



HAL
open science

étude comparative de méthodes de segmentation dans une approche orientée indexation.

Jérôme da Rugna, Hubert Konik

► **To cite this version:**

Jérôme da Rugna, Hubert Konik. étude comparative de méthodes de segmentation dans une approche orientée indexation.. RFIA 2004, Jan 2004, Toulouse, France. pp.13-20. ujm-00124898

HAL Id: ujm-00124898

<https://ujm.hal.science/ujm-00124898>

Submitted on 16 Jan 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Étude comparative de méthodes de segmentation dans une approche orientée indexation

Comparison of different segmentation algorithms for indexation oriented approach

J. Da Rugna¹

H. Konik¹

¹ Université Jean Monnet - Laboratoire LIGIV EA 3070

10 rue Barrouin - Bât. E - 42000 SAINT-ÉTIENNE
darugna@ligiv.org

Résumé

Cet article présente une étude objective et quantitative des méthodes de segmentation pour la recherche d'images par le contenu. Une méthode pyramidale de segmentation couleur par propagation d'étiquettes est introduite avec différentes possibilités d'initialisation des germes. Une nouvelle approche d'évaluation dans un contexte orienté indexation est proposée. Des descripteurs sont définis pour valider la performance de stabilité. Des résultats sur une banque d'images test conséquente sont présentés.

Mots Clef

évaluation de performance, segmentation couleur, segmentation multirésolution, recherche par le contenu image.

Abstract

This paper presents an objective and quantitative study of segmentation methods for content-based image retrieval. A pyramidal algorithm of color segmentation by labels propagation is introduced with different initialisation possibilities of the germs. A new approach evaluation in an indexation oriented context is proposed. Descriptors are defined to validate stability performance. Results on a large collection of test images are presented.

Keywords

performance evaluation, color segmentation, multiresolution segmentation, content-based image retrieval.

1 Introduction

L'objectif de cet article est de proposer une étude comparative de quelques méthodes de segmentation en vue de leur utilisation en indexation d'images. Cette thématique a en effet pris un essor conséquent ces dernières années mais l'évaluation objective des différents travaux proposés reste

insuffisante ([14]). En fait, parler d'indexation d'images sous-entend habituellement de rechercher des similarités entre images... Or, la notion de similarité espérée par l'utilisateur potentiel est bien difficilement appréhendable par simple recherche par le contenu, cette limite des moteurs actuels freinant encore leur utilisation dans une optique généraliste. En effet, un réel fossé existe encore entre l'attente sémantique de l'utilisateur et les capacités actuelles de l'ensemble des moteurs disponibles ([19]). Notre intention n'est alors pas d'entrer dans le débat sur l'opportunité même de parler de sémantique dans une image - toute interprétation restant relative- mais plutôt d'introduire une critique des traitements bas niveau mis en jeu la plupart du temps. Les travaux présentés ici se situent ainsi en amont de la phase d'indexation elle-même sur une évaluation critique de la robustesse des techniques utilisées.

Habituellement, notamment lorsque les images considérées sont dites de scène, la similarité naît de la corrélation entre les vecteurs d'attributs (texture, couleur, forme) attachés aux différentes régions d'une image. Par conséquent, la plupart des approches existantes restent locales au sens où elles reposent sur une étape de segmentation préalable. Dans ce contexte, de nombreux travaux ont conduit à l'évaluation des méthodes de segmentation elles-mêmes ([21], [22], [2], [18]). Néanmoins, plutôt que de valider le choix de telle ou telle méthode avec une attente "sémantique", au sens où les objets doivent être le plus proches possible d'une segmentation visuelle, nos travaux consistent plus précisément à valider telle ou telle option selon sa stabilité au contexte environnant. En effet, si l'objectif à terme est d'assumer une tâche de comparaison, la première propriété souhaitable d'une méthode de segmentation est de rester stable quel que soit le contexte. Ainsi, même si le résultat semble insuffisant par rapport à une approche visuelle, plus les

résultats obtenus seront invariants au contexte et plus la méthode sera jugée potentiellement efficace dans une optique indexation. Dans ce cadre, nous avons alors construit une base de référence en sélectionnant quelques objets que nous avons plongés dans une base d'images généraliste. Ensuite, chaque résultat sera comparé à une segmentation de référence obtenue en immergeant chaque objet dans un fond complexe homogène similaire.

Parallèlement à cette étude comparative, nous avons introduit une méthode de segmentation multirésolution basée sur un processus top-down de propagation d'étiquettes à travers les niveaux d'une pyramide couleur gaussienne. Initialisée au sommet par différentes versions d'extraction de germes potentiels, un parcours descendant à travers les différentes résolutions engendre une image segmentée étiquetée. La présentation de cette méthode constituera l'objectif de la première partie. Ensuite, dans une seconde partie, nous listerons l'ensemble des indicateurs de stabilité introduits pour réaliser une analyse critique des méthodes de segmentation dans une optique indexation. Tout en précisant la construction des images test, une troisième partie permettra d'évaluer quantitativement quelques méthodes sélectionnées pour leur diversité.

