



## POURSUITE VIDEO D'ICONES VARIANT EN FORME ET EN COULEUR

D. SAGE - P.Y. COULON

Laboratoire TIRF-ENSERG - Institut National Polytechnique de Grenoble  
46, avenue Félix Viallet 38031 GRENOBLE CEDEX - FRANCE

### RÉSUMÉ

Nous présentons dans ce papier un système de poursuite d'objets. Il détermine la trajectoire d'objets mobiles à partir d'une séquence d'images vidéo. L'objet est décrit par des icônes qui sont des petites matrices de pixels représentant un motif discriminant. Le système de poursuite est organisé en modules dont les trois principaux sont le prétraitement, le positionnement de la fenêtre de poursuite et la localisation de l'icône. Une modélisation dynamique des icônes permet de suivre les objets malgré des variations temporelles de forme et de couleur. Nous étudions un prétraitement couleur qui consiste à transformer une image couleur (3 plans RVB) en une image noir et blanc (1 plan). La transformation se ramène à une projection des couleurs de l'espace RVB sur un axe passant par la couleur de la cible et la couleur du fond afin d'améliorer au mieux le contraste. Ce système de poursuite est installé sur un poste de travail image piloté par un micro-ordinateur avec un magnétoscope commandable. Ce système est utilisé comme outil d'assistance pour analyser des films vidéo.

### SUMMARY

In this paper, we present a tracking objects system. It provides the trajectory of moving objects from a video image sequence. The object is described by icons which are small matrix of pixels representing a discriminant pattern. The tracking system is composed of several modules. The three main ones are preprocessing, tracking window opening and localization of the icon. A dynamic icon modeling allows to follow the objects even with time-varying form and color. We study a color preprocessing which transforms a color image (3 planes RGB) into a black and white image (1 plane). The transformation is a projection of colors belonging to RGB-space, on an axis. In order to have the best contrast, the choosen axis includes the target color and background one. The algorithms are performed on an image workstation controlled by a micro-computer with a video tape recorder. This tracking device is used as an assistance tool to analyse video films.

#### 1. Introduction

L'analyse automatique de séquences d'images, en vue d'étudier des mouvements suscite un vif intérêt, en raison de multiples applications dans des secteurs d'activité très variés (robotique, médical, sportif, météo ...). L'application qui a motivé notre étude concerne une assistance automatisée au dépouillement des films vidéo du Centre Technique du Papier de Grenoble. Il s'agit de prises de vue effectuées sur des machines de conditionnement de cartons sur lesquelles les cycles défectueux doivent être détectés ; nous devons alors analyser un cycle représenté par une séquence d'images. Ce type d'analyse est relativement difficile autant pour des problèmes théoriques que pratiques, le principal d'entre eux est l'énorme quantité d'information à manipuler.

Les méthodes d'analyse de séquences d'images relèvent de trois approches [1] :

- (a) Méthode de Fourier : étude du mouvement dans l'espace fréquentiel. Les déplacements (translations) d'objets ont des effets simples sur la transformée de Fourier d'une image.
- (b) Méthode différentielle : estimation complète du mouvement apparent, c'est à dire l'ensemble des vecteurs vitesse associés à chaque pixel (flot optique) [2].
- (c) Méthode de correspondance : mise en correspondance des caractéristiques de deux images successives. Ces caractéristiques sont plus ou moins élaborées. On distingue, en général, deux niveaux :

- Niveau intensité (bas niveau) : La correspondance est établie directement sur les pixels de l'image.
- Niveau structurel (haut niveau) : Avant la mise en correspondance, un processus extrait des indices visuels de l'image, typiquement des contours. La mise en correspondance s'effectue alors sur ces indices visuels.

Des réalisations avec des processeurs dédiés ont vu le jour comme par exemple [3] pour la mise en correspondance bas niveau et [4] pour le haut niveau.

Dans l'optique de la réalisation d'un système de poursuite, les deux premières méthodes sont inapplicables, elles conduisent à des coûts calcul excessifs qui sont en tout état de cause incompatibles avec les contraintes de temps imposées pour la poursuite. Nous nous orientons donc vers une méthode de mise en correspondance. Nous avons déjà montré comment estimer le mouvement d'un angle par une technique de mise en correspondance au niveau intensité ou au niveau structurel [5].

