel simple	3,8	32 68	
		bidimensionnel avec sauts	4,4		
CHATEAU	transition $T_1 \rightarrow T_2$ ou $T_2 \rightarrow T_1$	monodimensionnel	4,8	100	
		bidimensionnel simple	4,7	86 14	
		bidimensionnel avec sauts	4,7		
	transition $X \rightarrow T_3$	monodimensionnel	5,0	100	
		bidimensionnel simple	4,4	36 64	
		bidimensionnel avec sauts	4,3		

IV.3 : Résultats et Conclusion :

Les résultats pour les deux séquences (Figure 8) sont exprimés en entropie théorique calculée sur l'histogramme d'adresses et en pourcentage de répartition entre l'adressage mono- et bidimensionnel. La nature des transitions de zones est distinguée également au vu de la forte disparité de corrélation ligne-à-ligne entre les différents types de transitions (voir Annexe).

Il apparaît ainsi que l'introduction d'un codage d'adresses bidimensionnel n'est fortement efficace que dans les zones fortement corrélées verticalement, comme se trouvent être les zones de points non compensés de type T_3 . Dans le cas de séquence à mouvement peu complexe, le gain de l'adressage différentiel 2-D est toutefois sensible sur tous les types de transitions.

CONCLUSION :

L'ensemble des techniques algorithmiques présentées ici ont cherché, tant au niveau prédiction de l'intensité lumineuse qu'au niveau adressage de zones, à utiliser les redondances (corrélations spatiales et temporelles) pour réduire le débit binaire d'informations à transmettre. De telles techniques permettent d'espérer atteindre des débits faibles (avoisinant les 2bit/pixel) pour des séquences de Télévision complexes et de haute qualité.

FIGURE 8

REFERENCES :

- [1] : C. LABIT : Estimation de mouvement dans une séquence d'images de télévision
Thèse de Docteur-Ingénieur Rapport IRISA n°187 Fév. 82
- [2] : C. LABIT, A. BENVENISTE : Motion estimation in T.V sequences
NATO ASI Series (édité par T.S HUANG): Image Sequence Processing and dynamic scene analysis (83)
- [3] : P. ROBERT : Contribution à la définition d'un schéma de codeur multimodes avec compensation de mouvement
Thèse Docteur-Ingénieur IRISA Nov. 83
- [4] : A.N NETRAVALI, J.O LIMB : Picture Coding : a review
Proceedings of the IEEE Vol 68 n°3 Mars 80
- [5] : A.K JAIN : Image data compression : a review
Proceedings of the IEEE Vol 69 n°3 Mars 81
- [6] : F. KRETZ : Edges in visual scenes and sequences. Application to filtering, sampling and ADPCM coding.
Même référence que [2]
- [7] : A.N NETRAVALI, J.O ROBBINS : Motion compensated Television Coding
BSTJ Vol 58 n°3 Mars 79
- [8] : Y. YASUDA : Overview of digital Facsimile Coding Techniques
Proceedings of IEEE Vol 68 n°7 July 80

CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION
 Adaptive multimode coding scheme for broadcast television sequences .
 PHILIPPE ROBERT CLAUDE LABIT

ANNEXE (L'ensemble de ces résultats graphiques sont
 extrait de la référence 3)

FIGURE A : Partition de l'Image "JUDITH" en 3 classes suivant l'algorithme de segmentation hiérarchique
 Pour comparaison, est visualisée également l'image contours

