



SEGMENTATION D'IMAGES PAR REGROUPEMENT PERCEPTIF

S. Guillaudeux, M. Klefstad-Sillonville, P. Leray

CCETT
Centre Commun d'Etudes de Télédiffusion et Télécommunications
4 rue du Clos-Courtel _ BP 59
35512 Cesson-Sévigné

RESUME

Dans cette communication, nous décrivons une méthode de segmentation de l'image s'appuyant sur la mise en oeuvre d'un concept général de regroupement perceptif. L'originalité de ces travaux réside dans le fait que le processus de regroupement utilisé opère simultanément sur plusieurs types d'information en ne faisant appel qu'à des notions simples de similarité-homogénéité, indépendamment des données traitées. Dans l'application de segmentation spatiale présentée ici, le regroupement agit sur des caractéristiques de type contours orientés, extraites de l'image lors d'une étape préliminaire.

Ces travaux s'inscrivent dans le cadre d'études amont de codage et compression d'images vidéo animées (images TV) pouvant déboucher à terme sur la spécification d'un système de codage s'inspirant des mécanismes bas-niveau du cortex visuel.

I • MODELISATION DU REGROUPEMENT PERCEPTIF

Des études psychophysiques [ORBA 84] ont montré que le système visuel humain bas-niveau analyse localement l'image, à différentes échelles spatiales, à l'aide de cellules sensibles à divers paramètres.

Les travaux d'Hubel et Wiesel ont permis de mieux comprendre certains mécanismes intervenant au niveau de l'aire visuelle primaire V1. Ils ont par exemple différencié plusieurs types de cellules :

- Les cellules "simples" dont le champ récepteur, aire rétinienne sur laquelle agit le neurone, présente une forme allongée dans une direction privilégiée. Ces cellules répondent de façon optimale à une barre dont l'orientation est identique à celle de leur champ récepteur.
- Les cellules "hypercomplexes" qui réagissent aussi à des barres orientées, mais en tenant compte de la longueur du stimulus : la réponse sera d'autant plus faible que la longueur dépassera celle du champ récepteur. Ces cellules sont également sensibles à la courbure du stimulus.

ABSTRACT

In this communication, we describe a method of image segmentation based on the implementation of a general concept of perceptive grouping. The originality of this work lies on the fact the grouping process works simultaneously on several types of information, only using simple notions of similarity-homogeneity, independently of the processed data. In the spatial segmentation application presented here, the grouping acts on orientated-edges-type characteristics, extracted from the picture during a preliminary step.

This work appears as part of advanced studies in the field of moving video pictures (TV pictures) coding and compression, that could lead to the specification of a coding system based on the low-level mechanisms of the visual cortex.

Les traitements effectués par ces cellules contribuent à extraire les informations d'orientation, de courbure et de fréquence spatiale.

D'autres types de cellules captent parallèlement les caractéristiques de couleur, mouvement et relief.

Ainsi, dès les premières aires du cortex visuel [BURN 88], sont effectués des traitements initiaux correspondant à une décomposition de l'image qui fournissent en chaque point un codage multi-canaux et multi-échelles de l'information :

- Représentation parallèle sur canaux séparés de primitives spécifiques (fréquence spatiale, orientation, couleur, mouvement).
- Représentation multi-échelles simultanée de ces primitives, correspondant à différentes résolutions.

Une deuxième étape de combinaison et regroupement des caractéristiques est alors opérée afin d'aboutir à une perception du monde extérieur. Un modèle prenant en compte les mécanismes intervenant lors de cette deuxième phase peut être proposé [ANTO



92]. Reprenant certaines notions du concept de "Gestalt", [REED 90] fait intervenir, pour la mise en oeuvre du regroupement perceptif, des considérations de proximité, similarité, bonne continuation etc... Ce principe a par exemple été utilisé pour regrouper des éléments de contours dans un but de segmentation d'objets [MOHA 92].

Nous en proposons ici une application à la segmentation d'images texturées.

La modélisation s'appuie sur un processus de regroupement des éléments constitutifs de l'image reposant simplement sur l'application itérative de critères de similarité et d'homogénéité, sans faire intervenir de connaissances de plus haut-niveau.

