

# COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75



CARACTERISATION DES DEFAUTS APPORTES PAR  
LES CODAGES NUMERIQUES DE SIGNAUX D'IMAGE

Francis KRETZ et Jean-Louis BOUDEVILLE

C. C. E. T. T., Département NIM, 2 rue de la Mabilais, B. P. 1266, 35013 RENNES

## RESUME

Les techniques de transmission numérique des signaux d'image permettent d'assurer quasiment une qualité constante quelle que soit la longueur de la transmission. Pour des raisons économiques, il est nécessaire de réduire le débit d'information: on utilise pour cela diverses techniques de compression d'information. Les systèmes analogiques usuels utilisent certaines des propriétés de la vision pour réduire au mieux la bande passante à transmettre. Les divers défauts qui sont intervenus dans le choix des normes seront rappelés car les systèmes analogiques servent de support pour les codages numériques. Ceux-ci utilisent conjointement des modèles stochastiques et des propriétés de la vision. Ils se décomposent en une partie "numérisation" (échantillonnage, quantification), une partie codage et une partie transmission. Les divers défauts seront caractérisés en prenant comme exemples les systèmes existants. L'accent sera mis sur leur visibilité et sur les problèmes d'optimisation de ces systèmes.

## SUMMARY

The digital transmission techniques of video signals allow a nearly constant quality whatever the transmission distance is. For economical reasons, it is necessary to reduce the information rate: various techniques of information compression are used. The analog systems use some properties of the vision in order to reduce the bandwidth. The impairments involved in the normalization will be reviewed because the digital systems code the analog signals. These systems make use of both stochastic models and vision properties. They can be decomposed in a digitization part (sampling and quantization), a coding part and a transmission one. The various impairments will be characterised by using the existing systems as examples. Their visibility and the problems of optimization will be emphasized.



CARACTERISATION DES DEFAUTS APPORTES PAR  
LES CODAGES NUMERIQUES DE SIGNAUX D'IMAGE

I. INTRODUCTION

Le but de cet exposé est de présenter les types de défauts qui interviennent quand on numérise des signaux d'image et surtout quand on les code. Les systèmes étudiés doivent évidemment être définis et optimisés en fonction de ces défauts. Il ne sera envisagé ici que les systèmes temps réel (télévision ou visiophone) qui, même si le signal original est fixe (scène fixe analysée par la caméra), introduisent par le bruit et le balayage des défauts temporels en général beaucoup plus gênants que les défauts dus au même codage figés sur une photographie.

II. DESCRIPTION DES SYSTEMES ANALOGIQUES

La scène analysée peut être décrite par une fonction à quatre variables  $f(x, y, t, \lambda)$  - où  $f$  est l'énergie émise dans la direction associée à  $(x, y)$ , au temps  $t$  sur la longueur d'onde  $\lambda$  -. L'information à transmettre doit être à une seule dimension temporelle. Les propriétés trichromatiques de la vision font que l'information spectrale  $f(\lambda)$  peut être ramenée à un vecteur  $Y = (R, V, B)$ . Un balayage périodique de période  $T$  défini par deux fonctions de période  $T$ ,  $x(t)$  et  $y(t)$ , permet de ramener l'information à transmettre sous forme de trois signaux  $R(t) = \int r(\lambda) f[x(t), y(t), t, \lambda] d\lambda$ , de même  $V(t)$ ,  $B(t)$ . Le balayage s'effectue en général en lignes, la ligne suivante étant légèrement en-dessous de la précédente sauf en bas du champ où on remonte analyser le haut de la scène (période - image =  $T$ ). Le nombre de lignes, leur longueur et la vitesse d'analyse ne peuvent être que finis. En télévision et en visiophone, la distance d'observation ainsi que le champ choisi et une certaine correspondance entre les définitions verticale et horizontale permettent de choisir le nombre de lignes minimales ainsi que la bande passante une fois  $T$  choisie. Le choix de la période - image dépend de deux défauts: le papillotement qui aux luminances envisagées nécessite  $T < 20$  ms et le rendu non saccadé des mouvements qui nécessite seulement  $T < 70$  ms à 80 ms en gros; la persistance des phosphores joue un rôle de filtrage: une persistance trop grande diminue la visibilité du papillotement mais induit un traînage sur les mouvements. Le papillotement étant un phénomène temporel agissant sur les basses fréquences spatiales, une cadence d'images de 25 Hz associée à un entrelacement d'ordre 2 (les lignes impaires sont

