

Sélection de composantes couleurs ou comment trouver un espace couleur discriminant ?

Colour space analysis through selection feature methods.

Romain Raveaux¹

Jean-Christophe Burie¹

Jean-Marc Ogier¹

¹Laboratoire L3I

Université de La Rochelle, 17000, France
romain.raveaux01@univ-lr.fr

Résumé

Le choix d'un espace couleur pertinent est crucial dans les applications de traitements d'images (segmentation, recherche d'objets graphiques...). Partant de ce constat, nous répondons de manière générique à la question suivante : quel est le meilleur espace de représentation pour traiter une image donnée ? Dans cet article, nous proposons un système de sélection d'espace couleur. Le principe est fondé sur des méthodes de sélection de composantes couleurs ainsi que sur l'évaluation de la pertinence d'un modèle de représentation. La qualité d'un espace couleur est évaluée en terme de séparabilité des données. Notre système apporte donc une réponse sur le choix de l'espace de représentation en direction des applications de traitement d'image s'appuyant sur l'information couleur.

Mots Clef

Espace couleur, Espace hybride couleur, sélection de caractéristiques.

Abstract

The choice of a relevant colour space is crucial in the applications of image processing (segmentation, graphic recognition...). From this fact, we answer in a generic way the following question: What is the best representation space for a computational task on a given image? In this article, we propose a colour space selection system. The principle is based on selection feature methods and on the evaluation of the suitability of a representation model. The quality of a colour space is evaluated in term of data separability. Our system brings an answer about the choice of a meaningful representation space dedicated to image processing applications which rely on colour information.

Keywords

Colour space, selection feature methods, hybrid color space.

1 Introduction

La représentation couleur est un élément décisif dans beaucoup d'applications de vision par ordinateur et de traitement d'image. Plusieurs espaces couleurs ont été développés pour le traitement d'image numérique [1], exemple Rouge, Vert, Bleu (RVB) ou Teinte, Saturation, Intensité (TSI). Les performances de n'importe quel système s'appuyant sur la couleur sont hautement influencées par l'espace de représentation qu'il exploite. La qualité d'un espace couleur est déterminée par sa capacité de distinction entre les différentes couleurs ainsi que la robustesse au changement naturel, tel que les variations de couleur, les changements de lumière et les ombrages. Ces deux critères semblent être conflictuelles ce qui représente un réel challenge pour toute approche tentant de créer un espace de représentation optimal (maximisant la distance inter-classes et minimisant la distance intra-classe). Plusieurs travaux de recherches récents ont essayé de trouver un compromis entre ces deux contraintes. Par exemples, dans [2], [3], des joueurs de foot sont classifiés d'après leur information couleur en utilisant un apprentissage supervisé. Dans cette approche les caractéristiques dominantes de différents espaces couleurs sont sélectionnées pour construire un espace hybride couleur (EHC). Cependant, il y a quelques limitations dans l'EHC en général. Par exemple, pour choisir les caractéristiques dominantes, une fonction est itérativement optimisée en évaluant des combinaisons possibles des primaires des modèles candidats. Pratiquement, il est presque impossible d'examiner toutes les combinaisons possibles, puisqu'elles ont un nombre combinatoire égal au factoriel du nombre total des primaires. Par conséquent, l'EHC est construit en recherchant une combinaison acceptable des caractéristiques qui maximisent la fonction d'évaluation. Les inconvénients de l'approche [2] réside dans le fait qu'elle nécessite une connaissance a priori sur l'image, que seul un expert peut fournir, cette étape est donc fastidieuse et diminue la généralité de cette méthode. Un autre point faible se situe dans la manière d'explorer l'espace des solutions, un algorithme glouton « hill

climbing » est utilisé, ce type de recherche est coûteux en temps et ne garantit pas un maximum global.

Notre approche consiste donc en un prétraitement généraliste présent au début de tout traitement d'image et qui a comme objectif de trouver pour une image donnée, un espace couleur maximisant la distinction des ensembles chromatiques porteurs de sens. Le synopsis de notre méthode est présenté en Fig 1.

La suite de notre papier s'articule en quatre étapes. Tout d'abord, nous présentons les travaux déjà effectués dans ce domaine, nous introduisons ainsi les espaces couleurs et les espaces hybrides. Dans un second temps, un tour d'horizon décrit les méthodes de sélections de caractéristiques. Puis troisièmement, une étude expérimentale est menée, les résultats de notre processus de sélection d'espace couleur sont présentés. Enfin, nous concluons dans la section 5.