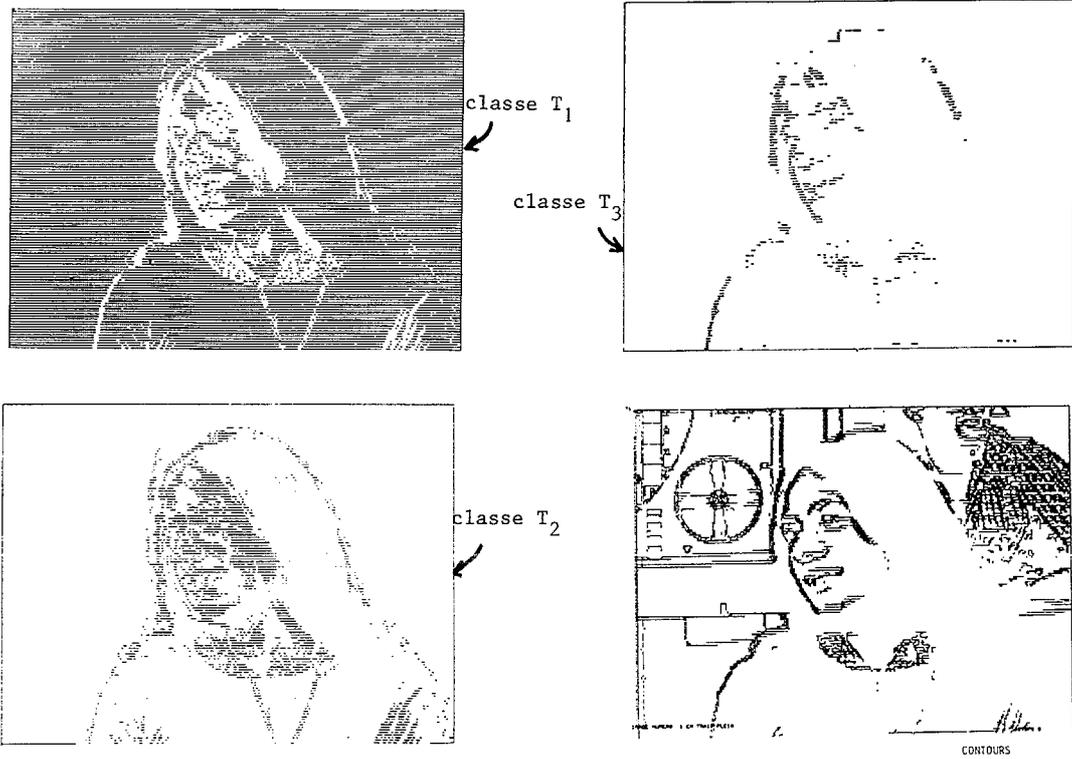
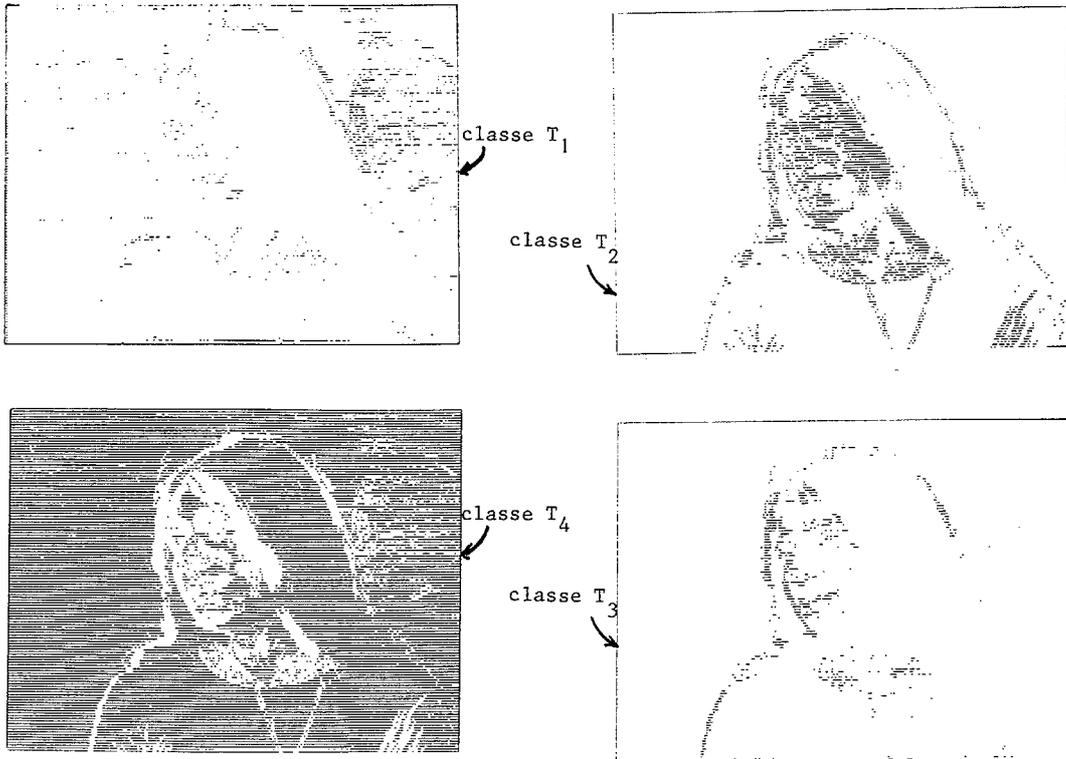