balayées pendant la première trame, les lignes paires pendant la seconde) induit un papillotement à 50 Hz au lieu de 25 Hz donc à la limite de visibilité. Des entrelacements d'ordre supérieur sont envisageables mais ils font apparaître un défaut de défilement ou de vibration dû au fait que l'œil s'accroche aisément à des structures à défilement continu. L'entrelacement d'ordre 2 produit en même temps une structure défilante montante et une descendante qui se masquent l'une l'autre et ne sont pas visibles sauf quand l'œil se déplace volontairement à leur vitesse.

Les signaux analogiques transmis (PAL et SECAM) sont  $Y' = 0,3 R^{1/\gamma} + 0,5 V^{1/\gamma} + 0,11 B^{1/\gamma}$  (pseudoluminance),  $D'_R = R^{1/\gamma} - Y'$ ,  $D'_B = B^{1/\gamma} - Y'$  (signaux de chrominance). Les bandes passantes sont respectivement 6 MHz et 1,5 MHz. En SECAM, chaque ligne n'a qu'une information  $D'_R$  ou  $D'_B$ , la restitution se faisant par répétition.

III. NUMERISATION: ECHANTILLONNAGE ET QUANTIFICATION

1) Echantillonnage Les systèmes analogiques comportent déjà deux échantillonnages: un temporel (structure d'image) et un vertical (structure de ligne et de trame du fait de l'entrelacement). Les systèmes numériques ne transmettent que des éléments binaires. Le signal vidéo doit donc être échantillonné (par rapport à l'image les échantillons sont des points formant un quadrillage régulier à maille rectangulaire ou losange). L'échantillonnage le plus simple à la fréquence de Shannon donne une structure rectangulaire; le caractère a priori isotrope <sup>des images</sup> autorise une meilleure compactification du spectre du signal échantillonné, correspondant à une structure de points en quinconce. On peut la réaliser par une fréquence d'échantillonnage égale à un multiple impair de la demi-fréquence ligne. Le signal vidéo échantillonné se replie alors en insérant de l'énergie entre les multiples de la fréquence ligne (le spectre vidéo a de l'énergie seulement autour des multiples de cette fréquence [1]). La restitution se fait par un filtre en peigne. Un mauvais échantillonnage fera apparaître des défauts sur les contours (moiré s'il y a des structures répétitives dans l'image). Le sous-échantillonnage du fait de l'entrelacement, de la non-identité de la reproduction des lignes consécutives et de l'imparité du nombre de lignes (la structure en quinconce s'inverse d'image à image) fait apparaître des structures défilantes sur



CARACTERISATION DES DEFAUTS APPORTES PAR  
LES CODAGES NUMERIQUES DE SIGNAUX D'IMAGE

les contours et réduit la définition horizontale si on ne le corrige pas. La faisabilité du sous-échantillonnage est en cours d'étude au C. C. E. T. T. (projet OCCITAN). On peut remarquer que la structure plus lâche ( $\sim 3$  fois) des points de chrominance rend les défauts de positionnement ou de rendu des contours beaucoup plus visibles.

2) Quantification: La dynamique est en général quantifiée avec un nombre entier d'e. b. Deux phénomènes entrent en jeu: dans les zones à peu près constantes, un nombre trop faible de niveaux fait apparaître des faux contours du fait d'une transition d'un niveau au suivant avec une corrélation spatiale de la position de commutation. L'oeil qui cherche les corrélations spatiales linéaires (contours) y est particulièrement sensible. Le deuxième défaut est le bruit granulaire dû au fait que la plus petite différence entre les niveaux est fixe, non nulle. La solution d'optimisation semble être d'uniformiser la visibilité de ces défauts dans l'espace de l'amplitude du signal à quantifier. En luminance, l'optimum est une quantification logarithmique (pour les deux défauts cités), la quantification linéaire de  $Y' = Y^1/Y$  est proche de l'optimum [2]; 7 e. b. suffisent [3]. DR' et DB' peuvent aussi être quantifiés avec 7 e. b., mais l'espace  $Y'$ , DR' et DB' n'est pas une représentation uniforme de l'espace coloré. Il est à remarquer que les faux contours et le bruit granulaire interviennent bien sûr dans les codages qui ont lieu après la représentation numérique.