#### 2. Poursuite d'icônes

##### 2.1 - Principe

Une poursuite débute par une phase d'initialisation où l'opérateur désigne la cible à suivre sur la première image, puis la poursuite, à proprement parler, est exécutée sur les images suivantes. Le système suit automatiquement la cible dans tous ses déplacements. En fin d'analyse il peut reconstruire la trajectoire de la cible.



La principale difficulté en analyse de séquences d'images vient de l'énorme masse d'information à traiter. L'espace de recherche est par nature tridimensionnel : les deux dimensions spatiales et la dimension temporelle. La technique de la poursuite est fondée sur une limitation de la zone de recherche :

- Dans le temps - On traite les images les unes à la suite des autres en stockant une seule image en un temps donné. L'information provenant des images précédentes est contenue dans un modèle de l'objet.
- Dans l'espace - On ne traite qu'une zone réduite de l'image que l'on appelle fenêtre de poursuite, susceptible de contenir l'information intéressante.

Ces deux limitations sont nécessaires pour réaliser un système de poursuite efficace qui puisse être installé sur des configurations matérielles réduites.

## 2.2 - Icône

Dans notre système de poursuite la cible est une icône. Une icône est une petite zone de l'image décrite par une matrice de pixels. Elle représente un motif particulier de la scène comme un trou ou un coin d'une pièce mécanique. Si l'objet à suivre est petit, il peut entrer complètement dans l'icône, sinon l'objet peut être caractérisé par plusieurs icônes (Figure 1).

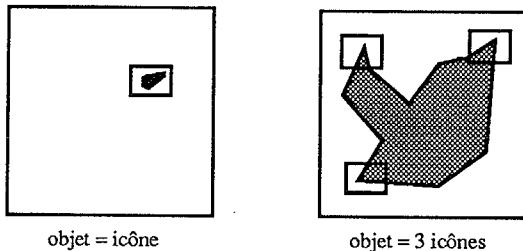


Figure 1 - Description d'un objet par des icônes

Une icône est a priori quelconque, la seule condition est qu'elle soit visuellement discriminante par rapport au reste de l'image. Pendant la séquence d'images l'icône subit des variations de forme et de couleur.

- Forme : Les déformations de la cible sont provoquées par la non-rigidité des objets. Les changements de forme sont également causés par la projection sur une image bidimensionnelle d'un objet tridimensionnel en mouvement.
- Couleur : Les variations de couleur au cours d'une séquence d'images sont faibles, elles sont provoquées par des modifications de la source lumineuse ou par des phénomènes d'ombre.

Pour effectuer la poursuite dans de bonnes conditions, il est nécessaire d'avoir une hypothèse de continuité de toutes ces variations. Comme nous traitons des images réelles, issues d'un monde physique qui est par essence continu, l'hypothèse est souvent vérifiée.

## 2.3 - Modélisation

Toutes les caractéristiques de l'icône sont regroupées dans le modèle de l'icône. Nous avons vu que ses caractéristiques évoluent dans le temps d'où la nécessité d'avoir un modèle dynamique. Nous avons décomposé le modèle de l'icône en trois parties qui servent chacune à une étape importante de la poursuite.

- Modèle cinématique : Il contient une description des mouvements de la cible (position, vitesse courante).
- Modèle morphologique : Il contient une description de la forme de l'icône décrite par une matrice de pixels.
- Modèle statistique couleur : Les propriétés statistiques de l'icône sont calculées à partir de la matrice de pixels.

## 2.4 - Structure

La structure générale du système de poursuite est présentée à la figure 2 [6].