2 Méthode top-down de segmentation par propagation d'étiquettes

La segmentation d'images est une étape essentielle de bas niveau qui consiste à définir une partition d'une image en régions visuellement distinctes et uniformes selon des critères de texture et/ou de couleur par exemple. De cette simple définition naît la problématique sous-jacente de l'évaluation théorique de ces méthodes... De plus, avec la diversité des choix pour les prédicats d'homogénéité, de nombreuses méthodes ont depuis des décennies vu le jour ([15]). Dans l'ouvrage [3], les auteurs confirment qu'aucune méthode ne semble prévaloir sur les autres, chacune ayant son domaine de prédilection. En absence de méthode universelle, il est classique de retrouver les différentes approches classifiées en quatre thèmes : clustering, approches contours, approches régions et méthodes hybrides. Qui plus est, dans un contexte de recherche par le contenu image, la rapidité et la polyvalence des méthodes est souhaitable. Ces hypothèses, ainsi que la capacité de l'outil à répondre aux quatre classes sus-citées, ont conduit à exploiter les techniques multirésolution ([7], [6], [16], [8]), quand bien même les critiques relatives à l'outil lui-même sont référencées ([1]).

Notre procédure de segmentation va alors reposer sur le schéma présenté figure 1, donnant naissance à trois étapes distinctes.

2.1 Étape 1 : Construction de la pyramide

Tout d'abord, il est nécessaire de construire la pyramide gaussienne, selon l'approche de Burt, qui se définit comme une suite d'images de résolutions réduites. Plus précisément, un noyau avec recouvrement est utilisé ce qui revient

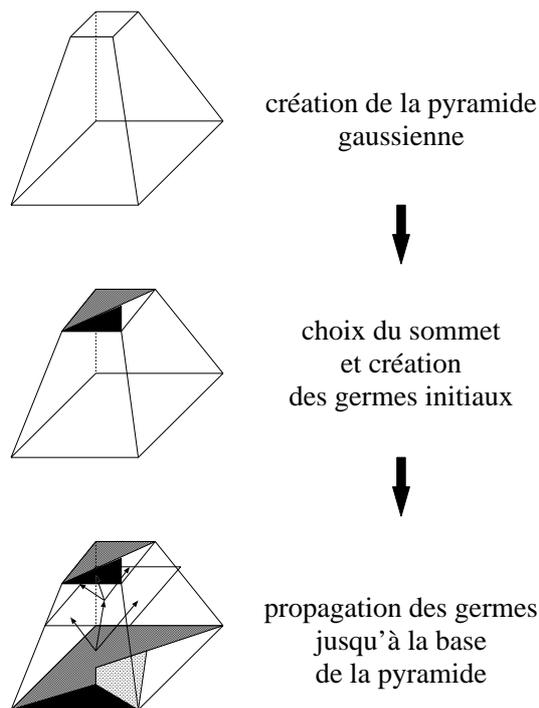


FIG. 1 – Procédure de segmentation.

à construire chaque pixel du niveau courant par moyennage pondéré de 16 pixels - ses fils. Parallèlement, chaque pixel du niveau courant intervient dans la construction de 4 pixels du niveau supérieur - ses pères. Par la nature linéaire de cette construction, l'espace *RGB* est utilisé ([10]). Par essence même cet outil permet d'évaluer conjointement l'information spatiale tout autant que l'information colorimétrique. Ensuite, les hauts niveaux offrent une vision globale de l'image conduisant à une tâche d'extraction des régions principales plus aisée, alors que les bas niveaux comportent seuls l'ensemble des détails, nécessaires à une description précise de chacune d'elle.

Pour s'adapter à des images de taille variable, la construction peut "perdre" une ligne et/ou une colonne à chaque nouveau niveau (de résolution divisée par 2 dans les deux directions par rapport au précédent) sachant qu'elle n'est jamais poursuivie jusqu'à des niveaux de résolution prohibitifs où l'intégralité de l'information initiale est de toute évidence perdue.

2.2 Étape 2 : Choix des germes

Maintenant que la pyramide est construite, l'objectif est d'extraire les germes d'un processus hiérarchique descendant. Initialement, le nombre de germes potentiel équivaut au nombre de pixels du sommet choisi.