FIGURE B : Segmentation de l'Image "JUDITH" en 4 classes avant filtrage de la classe T₄





CODAGE ADAPTATIF MULTIMODES POUR LES SÉQUENCES D'IMAGES DE TÉLÉVISION .
 Adaptive multimode coding scheme for broadcast television sequences .
 PHILIPPE ROBERT CLAUDE LABIT

FIGURE C : Partition de l'Image "JUDITH" en trois classes par segmentation avec filtrage.

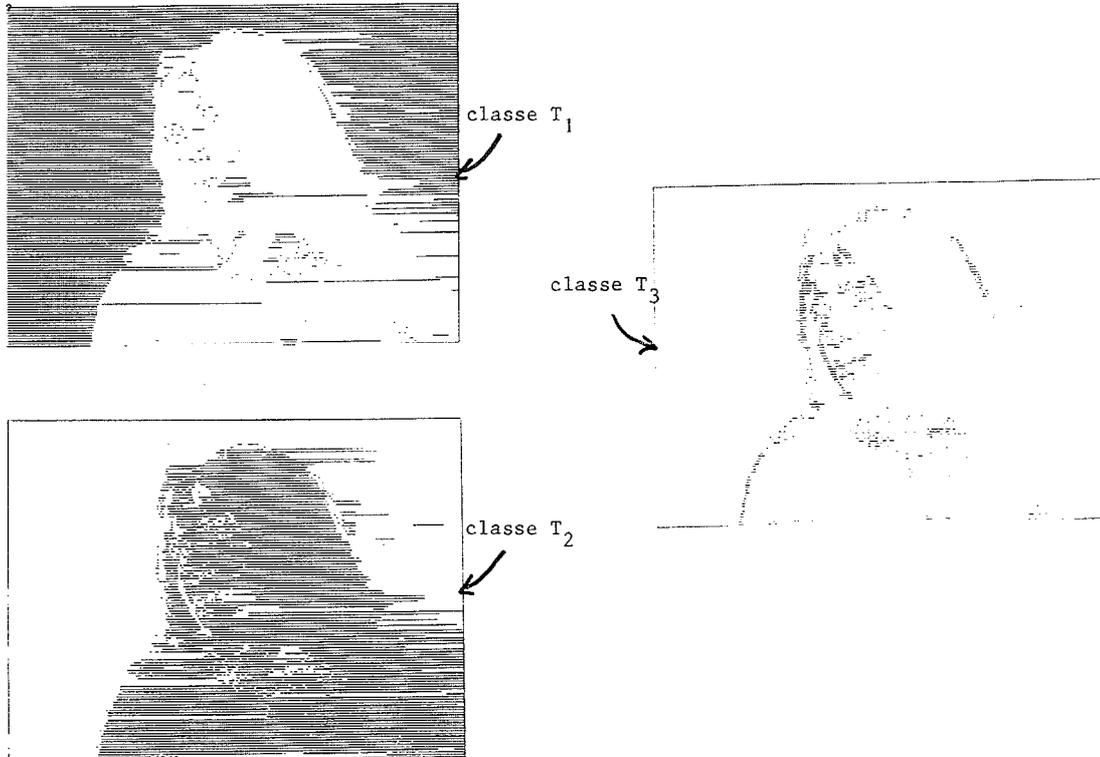


FIGURE D : Même résultats sur l'Image "CHATEAU"