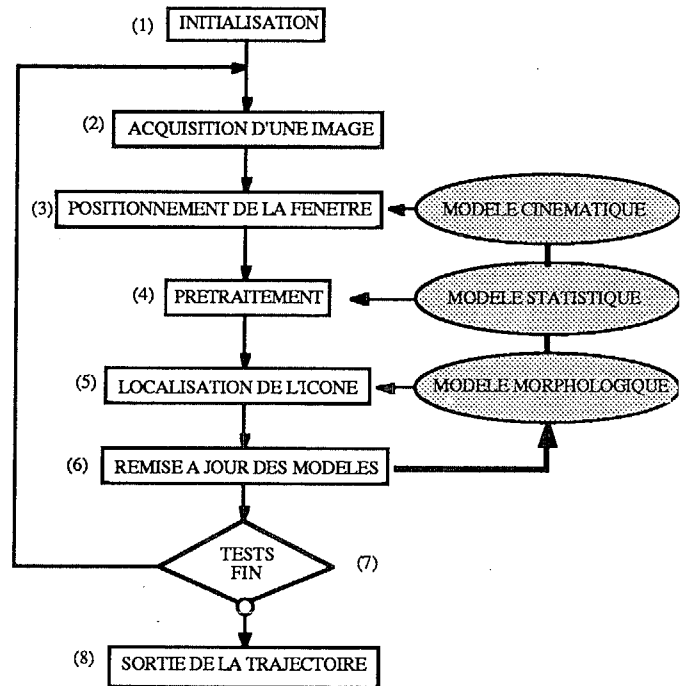


Figure 2 - Synoptique de la poursuite

### (1) Initialisation

Elle est réalisée sur la première image de la séquence, l'opérateur pointe l'icône choisie avec une souris.

### (2) Acquisition d'une image

L'acquisition d'une image vidéo dure 40 ms, elle provient d'une source d'images qui est soit une caméra soit un magnétoscope.

### (3) Positionnement de la fenêtre

La fenêtre est placée sur la position estimée de l'icône qui est déterminée par extrapolation de la trajectoire et avec le modèle cinématique.

### (4) Prétraitement

L'étape prétraitement est importante pour améliorer la qualité des images. [7] a montré qu'un filtrage du type passe-haut sur des images à niveau de gris rendait plus facile la localisation de cible. Nous étudions un prétraitement couleur dans le paragraphe 3.

### (5) Localisation de l'icône

Il s'agit de retrouver l'icône dans la fenêtre de poursuite. La méthode de recherche de l'icône est basée sur une mise en correspondance niveau intensité. La procédure consiste à évaluer une mesure de ressemblance (fonction de corrélation ou somme des différences absolues) entre l'icône et une partie de la fenêtre.

### (6) Remise à jour des modèles

Les modèles sont réactualisés après chaque analyse pour s'adapter aux variations.

### (7) Tests de fin

On a plusieurs événements qui stoppent la poursuite.

- Arrêt demandé par l'opérateur.
- L'icône sort de l'image, c'est à dire du champ visuel de la caméra.
- Le magnétoscope est en fin de bande.
- Le décrochage : La perte d'une icône est imputable à un changement brutal d'une des caractéristiques de l'icône (vitesse, forme, couleur). A ce stade, il faut distinguer une perte prolongée qui stoppe la poursuite d'une perte temporaire (occlusion) qui fait l'objet d'un traitement spécial d'extrapolation de la trajectoire.

### (8) Sortie des trajectoires

Les trajectoires sont visualisées sur un moniteur ou sur une imprimante.

Les diverses expérimentations que nous avons menées avec ce système de poursuite montrent la validité des différents algorithmes sous les hypothèses suivantes : icône discriminante et vitesse lente.

### 3. Prétraitement couleur

Notre système de poursuite exploite des images noir et blanc comme la plupart des algorithmes et processeurs dédiés de traitement d'images. Mais si l'on dispose au départ d'images couleur, il est intéressant de transformer l'image couleur en une image noir et blanc en essayant de conserver des propriétés pertinentes de l'image couleur.

Si la transformation est linéaire, il s'agit alors de prendre une couleur de coordonnées (r,v,b) qui est un point de l'espace RVB et de projeter cette couleur sur un axe. L'abscisse obtenue sur cet axe correspond à un niveau de gris. Le choix de l'axe est fonction du type d'information que l'on souhaite privilégier.

Dans le cadre du système de poursuite nous avons trois critères:

- augmenter localement le contraste
- minimiser la perte d'information
- minimiser le coût calcul

Nous proposons l'axe différence qui est basé sur la constatation suivante : l'histogramme trivarié RVB d'une icône est souvent bimodal. Nous nommons les deux classes de points : ensemble cible et ensemble fond par référence au cas de poursuite le plus courant, c'est à dire la poursuite d'une cible se déplaçant sur un fond. Mais ces deux ensembles de points peuvent appartenir à deux objets ou être de part et d'autre d'un contour. L'axe différence passe par la couleur de la cible et la couleur du fond (Figure 3).