Nous avons fixé un niveau initial situé à deux niveaux en-dessous du niveau maximal de construction, engendrant par exemple des images de taille 7×11 pour des images réelles de taille 249×378 . Ce choix repose également sur une volonté d'obtention d'un nombre de régions finales pas

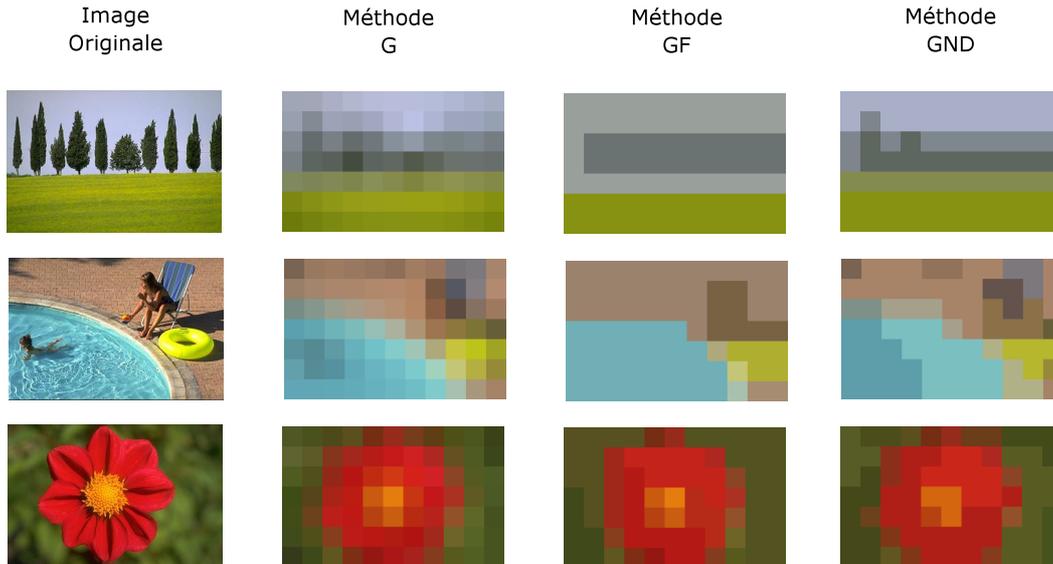


FIG. 2 – Construction des germes du processus de segmentation hiérarchique.

trop important. En effet, dans un contexte applicatif de recherche par le contenu dans de grandes bases de données image, un nombre conséquent de régions obtenues peut être inadapté et devenir prohibitif ([11]).

À partir de là, différentes approches sont possibles pour réaliser un premier étiquetage du sommet :

- *méthode G* :
Chaque pixel devient germe et le nombre de régions initial dépend de la seule résolution du niveau sommet ([17]).
- *méthode GF* :
Chaque pixel devient germe mais une étape de fusion permet à certains germes de se regrouper en cas de forte ressemblance. En effet, il semble logique d'agréger les germes d'une région commune. Le critère de proximité peut classiquement reposer sur une distance faible calculée dans l'espace RGB ou l'espace HSV ([12]). Perceptuellement plus uniformes, les espaces comme $L^*a^*b^*$ semblent encore plus difficiles à être utilisés puisque généralement l'illuminant ayant servi à éclairer la scène acquise est inconnu et, dans le cas de bases généralistes, peut être très varié. Dans ce cas, le nombre de germes est contraint par la taille du sommet et par le seuil de proximité.
- *méthode GND* :
Utilisation d'un algorithme de classification par nuées dynamiques ou k-means en fixant le nombre de germes souhaités. Afin de couvrir au maximum l'information colorimétrique de l'image, pour également améliorer la robustesse de la méthode, les germes initiaux du processus de classification occupent au mieux le nuage couleur dans le cube RGB . Pour cela, un découpage

en sous-cubes est utilisé. Dans ce dernier cas, le nombre de germes est indépendant de la taille du niveau initial mais repose sur un choix empirique. Notons que cette borne supérieure peut ne pas être atteinte selon la densité colorimétrique du nuage de points.

La figure 2 présente ces trois possibilités sur des images de scène variées.

2.3 Étape 3 : Processus top-down d'agrégation

À partir du sommet, représentation grossière de l'image initiale, il est nécessaire d'emprunter les liens de descendance pour appréhender les relations spatiales et les similarités entre éléments. Les relations spatiales sont données d'une part par les notions de père et de fils présentées précédemment entre les niveaux et par les notions de voisinage classiques V_4 ou V_8 à un même niveau. Les similarités sont dans notre cas obtenues en comparant la proximité colorimétrique entre un fils et ses pères potentiels. Dans le but d'élire le meilleur représentant, un pixel du niveau courant aura tendance à recevoir l'étiquette de son père le plus similaire.