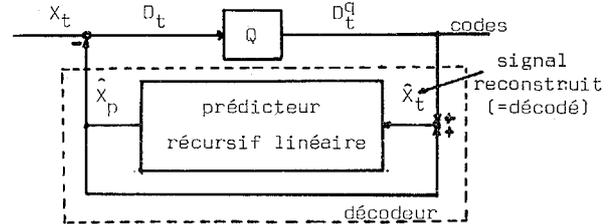
#### IV. SYSTEMES DE CODAGE

Trois types de systèmes seront décrits brièvement ici:

systèmes prédictifs d'ordre 0: on utilise les statistiques de plage à  $\pm \epsilon$ . L'amplitude  $X$  du premier point est transmise ainsi que la longueur de la plage  $X \pm \epsilon$ . La compression se fait par régularisation du débit d'information.  $\epsilon$  doit être assez petit pour éviter les faux contours, le bruit granulaire est lissé. Le système rend correctement les contours francs quoique des fluctuations sur la position du début de plage et de la fin de plage existent quand  $\epsilon$  est petit (flottement de contour).

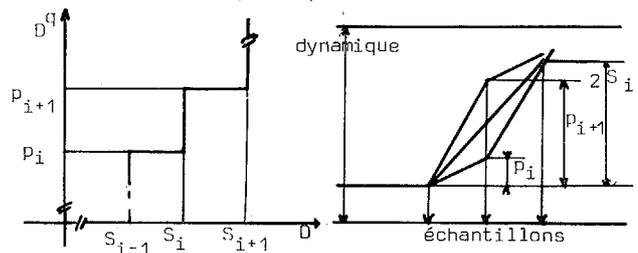
systèmes différentiels à prédiction: la compression se fait en quantifiant avec peu d'e. b. (3 ou 4) le signal de différence entre le signal  $X_t$  et une prédiction récurrente linéaire [4] de cette valeur:  $\hat{X}_p = P(\hat{X}_t - 1, \dots)$ , qui utilise les valeurs des points déjà transmis (décodés). Le décodeur est ainsi inséré dans la boucle de différenciation éliminant

l'accumulation des erreurs de quantification [5]:



Les petites différences sont quantifiées finement pour éviter le bruit granulaire. Les propriétés de la vision font que l'on peut se permettre une imprécision dans le rendu des variations, d'autant plus grande que la variation est importante. D'où la forme générale de la quantification utilisée: niveaux de reproduction de plus en plus écartés pour les différences de plus en plus grandes. Deux défauts interviennent dans les systèmes: le trainage qui est dû au fait que le plus grand niveau de reconstruction est plus petit que la dynamique donc le signal décodé ne rattrape les contours raides qu'après quelques points, le flottement de contour qui apparaît dès que le rendu d'un contour par un codeur fluctue de ligne à ligne ou d'image à image. Ce défaut est très gênant car il apparaît sur des configurations corrélées dans l'espace et le temps (contours fixes ou mobiles - quoique le mouvement fasse diminuer la visibilité). Une fluctuation d'un quart de point en visiophone pour un contour franc est visible.

Le phénomène peut être schématisé ainsi (prédiction par le point précédent):



Si le contour a une pente proche d'un des seuils  $s_i$  du quantificateur du signal de différence  $D$ , la différence sera codée  $p_i$  ou  $p_i + 1$  suivant que la différence  $D$  est supérieure ou inférieure à  $s_i$ , un très faible bruit peut faire commuter le premier code entre  $p_i$  et  $p_i + 1$  de ligne à ligne (contour vertical par exemple) et d'image à image. Si le premier code est  $p_i$ , la nouvelle différence dépassera  $s_i$ , le second code sera  $p_i + 1$  pour rattraper le contour d'où le rendu ①, dans le cas contraire le rendu sera ②. Le changement aléatoire entre ① et ② qui apparaît sur le contour à coder produit le flottement de contour.