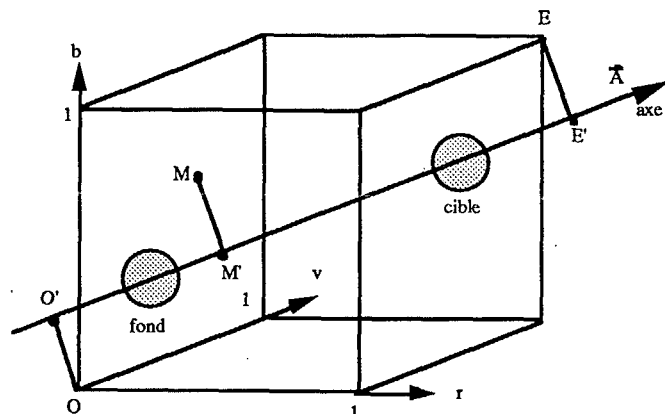


Figure 3 - Espace RVB et axe de projection

Pour une couleur M, le niveau de gris correspondant sur l'image noir et blanc est :

$$n = \frac{\vec{A} \cdot \vec{OM}}{\vec{A} \cdot \vec{OE}} \text{ avec } \vec{A}(R_a, V_a, B_a)$$

Nous normalisons avec O'E' qui a la projection d'une diagonale du cube RVB. Les couleurs de base (sommets du cube) servant d'origine O (Ro,Vo,Bo) et d'extrémité E (Re,Ve,Be) sont déterminées de la façon suivante:

- Si  $R_a > 0$  alors  $R_e = 1, R_o = 0$  sinon  $R_e = 0, R_o = 1$
- Si  $V_a > 0$  alors  $V_e = 1, V_o = 0$  sinon  $V_e = 0, V_o = 1$
- Si  $B_a > 0$  alors  $B_e = 1, B_o = 0$  sinon  $B_e = 0, B_o = 1$

Pour étudier le comportement de l'axe différence face aux trois critères nous faisons les simplifications et l'hypothèse suivantes qui ne remettent pas en cause les résultats obtenus:

Simplification 1: Nous traitons le cas bidimensionnel (RV) au lieu de travailler sur l'espace RVB.

Simplification 2: Nous étudions le cas où  $R_a > 0, V_a > 0, B_a > 0$  c'est à dire que la diagonale servant d'unité est la première diagonale (Noir-Blanc).

Hypothèse : La couleur de la cible et du fond sont des distributions gaussiennes indépendantes et de covariance identique sur chaque composante.

Notation : Axe de projection A de direction  $\alpha$

Description statistique avant projection :

cible moyenne  $\begin{pmatrix} mr1 \\ mv1 \end{pmatrix}$  matrice de covariance  $C_1 = \sigma_1^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$   
 fond moyenne  $\begin{pmatrix} mr2 \\ mv2 \end{pmatrix}$  matrice de covariance  $C_2 = \sigma_2^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Description statistique après projection :

cible moyenne  $M_1$  variance  $S_1^2$

fond moyenne  $M_2$  variance  $S_2^2$

Avec ces notations la formule de projection s'écrit pour une couleur de coordonnées (r,v) :

$$n = \frac{r \cdot \cos \alpha + v \cdot \sin \alpha}{\cos \alpha + \sin \alpha}$$

#### Maximiser le contraste local

C'est le critère auquel nous attachons la plus grande importance. Nous définissons le contraste de la façon suivante (différence statistique) :

$$\Gamma = \frac{|M_1 - M_2|}{\sqrt{S_1^2 + S_2^2}}$$

Nous avons :

$$M_1 = \frac{mr1 \cdot \cos \alpha + mv1 \cdot \sin \alpha}{\cos \alpha + \sin \alpha} \quad S_1 = \frac{\sigma_1}{\cos \alpha + \sin \alpha}$$

$$M_2 = \frac{mr2 \cdot \cos \alpha + mv2 \cdot \sin \alpha}{\cos \alpha + \sin \alpha} \quad S_2 = \frac{\sigma_2}{\cos \alpha + \sin \alpha}$$

d'où :

$$\Gamma = \frac{|(mr1 - mr2) \cdot \cos \alpha + (mv1 - mv2) \cdot \sin \alpha|}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

En annulant la dérivé du contraste, nous trouvons que l'axe qui le maximise est tel que :

$$\alpha = \text{Arctg} \left( \frac{mv1 - mv2}{mr1 - mr2} \right)$$

C'est donc l'axe différence qui passe par le centre de la couleur cible et par le centre de la couleur fond qui maximise le contraste.