Néanmoins, dans une approche région, on va mettre à jour à chaque niveau des descripteurs pour chacune des régions définies. À partir de là, un pixel va alors recevoir l'étiquette de la région la plus proche à laquelle appartient un de ses pères. De plus, les niveaux élevés représentent une version grossière de l'image. Du coup, il se peut que par moyennage des régions distinctes sur la base donnée naissance à une région commune à un niveau de faible résolution. Par conséquent, il faut donner la possibilité à une région de se scinder lors de la descente. Pour se faire, lorsqu'un fils va être jugé trop éloigné de ses pères potentiels, il va donner naissance à une nouvelle région. Malgré tout, en absence

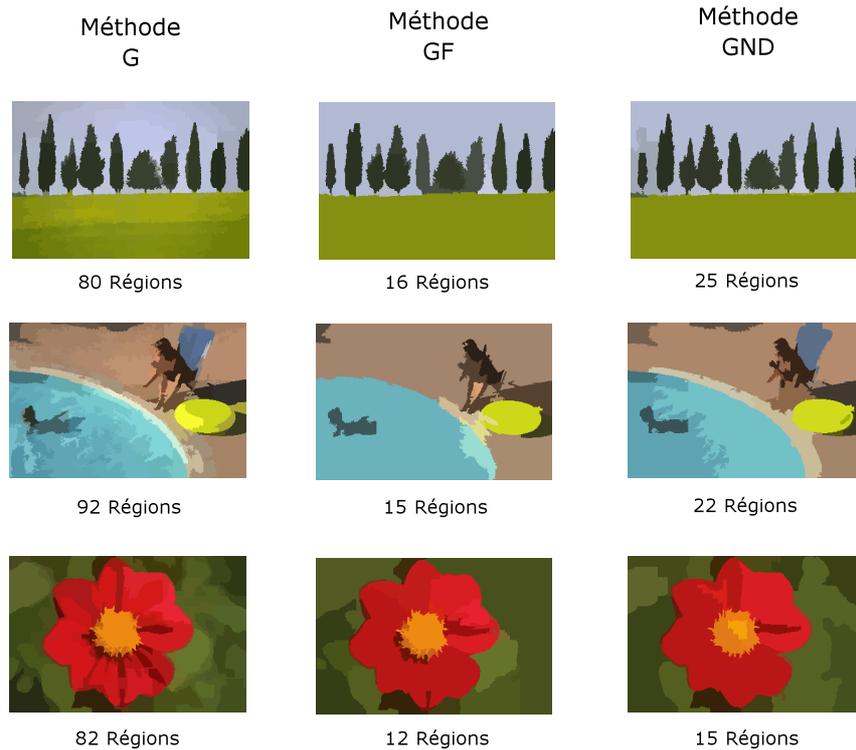


FIG. 3 – Exemples de segmentations obtenues selon l'initialisation des germes.

d'un post traitement de type fusion à chaque niveau de segmentation, plutôt que de partir isolé, il va d'abord chercher à se rapprocher de ses frères - pixels voisins du même niveau - pour éventuellement former des régions communes. La figure 4 résume alors cette procédure descendante itérée depuis le niveau situé juste en-dessous du sommet jusqu'à la base de la pyramide.

En définitive, présentons quelques résultats de cette segmentation sur les images dont les sommets ont déjà été présentés figure 2. La figure 3 présente ces résultats, où chaque région reçoit la valeur moyenne couleur calculée sur l'ensemble de ses pixels.

3 Méthodes d'évaluation de la robustesse de la segmentation

Habituellement, il est d'usage d'évaluer la qualité d'une segmentation en la comparant à une segmentation idéale bien souvent construite manuellement ([21]). Dans ce cadre, la méthode sera d'autant meilleure qu'elle optimisera l'adéquation entre les résultats automatiques et les résultats espérés.

Néanmoins, il semble délicat de suivre ce processus dans une optique orientée recherche par le contenu, puisqu'il est illusoire d'espérer une segmentation "sémantique" capable de s'adapter quelle que soit la recherche envisagée. De même, le coté subjectif des résultats attendus peuvent ne pas correspondre selon la requête d'une part et selon

l'utilisateur d'autre part.

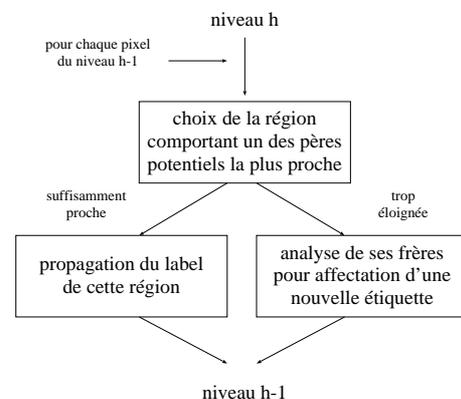


FIG. 4 – Processus de propagation des étiquettes entre les niveaux.

Par contre, pour pouvoir envisager de retrouver automatiquement certains objets particuliers d'une image à l'autre, il est préférable de donner crédit à une méthode stable eu égard au contexte environnant dans lequel apparaît l'objet de référence. Du coup, nous avons introduit un processus d'évaluation en sélectionnant plusieurs objets de référence que nous avons plongés aléatoirement au coeur d'une base généraliste (figure 5).