2 Travaux antérieurs

La plupart des dispositifs d'acquisition et d'affichage couleur, tels que les appareils photo numériques et les scanners, ont leurs entrées ou leurs signaux de sortie dans l'espace RVB. C'est pourquoi l'espace RVB est employé couramment dans les applications de traitement d'image. Cependant, la représentation RVB a plusieurs inconvénients qui diminuent les performances des systèmes qui s'appuient sur ce système couleur. L'espace RVB n'est pas uniforme, c.-à-d. les distances relatives entre les couleurs ne reflètent pas les différences perceptuelles. Par exemple, dans la gamme de couleur de l'espace de RVB la distance entre le bleu= (0.0.1) et le magenta= (1.0.1) est égale la distance entre le magenta et le blanc= (1.1.1). Cependant, le système humain de vision considère la distance perceptuelle entre bleu et le magenta inférieur à la distance entre le blanc et le magenta. Une autre limitation importante de l'espace RVB est l'influence élevée de la lumière et de la spécularité sur les valeurs des primaires couleurs. En d'autres termes, une surface particulière a des valeurs RVB moindre que cette même surface soumise à un éclairage plus lumineux. Pour diminuer la dépendance à l'égard de l'intensité d'illumination, le RVB normalisé (NRGB) a été développé [1], suivant Eq.(1).

Équation 1 : RVB normalisée

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \quad (1)$$

Pour le cerveau humain, il est difficile d'interpréter les valeurs de couleur dans le modèle de RVB. Par conséquent, l'espace TSI a été développé comme une représentation plus proche du système de perception humain, l'homme peut facilement interpréter les primaires de cet espace. Dans l'espace TSI, la longueur d'onde dominante de couleur est représentée par la teinte. La pureté de couleur est représentée par la composante saturation. Enfin, l'obscurité ou la luminosité de la couleur est déterminée par le composant d'intensité. L'équation

Eq.(2) montre la transformation entre les espaces RVB et TSI [4].

Équation 2 : Espace Teinte Saturation Intensité

$$\begin{aligned} I &= \frac{1}{3}(R+G+B) \\ S &= 1 - \frac{3}{R+G+B}[\min(R,G,B)] \\ H &= \begin{cases} \theta & B \leq G \\ 360 - \theta & B > G \end{cases} \\ \text{where } \theta &= \cos^{-1} \left\{ \frac{0.5[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \right\} \end{aligned} \quad (2)$$

Bien que l'espace TSI soit plus approprié aux applications informatique, particulièrement la segmentation couleur, il reste néanmoins quelques problèmes. Spécifiquement, il y a des singularités non-évitable dans la transformation du RVB au TSI, suivant l'Eq.(2). Les valeurs TSI sont très sensibles aux changements des valeurs RVB près de ces singularités. De plus, il y a un problème semblable à celui de l'espace de RVB, qui est la correspondance non-uniforme entre les distances couleurs perçu par le système perceptuel humain et dans l'espace TSI. C'est parce que le changement uniforme du plan SI ne correspond pas à un changement uniforme du point vue de la perception humaine. Par exemple, le changement de la couleur perçue due à une variation de saturation avec une intensité basse est inférieur celle avec une intensité élevée.

L'espace de couleur de XYZ a été développé par la Commission Internationale De L'éclairage (CIE) en 1931 [5]. La base couleur dans le XYZ n'est pas semblable à celle de l'espace de RVB. La primaire R dans le RVB correspond à la quantité physique de la lumière réfléchie dans la bande rouge. Mais dans l'espace de XYZ, chaque primaire ne correspond pas à une lumière physique spécifique. Au lieu de cela, les couleurs sont représentées par la luminance dans une dimension et la chromaticité dans les deux autres dimensions. Cette conception de la base de XYZ garantit qu'il n'y a aucun tristimulus négatif pour n'importe quelle couleur dans le spectre visible.