Dans cette modélisation, chaque pixel (x,y) de l'image est affecté d'un vecteur ou colonne (en analogie avec les colonnes corticales du cortex) notée $C(A_1, A_2, \dots, A_n)$ de dimension n, la composante A_j de la colonne C représentant la caractéristique j.

Chaque colonne est soumise à des interactions latérales tendant à renforcer localement la cohésion des informations. C'est au niveau de ces interactions que le regroupement perceptif va intervenir, la modulation de l'activité de chaque colonne en fonction de celle de ses voisines pouvant être exprimée de la manière suivante :

$$C(t) = (\alpha * C(t-1)) + \sum_{i=1,4} (\lambda_i * C_i(t-1))$$

t : pas d'itération, λ_i = fonction $f(C, C_i)$ d'interaction entre 2 colonnes voisines C et C_i .

Critères de Regroupement

Pour introduire une règle de similarité, le facteur $\lambda_i = f(C, C_i)$ doit tenir compte de la corrélation entre les activités des colonnes voisines C et C_i sous la forme d'une fonction comportant deux termes : l'un constant, l'autre croissant selon la similarité des colonnes C et C_i . (Nota bene : la similarité est liée à la distance dans l'espace des caractéristiques entre les colonnes C et C_i). L'activité de la colonne C tend alors :

- vers la moyenne des activités des 4 colonnes voisines dans l'espace des caractéristiques, ce qui correspond à une uniformisation de l'information avec un facteur de pondération égal au premier terme constant de la fonction $f(C, C_i)$.
- vers l'activité de sa plus proche voisine, pour laquelle λ_i sera maximum (action du second terme de $f(C, C_i)$). L'influence de ce deuxième facteur contribue à renforcer les frontières des régions.

Le choix du plus proche voisin dépendra d'autre part d'une contrainte supplémentaire prenant en compte l'homogénéité des informations au voisinage de chacune des colonnes candidates. Complémentairement à la règle locale de similarité qui

ne met en jeu que 2 colonnes, le critère d'homogénéité introduit donc la notion de population de colonnes C_i les plus cohérentes entre elles et voisine de la population de colonnes C. Cette contrainte d'homogénéité favorise significativement l'uniformisation à l'intérieur des régions tout en préservant bien les frontières. En effet, nous avons constaté que la prise en compte de la seule similarité permet effectivement une bonne uniformisation, mais ceci au détriment de la délimitation des frontières qui deviennent floues.

Après quelques itérations, l'algorithme converge et fait apparaître des plages uniformisées.

II • APPLICATION DU REGROUPEMENT EN VUE D'UNE SEGMENTATION DE L'IMAGE

Le principe de regroupement perceptif a été ici modélisé à des fins de segmentation spatiale d'images. Il est appliqué sur des caractéristiques (orientation des contours) extraites dans une phase préliminaire de décomposition de ces images par des opérations de filtrage.

Le système visuel humain extrait l'information de contours par l'intermédiaire de cellules sensibles aux contrastes locaux pour une orientation et une fréquence spatiale privilégiées. Ces cellules peuvent être modélisées par des filtres ayant une réponse impulsionnelle en forme de "chapeau mexicain". Plusieurs variantes existent : les filtres de Gabor, les fonctions DOG (Difference Of Gaussians), les dérivées secondes de Gaussienne. Ils ont par exemple été utilisés pour la discrimination de textures [MALI 90]. Nous avons simulé des cellules de ce type par un jeu de filtres correspondant à des dérivées secondes de Gaussiennes. Ces filtres, caractérisés chacun par une orientation β et une échelle (cf. Fig 1), ont servi à l'extraction des caractéristiques d'orientation des contours.

Ce même type de filtres peut aussi être utilisé pour estimer la courbure de ces contours [MARC 92].