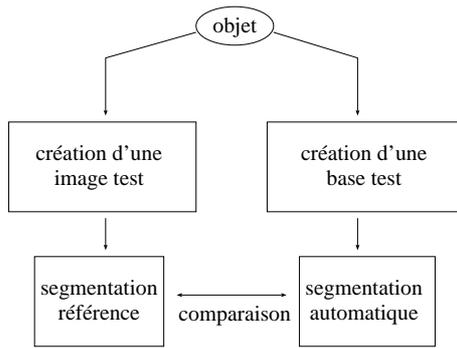


FIG. 5 – Création d'un protocole d'évaluation objective de la segmentation.

3.1 Création d'une base test

Pour créer une base de référence, nous avons sélectionné certains objets choisis pour leur diversité en termes de couleur (jaune, rouge, marron...), de texture (citron, perroquet...) et de forme (rond, allongé, irrégulier...). De plus, pour affiner la performance de stabilité, chaque objet existe dans plusieurs échelles (depuis 75 jusqu'à 300 pixels de largeur) et, précisons-le, n'est pas forcément segmentable en une unique région homogène.

Ensuite, de façon aléatoire, nous avons immergé chacun d'entre eux dans différentes images de scène diverses et variées en prenant bien évidemment soin de conserver la position du rectangle englobant l'objet dans tous les cas. La base ainsi construite inclut 24 objets de référence et 4500 images de scènes (la taille des images, variant de 250x380 à 1000x1000).

La figure 6 présente des images de quelques objets apparaissant au coeur de différentes scènes.



FIG. 6 – Quelques images de la base de référence.

De plus, il est nécessaire de construire pour chaque objet une segmentation de référence à chaque échelle que l'on viendra comparer aux résultats obtenus. Pour se faire, chaque objet a été immergé au coeur d'un même fond colorimétriquement dense (image d'une Macbeth DC Color Checker) afin de ne privilégier aucune approche de segmentation par rapport à une autre, et de ne créer aucun biais par rapport à l'immersion dans les images de scènes.

3.2 Descripteurs de similarité

Maintenant, il faut introduire des mesures de ressemblance entre deux segmentations. Dans notre cas, l'une sera celle de référence de l'objet en cours alors que l'autre sera celle obtenue sur l'image courante. Classiquement, les approches en usage reposent sur deux aspects : l'un colorimétrique, notamment pour évaluer la qualité d'une compression d'image par exemple, et l'autre plus morphométrique mesurant le taux de recouvrement entre régions. Si un objet est découpé en deux zones, encore faut-il qu'elles soient toujours physiquement positionnées de façon stable.

Dans ce dernier cas, nous allons utiliser deux quantificateurs : le premier mesurant la pertinence des régions obtenues au sens où elles ne doivent pas trop empiéter sur le contexte environnant, le second mesurant la stabilité des régions obtenues.

Le premier descripteur, dénommé *coefficient de mélange*, mesure le pourcentage de l'objet mélangé au fond. En effet, une trop grande dispersion sur les régions voisines risquent de limiter la reconnaissance de l'objet. Néanmoins, on se donne une certaine souplesse, et si l'objet sort de moins de $n\%$ du support idéal, il sera considéré comme valide (n bien évidemment assez faible).

Si $X = \bigcup_{i=1}^m X_i$ est une segmentation de l'objet dans l'image

référence et $Y = \bigcup_{j=1}^n Y_j$ une segmentation de l'objet dans

l'image de scène, le coefficient de mélange vaut :

$$CM = \frac{\sum_{j=1}^n (Card(Y_j \cap X) \times \delta_j)}{Card(X)}$$

avec

$$\delta_j = \begin{cases} 1 & \text{si } \frac{Card(Y_j \cap \bar{X})}{Card(Y_j)} \geq n \\ 0 & \text{sinon} \end{cases}$$

Pour le second descripteur, les approches classiques ([21]) reposent sur une matrice de confusion où chaque entrée C_{ij} représente le nombre de pixels de classe j classifiés en tant que classe i par l'algorithme de segmentation. Suivant la démarche adoptée en [3], nous allons plutôt utiliser la *mesure de Vinet* reposant sur la détermination des couples de régions assurant un recouvrement maximum entre les deux segmentations et la caractérisation de la dissimilarité par la proportion de pixels ne participant pas à ce recouvrement. Posant N comme le nombre de pixels total de l'objet initial considéré, le principe est alors le suivant :

– définition de la table de superposition T par :

$$T(i,j) = Card(X_i \cap Y_j) \quad \forall i = 1..m \\ \forall j = 1..n$$

- recherche du couple de régions (X_{i_1}, Y_{j_1}) de recouvrement maximal tel que :

$$T(i_1, j_1) \geq T(i, j) \quad \forall i, j$$

On note $c_1 = T(i_1, j_1)$.

- itération jusqu'à obtenir (X_{i_k}, Y_{j_k}) , avec $k = \min(m, n)$, tel que :

$$T(i_k, j_k) \geq T(i, j) \quad \begin{array}{l} \forall i \neq i_1, i_2, \dots, i_{k-1} \\ \forall j \neq j_1, j_2, \dots, j_{k-1} \end{array}$$

- calcul de la mesure de dissimilarité par :

$$d_{Vinet} = \frac{N - \sum_{i=1}^k c_i}{N}$$

Présentons un exemple simple de cette mesure de Vinet en comparant deux segmentations V et W avec une segmentation témoin notée R , présentées figure 7.