systèmes à transformation: on utilise des transformations linéaires orthogonales (Fourier, Hadamard, Haar etc...) pour transformer une image  $F(x, y)$  en sa transformée  $\hat{F}(u, v) = \sum_{x,y} A(u, v, x, y) F(x, y)$ . Les composantes  $\hat{F}(u, v)$  sont en général décorréliées et leur énergie varie beaucoup suivant  $u, v$  permettant de quantifier grossièrement et même de ne pas transmettre celles qui ont une énergie faible. On a inversement  $F(x, y) = \sum_{u,v} A^{-1}(u, v, x, y) \hat{F}(u, v)$ . Une erreur de quantification sur  $\hat{F}(u_0, v_0)$  fait apparaître en superposition une structure qui est l'image de base  $A^{-1}(u_0, v_0, x, y)$  de la transformation. Des faux contours apparaissent ainsi que des rendus en escalier des contours qui sur une séquence d'images produiront du flottement de contour.

effets des erreurs de transmission: Les canaux de transmission ont en général un taux d'erreur ne portant pas à conséquence sauf pendant un certain pourcentage du temps où il convient que les systèmes ne soient pas trop sensibles aux erreurs. Plus le codage comprime l'information, plus il est évidemment sensible aux erreurs de transmission quoique la visibilité de ces erreurs puisse varier beaucoup pour de petites variations de paramètres de codage. La visibilité dépend en effet de l'étendue de la propagation des effets d'une erreur sur l'image. Par exemple, dans le cas d'un codage différentiel, la prédiction par le point précédent fait propager les effets d'une erreur jusqu'à la fin de la ligne, une prédiction bidimensionnelle fait propager les effets jusqu'à la fin de l'image mais l'erreur s'amortit rapidement et s'étale ce qui rend le défaut moins visible que dans le premier cas.

#### V. LES PRINCIPAUX DEFAUTS ET LEUR VISIBILITE.

- faux contours. Ils sont dûs à la commutation d'un niveau au suivant dans les zones quasiment uniformes. Leur visibilité est décrite par les données sur les seuils différentiels [6], [7]. Elle décroît rapidement avec l'angle visuel du test et est optimum quand l'environnement (adaptation) est à la même luminance moyenne que les test [6].

- bruit granulaire. Il est dû au fait que les plus petites variations sont rendues par des variations fixes. C'est un effet temporel analogue au bruit analogique, il apparaît plus sur les zones uniformes et semble suivre les mêmes lois que les seuils différentiels en fonction de la luminance (le premier étage de la rétine est logarithmique) [8]. Sa visibilité dépend aussi de son spectre [9].

- traînage. C'est une saturation du temps de montée des contours, qui donne un effet de flou, mais uniquement sur les contours (différant ainsi du flou usuel - linéaire - qui ne dépend pas de l'amplitude du signal). Il est d'autant plus visible que le contour est net et contrasté.

- flottement de contour. Il est dû aux fluctuations du rendu d'un même contour de ligne à ligne et d'image à image. Il apparaîtra dans tous les systèmes non-linéaires pour lesquels les transitions brutales (ici contours) excitent des instabilités. La visibilité vient du fait que les erreurs sont corrélées spatialement avec un contour, elle croît d'autant plus que le contour est net et contrasté. Son effet subjectif ressemble à un grouillement de points sur le contour; si les erreurs sont à des fréquences fixes (ligne par exemple), il apparaît comme un défillement de points sur le contour (cas du sous-échantillonnage où le rendu diffère suivant la parité des lignes).

- combinaison de défauts. Des essais ont eu lieu pour mesurer la dégradation d'une combinaison de défauts connaissant la dégradation produite par chacun séparément. Des expériences sur des défauts analogiques indépendants ont montré une additivité des "imps", mesure de la dégradation ("impairment") égale à  $\frac{1}{p} - 1$ , si  $p$  est la note moyenne attribuée (échelle normalisée de 0 à 1): [10], [11]. A priori, ce résultat devrait être vrai pour les défauts numériques. Une expérience plus récente [12] consiste à comparer le défaut étudié à du bruit blanc ajouté à l'image, les observateurs devant chercher la puissance qui produit la même dégradation subjective. La méthode donne des résultats avec des observateurs expérimentés mais semble peu réaliste.

On peut remarquer qu'il est nécessaire que les défauts combinés dans ces expériences n'interfèrent pas: en effet deux défauts peuvent se masquer [13] l'un l'autre, voir s'annihiler (cas des deux structures défilantes de l'entrelacement d'ordre 2).