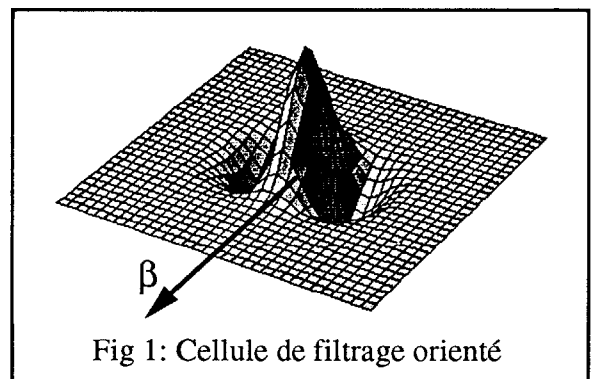


Fig 1: Cellule de filtrage orienté

L'étape de filtrage permet ainsi de générer un ensemble de plans image représentatifs de forts contrastes locaux (contours) caractérisés par leur orientation. Ainsi pour tout point de l'image, chacun de ces plans fournit une des composantes de la colonne C décrite dans la partie I.

Le regroupement perceptif est alors appliqué sur ces caractéristiques d'orientations à des fins de segmentation spatiale d'images. La mise en oeuvre de ce même principe permettrait aussi bien de regrouper des informations de déplacement temporel local en vue de déterminer des ensembles de pixels cohérents au sens du mouvement, et d'approcher ainsi une segmentation orientée "objet".

III • RESULTATS

Le regroupement perceptif, implanté en langage C sur Macintosh Quadra, a été testé sur différents types d'images.

Les filtres utilisés se révèlent être bien adaptés à l'extraction des caractéristiques de contrastes orientés. Le nombre de directions de filtrage a été fixé à 4 (0, $\pi/4$, $\pi/2$, $3\pi/4$) et pourrait être augmenté pour plus de précision. Le filtrage n'est effectué qu'à une seule résolution, le choix de la variance (taille du filtre) doit donc être adapté à la dimension des motifs formant les textures.

Lors de la phase de regroupement, la détermination du plus proche voisin va faire appel à 2 mesures :

- Une mesure de similarité, entre les populations de colonnes voisines, calculée simplement comme la distance euclidienne dans l'espace des caractéristiques.
- Une mesure d'homogénéité qui reflètera la variation d'activité de la population de colonnes. Comme critère d'homogénéité, nous avons utilisé la variance calculée pour chaque caractéristique sur l'ensemble de la population. L'homogénéité permet ainsi de tenir compte du "contexte" pour le choix du plus proche voisin.

Après une dizaine d'itérations, le processus de regroupement se stabilise et fait apparaître des plages uniformisées.

Afin d'extraire les frontières de ces régions, un dernier traitement effectuée à partir des images obtenues par le regroupement, une segmentation sur un critère de luminance. Cette méthode issue de la théorie des graphes [MORR 86] met aussi en jeu les interactions latérales entre chaque pixel en leur associant une fonction de coût basée sur un gradient de luminance. En premier lieu, on construit un chemin passant par tous les pixels de telle façon que le coût de ce chemin, défini comme la somme des coûts de chaque lien, soit minimal. La segmentation consiste

ensuite à "couper" les liens les plus forts du graphe établi. Tous les pixels situés entre 2 coupures vont former une région et seront affectés d'une même étiquette, on obtient ainsi des plages parfaitement homogènes dont les frontières sont faciles à localiser.

Cette étape, indépendante et postérieure au regroupement, n'intervient que pour la détermination des frontières entre régions. Des techniques plus simples basées sur un calcul de disparité locale (opérateur gradient par exemple) peuvent aussi être envisagées, mais ne fournissent pas des frontières continues et fermées comme c'est le cas ici.

Du fait de la nature des caractéristiques utilisées, le regroupement sera plus performant pour des textures composées de motifs simples pouvant être discriminés par leur orientation. Par exemple, utilisée sur l'image 1, la segmentation isole une région pour chaque plan issu du regroupement et représentatif d'une orientation β . Sur les images résultats 1,2,3, figurent l'ensemble des frontières de ces régions (traits blancs), superposées à l'image originale.

Néanmoins, le processus s'applique aussi sur des textures plus complexes comme celles de l'image 2 formée de textures naturelles (canevas, raffia, bois et laine). La région du canevas est même segmentée en plusieurs bandes en raison de la structure en "chevrons" formant des zones verticales où l'orientation du motif est constante.

Sur une image réelle (image 3), le regroupement réussit à détacher des régions (buste du personnage, pantin) du fond de l'image formé par les structures verticales de la tapisserie.