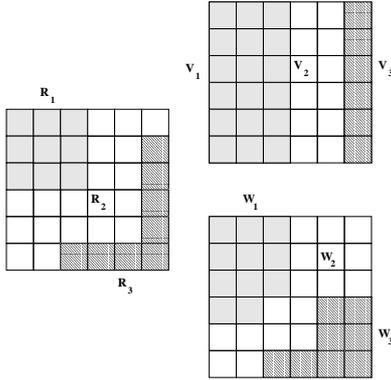


FIG. 7 – Exemples de segmentation pour définir la mesure de Vinet.

Pour cet exemple, on obtient :

$$\begin{cases} d_{Vinet}(R, V) & = & 0.33 \\ d_{Vinet}(R, W) & = & 0.16 \end{cases}$$

Visuellement, la segmentation W semble en effet plus proche de la segmentation R en terme morphométrique que ne l'est la segmentation V .

Le dernier descripteur retenu mesure l'adéquation colorimétrique entre la segmentation obtenue et la segmentation de référence. En effet, selon le contexte, il est parfois raisonnable d'envisager que les régions peuvent s'étendre sur les pixels voisins suffisamment proches pour être agrégés. Ce paramètre mesure si l'écart commis est influent ou non. Nous avons opté pour le classique $RMSE$.

Par définition, on note \bar{Y}_k la valeur moyenne RGB de la région segmentée Y_k . Soit $r(i, j)$ la région à laquelle appartient le pixel (i, j) , de couleur $f(i, j)$. On a alors :

$$RMSE = \frac{1}{Card(X)} \sum_{(i,j) \in X} (f(i,j) - \bar{Y}_{r(i,j)})^2$$

4 Résultats et comparaison

Une fois cette création d'une banque d'images test, l'objectif actuel est de comparer certaines méthodes de segmentation selon le protocole précédemment décrit. Bien évidemment, avec la multitude de méthodes de segmentation existantes, cette étude reste certes non exhaustive. Commençons donc par brièvement présenter les méthodes retenues.

4.1 Différentes méthodes de segmentation

- Approche clustering
Les approches basées sur un pré-clustering d'un espace de caractéristiques sont souvent utilisées pour le bon compromis entre efficacité et rapidité. Une illustration basée sur des clusters flous [13] a été choisie dans ce test car elle permet d'ajouter un aspect spatial fort intéressant dans le contexte de notre étude.
- Approche couleur via l'algorithme mean shift
Les auteurs proposent dans [4] d'utiliser dans l'espace des couleurs un algorithme classique de reconnaissance de formes : le mean shift. Appliquée soit en vue d'une quantification soit en vue d'une segmentation, cette méthode fournit des résultats qui visuellement paraissent performants, mais dont la tendance est de générer un nombre important de régions.
- Approche morphologique
Les méthodes basées sur des lignes de bassins versants [20] sont incontournables dans de nombreux domaines de traitement d'images. Via un gradient représentant les contours (par des valeurs élevées), cet algorithme effectue une croissance de région autour des minima locaux. Créant initialement beaucoup de régions, nous avons opté pour une version améliorée de cet algorithme [5]. Ainsi, il est possible de contrôler le nombre de bassins résultants.
- Approche multirésolution
Quant à l'approche multirésolution, notre choix s'est porté sur la méthode précédemment décrite, en utilisant les différents types de germes.

4.2 Résultats

Dans le but d'une robustesse des résultats obtenus, chaque objet, pour chaque résolution, a été inséré dans près de 4500 images, renforçant d'autant la validité de l'étude. On note dans les tableaux suivants Cluster l'approche clustering, WaterShed l'approche morphologique, MeanShift l'approche couleur et Pyra l'approche pyramidale. Chaque méthode a été paramétrée de façon à obtenir les meilleurs résultats possibles.

Pyra G	Pyra GND	MeanShift	Cluster	WaterShed
22.5	5	28.5	4.6	8.32

TAB. 1 – Nombre moyen de régions sur l'ensemble des objets.

Méthode	CM	d_{Vinet}	$RMSE$
Pyra G	0.12	0.63	0.0022
Pyra GND	0.17	0.52	0.0026
MeanShift	0.09	0.62	0.0019
Cluster	0.46	0.68	0.0017
WaterShed	0.13	0.53	0.0030

TAB. 2 – Objets texturés.