#### VI. MESURES ET OPTIMISATION.

L'équilibre est à trouver entre les deux méthodes extrêmes: la première consiste à s'appuyer sur des données psychophysiques (déjà disponibles ou à étudier), la seconde consiste, une fois le codeur réalisé avec un choix suffisant de paramètres, à lui faire subir des tests subjectifs systématiques pour définir la configuration optimale. Les données psychophysiques sont des données de visibilité (contraste

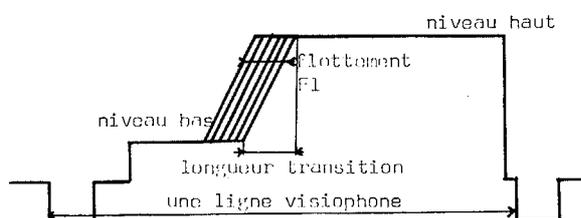


CARACTERISATION DES DEFAUTS APPORTES PAR  
LES CODAGES NUMERIQUES DE SIGNAUX D'IMAGE

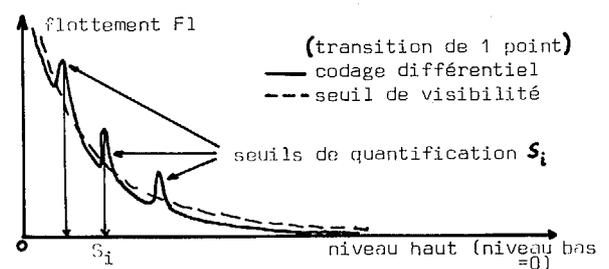
minimum, angle apparent minimum etc...) sur des configurations visuelles (spatiotemporelles, colorées dans le cas le plus général) schématiques donc plus critiques que les configurations apparaissant en général dans les images. Elles constituent une bonne base pour une qualité irréprochable mais l'introduction de statistiques sur les images pour compenser le caractère schématique des tests psychophysiques n'est pas facile. Les tests subjectifs de systèmes complets seront toujours utilisés pour en définir la qualité mais ils sont assez lourds d'emploi et ne permettent pas d'optimiser des systèmes à grand nombre de paramètres. Une certaine schématisation des défauts est donc nécessaire et l'emploi, à ce stade, de tests subjectifs est essentiel, le problème étant de modéliser le défaut et d'en trouver une mesure objective nécessaire à l'optimisation, mesure qui soit suffisamment corrélée avec la perception du défaut.

Un bon exemple de la différence entre les deux méthodes est l'étude du flottement de contour ("edge busyness"). La méthode de J. C. Candy [12] s'applique au codage différentiel qui code l'erreur  $x$  de prédiction. Un flottement de contour est introduit dans une image réelle en rajoutant du bruit sur le signal vidéo quand la différence dépasse un seuil  $d$ . La dégradation est mesurée par la puissance de bruit équivalente  $p(d)$ . La dérivée  $f(d)$  de  $p(d)$  donne une estimation de la visibilité d'un bruit ajouté quand  $x = d$ . La quantification de  $x$  est déduite de la courbe  $f(d)$ . Cette méthode peut s'appliquer à la quantification de tout signal  $x$  issu d'un codage: elle est actuellement mise en oeuvre aux Laboratoires Bell pour quantifier l'espace coloré et quantifier les composantes d'une transformée tri-dimensionnelle ( $2 \times 2 \times 2$  en  $x, y, t$  de Hadamard). La critique en est que l'addition d'un bruit blanc important sur un signal n'équivaut pas à le quantifier grossièrement et d'autre part la comparaison du flottement de contour avec du bruit blanc est surprenante.

La seconde méthode a été étudiée par les auteurs et a abouti à la réalisation d'unemaquette simulant le flottement de contour sur un signal schématisé [14] (standard visiohone):



Un déphasage aléatoire a été utilisé pour simuler le flottement de contour en partant de l'hypothèse que sa visibilité était d'ue aux variations dans la position détectée par l'oeil du contour. Le système électronique génère une ligne comme ci-dessus; on peut programmer le niveau bas, le contraste, le nombre de points de la transition, le nombre de phases différentes existées, le mode d'excitation (pseudo-aléatoire binomial ou uniforme, éventuellement calculé à volonté) et l'inclinaison du contour sur l'écran. L'écart minimum entre deux phases est de 1/16ème de point. L'écran peut être séparé en diverses zones pour jouer sur la longueur de la transition et pour faire des comparaisons avec un flottement réel produit par un codeur quelconque sur le même signal, le mode de production étant d'ajouter du bruit analogique au signal avant codage. Pour certaines configurations (quand la pente de la transition correspond à un seuil pour un codeur différentiel), le flottement est visible sans que le bruit granulaire le soit sur les plages. Des données de visibilité ont été obtenus pour le flottement simulé: le seuil de visibilité croît quand le contraste diminue et quand le contour est moins raide. Un essai de calcul du flottement sur un codage simulé sur ordinateur (mesure de la largeur significative  $F1$  du nuage des réalisations d'un contour suivant la réalisation du bruit de source) a été faite. On obtient des courbes comme ci-dessous (codage différentiel):