#### Minimiser la perte d'information

L'objectif est de garder un maximum d'information sur l'axe de projection. D'après la théorie de l'analyse en composantes principales (transformée de Karhunen-Loeve), l'axe qui minimise la perte d'information est l'axe principal. [8] a réalisé un processeur qui transforme une image couleur en une image noir et blanc par projection sur l'axe principal. On peut montrer très facilement que l'axe différence est l'axe principal avec le modèle statistique que nous avons choisi, il réduit donc la perte d'information due à la projection sur un axe d'un espace tridimensionnel. Mais l'axe différence est beaucoup plus facile à calculer.

#### Minimiser le coût calcul

Le temps de calcul se décompose en trois parties:

- détermination des caractéristiques statistiques
- calcul de l'axe de projection
- projection effective des couleurs sur l'axe

Le temps projection est le même quelque soit le type d'axe choisi. En ce qui concerne les deux autres, l'axe différence a un coût faible. En effet, pour déterminer les caractéristiques statistiques utiles il suffit de prendre quelques points de la cible et du fond et d'en déduire les vecteurs moyennes par lesquels passera l'axe de projection.

Nous avons montré que l'axe différence est celui qui répond au mieux aux trois critères que nous avons définis.

Le prétraitement couleur est appliqué à toute la fenêtre de poursuite mais les paramètres de la transformation sont calculés avec les caractéristiques de l'icône uniquement.

D'autre part, si l'on considère des covariances différentes sur chaque composantes, l'optimalité de l'axe différence n'est pas démontrable ; néanmoins, au vu des problèmes de séparabilité des classes en reconnaissance des formes l'heuristique 'Axe Différence' reste performante.



#### 4. Expérimentations

Les essais sont réalisés sur un poste de travail développé au laboratoire. Il est constitué de trois unités principales (figure 4):

- unité de traitement: ordinateur compatible PC/AT
- source d'images vidéo:
  - Caméra (traitement direct)
  - Magnéto standard VHS 1/2 pouce (traitement différé sans contrôle)
  - Magnéto de montage standard U-Matic 3/4 pouce (traitement différé avec contrôle) Sur la piste audio du magnéto nous enregistrons un code temps numérique SMPTE qui sert de référence temporelle.
- carte de digitaliseur et mémorisation d'images (512\*512 pixels codés sur 8 bits).

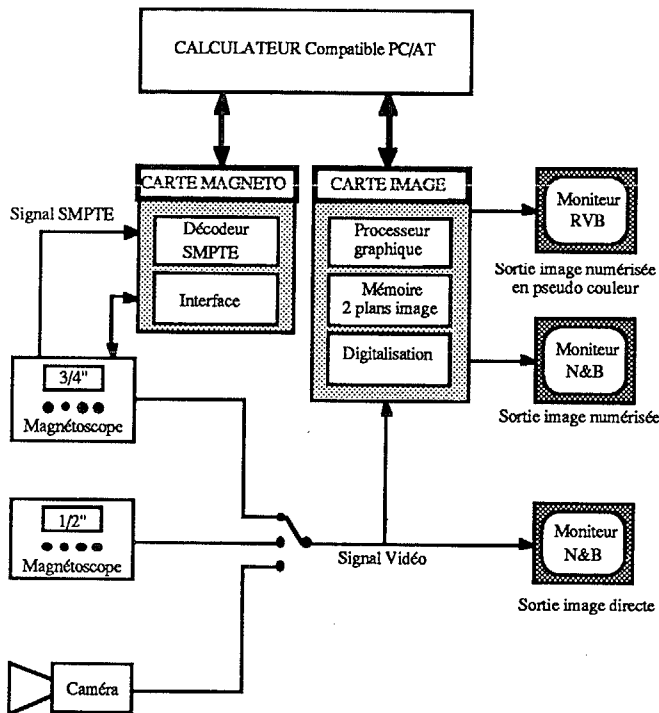


Figure 4 - Synoptique du poste de travail

La limite du système de poursuite est atteinte pour un mobile se déplaçant à plus de 80 pixels / seconde. Dans le meilleur des cas nous pouvons traiter une image en 250 millisecondes. L'algorithme de prétraitement couleur doit être installé sur une nouvelle version de carte couleur. Nous montrons deux exemples de poursuite d'une icône 15\*15 pixels sur des images à niveau de gris (Figure 5 et 6).