La première constatation est la relative faiblesse des méthodes dans leur ensemble au niveau des résultats, même si visuellement les segmentations semblent convaincantes. En effet, une grande partie de l'objet est généralement mélangée au contexte, malgré le principe de création (collage d'un objet au milieu d'une image) engendrant en général des objets contrastés par rapport au fond. Quant à la mesure de Vinet, on voit clairement qu'en moyenne on ne peut apparier qu'une petite moitié d'un objet segmenté avec la segmentation référence. On notera à ce sujet que la référence telle qu'elle est choisie correspond à la référence la plus objective, et celle donnant les meilleurs résultats. De plus, l'expérience montre qu'il n'y a pas de régions privilégiées, c'est-à-dire qu'il n'est pas possible d'affirmer que la valeur de la mesure de Vinet correspond à l'appariement de la même région sur toute la base (écart type élevé).

La deuxième constatation est la relation ambiguë entre les différentes mesures. Cela semble naturel de penser qu'un meilleur mélange engendre une meilleure mesure de Vinet ou un RMSE plus faible. L'analyse des résultats montre en effet que suivant la méthode utilisée, le lien entre les différentes mesures est très variable. L'approche clustering de l'espace couleur, par exemple, même si son coefficient de mélange est élevé, aura un RMSE très faible, l'objet étant "mêlé" avec un fond colorimétriquement semblable. L'al-

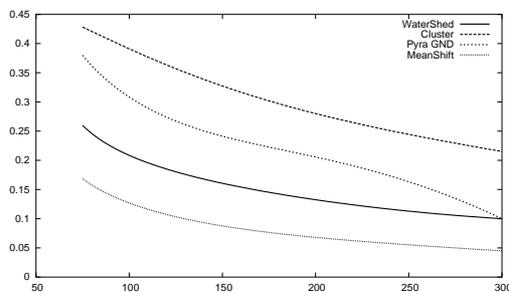


FIG. 8 – Influence de la taille de l'objet sur le coefficient de mélange.

Méthode	CM	d_{Vinet}	$RMSE$
Pyra G	0.20	0.65	0.003
Pyra GND	0.30	0.49	0.003
MeanShift	0.20	0.50	0.0034
Cluster	0.33	0.62	0.0030
WaterShed	0.39	0.52	0.0057

TAB. 3 – Objets colorés multitexturés.

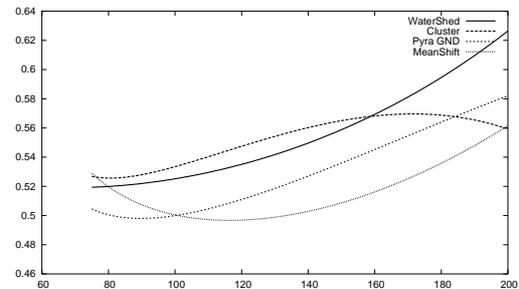


FIG. 9 – Influence de la taille de l'objet sur la mesure de Vinet.

gorithme mean shift, d'un autre côté, est le plus performant pour extraire l'objet de son contexte mais, du fait d'un grand nombre de régions, réalise de moins bons résultats pour la mesure de Vinet.

Les figures 8 et 9 présentent l'influence de la taille de l'objet sur respectivement le coefficient de mélange et la mesure de Vinet. Comme attendu, plus l'objet sera grand, moins la probabilité de mélange au fond sera forte. Par contre, la segmentation a tendance à perdre en stabilité eu égard au nombre croissant de régions extraites. Une fois encore, aucune méthode ne semble véritablement émerger par rapport aux autres.

La méthode multirésolution que nous avons proposée donne en définitive des résultats homogènes et convaincants, notamment dans le cas de germes GND. Suivant le type de germes, cela semble d'ailleurs très logique, la méthode aura tendance, soit à bien épouser la forme de l'objet et donc à ne pas le mélanger au contexte, soit à conduire à une segmentation plus grossière, qui aura tendance à moins dissocier l'objet du fond, mais sera en revanche plus stable.

En conclusion, rejoignant les résultats de [3], il n'y a pas à encore de méthodes génériques ou véritablement meilleures que les autres. Néanmoins, via ce protocole de test, il devient par exemple possible de choisir efficacement une méthode en fonction des objets et des descripteurs utilisés pour finaliser la recherche des images similaires.

5 Perspectives

L'évaluation objective des performances réelles des systèmes de recherche par le contenu est un problème crucial. En effet, la majorité des évaluations récentes se résume la plupart du temps à afficher des planches de résultats d'une

ou plusieurs requêtes. Néanmoins, cela ne constitue en rien un procédé efficace de validation réelle de telle ou telle approche. Or, l'essor applicatif de ces outils, presque inévitable à la vue de l'explosion du tout numérique, ne pourra que se reposer sur des évaluations précises et objectives de leur robustesse.

Dans ce contexte, et dans une moindre échelle, nous avons essayé de proposer un protocole d'évaluation objectif de l'outil habituellement situé en amont de la phase de recherche elle-même, à savoir la segmentation. En effet, cette méthode est antérieure à toute recherche de similarités. Certaines méthodes couvrant au maximum la gamme des techniques en usage (clustering, watershed, algorithme mean shift, processus multirésolution) ont été testées et critiquées selon ce protocole d'évaluation.