Les maxima sont bien corrélés avec les seuils du quantificateur des différences. L'allure générale des courbes (décroissance du seuil de perception avec le contraste) correspond bien aux tests psychophysiques. Des tests complets ont lieu actuellement pour conclure l'étude. Nous souhaitons montrer le bien-fondé de la méthode qui permettra d'utiliser les données psychovisuelles qui fournissent la visibilité en fonction du flottement  $F1$  suivant les paramètres du contour, en ayant au préalable estimé suivant les configurations le flottement  $F1$  effectivement produit par tel ou tel codeur. Une optimisation des paramètres du codeur pourra ainsi être menée par programmation.



VII. CONCLUSION

Les défauts des codeurs numériques ont été détaillés ci-dessus avec leurs causes et les paramètres de leur visibilité. Même si l'"additivité" des dégradations a été mentionnée, il reste un problème qui se pose fréquemment: c'est l'équilibre à trouver entre deux défauts dont l'un apparaît quand l'autre est réduit et vice-versa. Un bon exemple est le choix entre le traînage et le flottement de contour (en général un rendu correct des transitions brutales est sujet à une instabilité de phase): [15]. Une règle générale semble impossible à définir; une approche lente et difficile mais certainement fructueuse est d'approfondir les données fort nombreuses sur la perception visuelle [13].

REFERENCES

- [1] L. E. FRANKS, "A Model for the Random Video Process", B. S. T. J., avril 1966, pp 609-629.
- [2] F. KRETZ, "Subjectively Optimal Quantization of Pictures", à paraître dans IEEE Tr. on Com.
- [3] T. A. MOORE, "Digital Video: Number of Bits per Sample Required for Reference Coding of Luminance and Colour-Difference Signals", B. B. C. Res. Depart. 1974/42, dec. 1974.
- [4] J. B. O'NEAL, "Predictive Quantizing systems (DPCM) for the Transmission of Television Signals", B. S. T. J., mai-juin 1966, pp 689-721.
- [5] J. O. LIMB, F. W. MOUNTS, "Digital Differential Quantizer for Television", B. S. T. J., sept. 1969, pp 2583-2599.
- [6] P. MOON, D. E. SPENCER, "The Visual Effect of Nonuniform Surrounds", J. O. S. A., mars 1945, pp 233-248.
- [7] D. L. MACADAM, "Visual Sensitivities to Color Differences in Daylight", J. O. S. A., mai 1942, pp 247-274.
- [8] K. HACKING, "The Relative Visibility of Random Noise over the grey-Scale", J. BRIT. I. R. E., avril 1962, pp 307-310.
- [9] Z. L. BUDRIKIS, "Visual Thresholds and the visibility of random noise in T. V.", Pr. I. R. E. Austr., dec. 1961, pp 751-759.
- [10] R. D. PROSSER et al., "Quality grading of impaired television pictures", Pr. IEE, mars 1964, pp 491-502.
- [11] N. W. LEWIS et J. W. ALLNATT, "Subjective Quality of Television Pictures with Multiple Impairments", Electr. Letters, sept. 1965, pp 187-188.
- [12] J. C. CANDY et R. H. BOSWORTH, "Methods for Designing Differential Quantizers Based on Subjective Evaluations of Edge Busyness", B. S. T. J., sept. 1972, pp 1495-1516.
- [13] Z. L. BUDRIKIS, "Visual Fidelity Criterion and Modeling", Pr. IEEE, juillet 1972, pp 771-779.
- [14] J. PONCIN et F. KRETZ, "Problèmes de mesures objectives de la qualité des images codées", Congrès AFCET 1973, Tome 1, pp 33-42.
- [15] R. T. BOBILIN, "Prefilters, Sampling and Transmission Rates for Intraframe Codecs for Picturephone Service", B. S. T. J., avril 1973, pp 497-525.