#### 5. Conclusion

La trajectométrie est un outil très intéressant pour faire du dépouillement automatique des films. Les grandes lignes qui caractérisent le système de poursuite proposé dans cet article sont:

- Adoption d'une structure modulaire.
- Modélisation dynamique de l'icône qui permet de la suivre malgré des déformations et des changements de couleur.
- Application à des scènes et à des cibles variées. A ce titre, le choix de l'icône comme cible est essentiel pour se rendre indépendant du domaine d'application.
- Développement d'algorithme à faible coût calcul.
- Réalisation sur du matériel standard sans développement de hardware particulier.
- Facilité d'utilisation puisqu'il suffit à l'opérateur de pointer l'icône avec une souris.

L'emploi d'un matériel standard limite les performances temporelles du système. Pour assurer un temps réel vidéo, une

réalisation câblée ou l'utilisation d'un coprocesseur de traitement du signal [9] doit être envisagée.

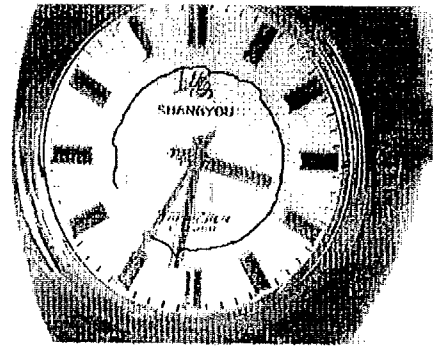


Figure 5 - Poursuite de la trotteuse d'une montre

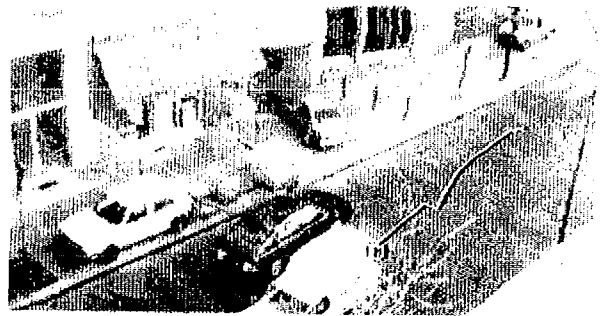


Figure 6 - Poursuite d'un véhicule

#### Références

- [1] T.S. HUANG and al. 'Image sequence analysis' T.S Huang, Springer Verlag, 1981.
- [2] P. BOUTHEMY 'Modèles et méthodes pour l'analyse du mouvement dans une séquence d'images' 2ème Atelier TIPI, Traitement d'Images: du Pixel à l'Interprétation, CNRS, 1988.
- [3] J.K. AGGARWAL, L.S. DAVIS, W.N. MARTIN 'Correspondance processes in dymanic scene analysis' Proceedings of the IEEE, Vol 69, N° 5, 1981.
- [4] P. STELMASZYK, DISCOURS, A. CHEHIKIAN 'A fast and reliable token tracker' Intelligence Artificiel and Pattern Recognition Workshop on Computer Vision, Tokyo, 1988.
- [5] D. SAGE, P.Y. COULON 'A contour feature tracking system' 7th Conference on Robot Vision and Sensory Controls (ROVISEC7), Zurich, 1988.
- [6] D. SAGE 'Poursuite d'objets par analyse d'une séquence d'images couleur' Thèse de doctorat, INPG, Grenoble, 1989.
- [7] J. CICCOTELLI 'Contribution du capteur d'image hybride N-F à la poursuite automatique de cible' Thèse de doctorat d'état, INPL, Nancy, 1987.
- [8] J.A. OCADIZ-LUNA 'Analyse en composantes principales d'une image couleur' Thèse de doctorat, INPG, Grenoble, 1985.
- [9] P. ASCHWANDEN 'Real-time tracker with signal processor' Signal Processing IV, EURASIP, Grenoble, 1988.

#### Remerciements

Cette étude a été menée en collaboration avec le Centre Technique du Papier de Grenoble et avec l'aide financière de l'Etablissement Public Régional de Rhône-Alpes.