Maintenant, nos développements futurs devront être menés sur la pertinence même parfois de segmenter les images. En effet, dans un contexte généraliste, certaines images apparaissent inévitablement bien illusoire à segmenter, soit parce que la complexité est trop importante, soit parce qu'au contraire l'homogénéité est telle que l'image forme une unique région. L'objectif de nos développements est donc de définir un continuum sur les images depuis les images non segmentables jusqu'aux images fortement segmentables en passant par les catégories intermédiaires... selon une approche similaire à celle développée dans [9] pour la thématique particulière de la texture. Avec cette description, en fonction de l'image cible considérée - dans un contexte recherche par l'exemple - la restriction aux seuls descripteurs pertinents devrait permettre d'améliorer les résultats de la requête.

Références

- [1] M. Bister, J. Cornelis, A. Rosenfeld, A critical view of pyramid segmentation algorithms, *Pattern Recognition Letters*, Vol. 11, pp. 802-809, 1990.
- [2] M. Borsotti, P. Campadelli, R. Schettini, Quantitative evaluation of color image segmentation results, *Pattern Recognition Letters*, Vol. 19, pp. 741-747, 1998.
- [3] J.P. Cocquerez, S. Philipp, *Analyse d'images: filtrage et segmentation*, Éditeur Masson, 1995.
- [4] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, *IEEE Trans. Pattern Analysis Machine Intelligence*, Vol. 24, pp. 438-445, 2001.
- [5] M.C. de Andrade, G. Bertrand, A.A. Araújo, Segmentation of microscopic images by flooding simulation: A catchment basins merging algorithm, *Proc. of the IST/SPIE Symp. On Electronic Imaging*, Vol. 8, 1997.
- [6] A. Del Bimbo, M. Mugnaini, P. Pala, F. Turco, Visual querying by colour perceptive regions, *Pattern Recognition*, Vol. 31, pp. 1241-1253, 1998.
- [7] Y. Deng, B. Manjunath, Unsupervised segmentation of color-texture regions in images and video, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 800-810, 2001.
- [8] C. Fuh, S.W. Cho, K. Essig, Hierarchical color image region segmentation for content-based image retrieval system, *IEEE Transactions on Image Processing*, Vol. 9, pp. 156-162, 2000.
- [9] K. Karu, A.K. Jain, R.M. Bolle, Is there any texture in an image?, *Pattern Recognition*, Vol. 29, pp. 1437-1446, 1996.
- [10] H. Konik, V. Lozano, B. Laget, Color pyramids for image processing, *Journal of Imaging Science and Technology*, Vol. 40, pp. 535-542, 1996.
- [11] H.F. Lau, M.D. Levine, Finding a small number of regions in an image using low-level features, *Pattern Recognition*, Vol. 35, pp. 2323-2339, 2002.
- [12] M.S. Lew, *Principles of Visual Information Retrieval*, Springer-Verlag, London, 2001.
- [13] A.W.C. Liew, S.H. Leung, W.H. Lau, Fuzzy image clustering incorporating spatial continuity, *Vision, Image and Signal Processing*, Vol. 147, pp. 185-192, 2000.
- [14] H. Muller, W. Muller, D. McG. Squire, S. Marchand-Maillet, T. Pun, Performance evaluation in content-based image retrieval: overview and proposals, *Pattern Recognition Letters*, Vol. 22, pp. 593-601, 2001.
- [15] N.R. Pal, S.K. Pal, A review on image segmentation techniques, *Pattern Recognition*, Vol. 26, pp. 1277-1294, 1993.
- [16] D. Prewer, L. Kitchen, Soft image segmentation by weighted linked pyramid, *Pattern Recognition Letters*, Vol. 22, pp. 123-132, 2001.
- [17] M.R. Rezaee, P.M.J. van der Zwet, B.P.F. Lelieveldt, R.J. van der Geest, J.H.C. Reiber, A multiresolution image segmentation technique based on pyramidal segmentation and fuzzy clustering, *IEEE Transactions on Image Processing*, Vol. 9, pp. 1238-1248, 2000.
- [18] R. Roman-Roldan, J.F. Gomez-Lopera, C. Atae-Allah, J. Martinez-Aroza, P.L. Luque-Escamilla, A measure of quality for evaluating methods of segmentation and edge detection, *Pattern Recognition*, Vol. 34, pp. 969-980, 2001.
- [19] S. Santini, *Exploratory Image Databases: Content-Based Retrieval*, Academic Press, San Diego, 2001.
- [20] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 13, pp. 583-598, 1991.
- [21] Y.J. Zhang, A survey on evaluation methods for image segmentation, *Pattern Recognition*, Vol. 29, pp. 1335-1346, 1996.
- [22] Y.J. Zhang, Evaluation and comparison of different segmentation algorithms, *Pattern Recognition Letters*, Vol. 18, pp. 963-974, 1997.