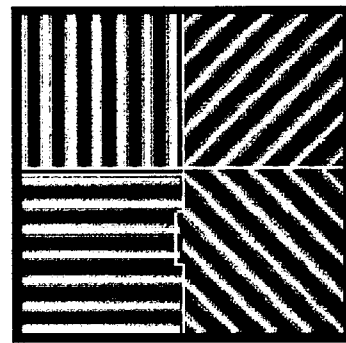


Image 1: Segmentation de textures artificielles

IV • CONCLUSION

Dans cet article, nous avons présenté une méthode de segmentation appliquant le principe de regroupement perceptif sur des caractéristiques d'images. La force des interactions latérales entre les populations de cellules représentatives de ces caractéristiques provoque un regroupement, suivant un critère de similarité et d'homogénéité, des



informations issues d'une étape préliminaire de décomposition de l'image par filtrage de type Gaussien.

Nous comptons appliquer ce principe sur des séquences d'images TV dans une perspective de codage et de compression. En effet une perception de l'image à plus haut niveau (ex : régions, plages texturées, zones statiques ou en mouvement) permettrait d'accéder à une représentation plus symbolique qui réduirait notablement la quantité d'informations à transmettre.

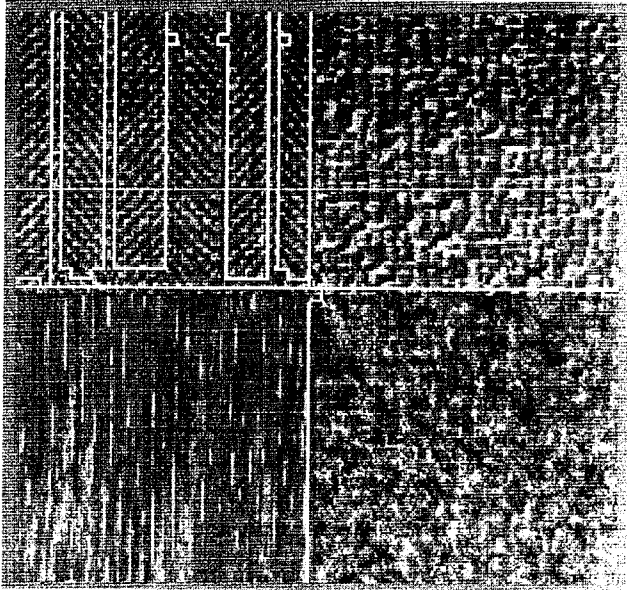


Image 2: segmentation de textures naturelles

Remerciements

Nous tenons à remercier Monsieur le Professeur Y.BURNOD, ainsi que Messieurs J.L.ANTON, F.GUYOT et P.MARCHAL, de l'Institut des Neurosciences de la Vision au CNRS pour leurs précieux conseils et dont les travaux ont servi de base à l'étude présentée ici.



Image 3: segmentation permettant d'isoler les objets du fond de l'image

BIBLIOGRAPHIE

- [ANTO 92] J.L. ANTON - "Modélisation des Bases Neurales du Regroupement Perceptif - Discrimination de Textures" - Rapport de DEA de Sciences Cognitives- Université PARIS VI, 1992.
- [BURN 88] Y. BURNOD - "An Adaptive Neural Network - The Cerebral Cortex" - Livre 370 pages - Masson, Paris 1988.
- [MARC 92] P.MARCHAL - "Populations of end-stopped neurons can efficiently guide ocular scanning of natural images : a neural net model." - soumis à Biological Cybernetics.

- [MALI 90] J.MALIK, P. PERONA - "Preattentive Texture Discrimination with Early Vision Mechanisms" - J.Opt.Soc.Am. A, Vol 7, No 5, 1990.
- [MOHA 92] R.MOHAN - "Perceptual Organization for Scene Segmentation and Description" - IEEE PAMI, vol 14, No 5, 1992.
- [MORR 86] O.J. MORRIS "Graph theorie for image analysis: an approach based on the shortest spanning tree." IEE Proceeding, vol 133, n°2, 1986
- [ORBA 84] G.A.ORBAN - "Neuronal Operations in the Visual Cortex" - Springer Verlag 1984.
- [REED 90] T.R. REED, H. WECHSLER - "Segmentation of Texture Images and Gestalt Organization using Spatial/Spatial-Frequency Representations" - IEEE PAMI, Vol 12, No 1, 1990.