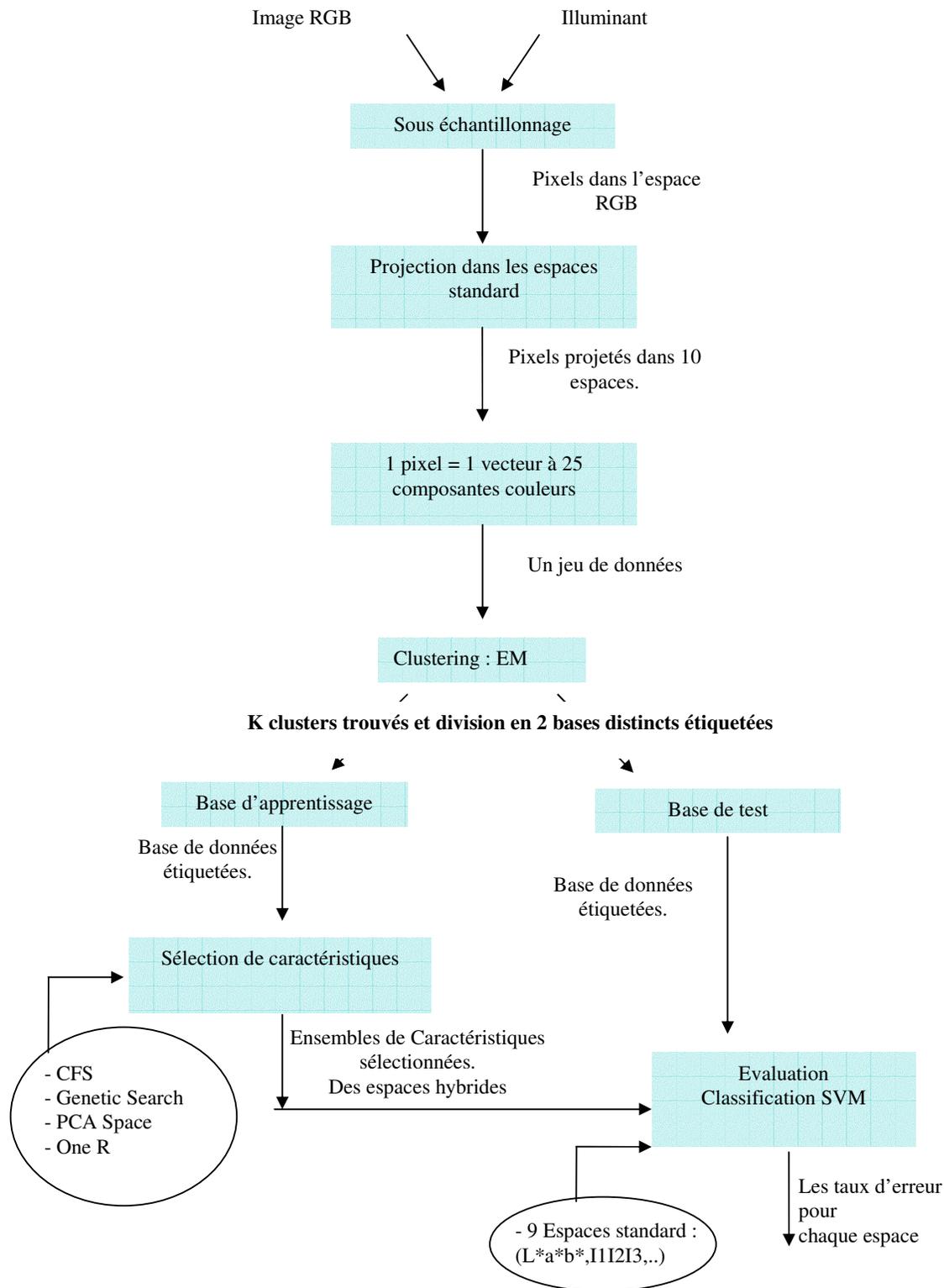


Figure 1 : Sélection automatique d'espace couleur maximisant la séparabilité des données

Le modèle YUV est l'une des représentations de XYZ. Il a été développé pour des applications d'affichage. Le modèle YUV exploite la couleur d'une manière semblable à celle dont le cerveau humain perçoit la couleur. Le modèle YUV comprend la luma (y), qui est semblable à la composante d'intensité dans l'espace de TSI. L'information couleur est représentée par deux composants U (ou Cb) et V (ou Cr). U est la mesure bleue-jaune tels que les valeurs positives représentent des couleurs bleuâtres et les valeurs négatives renvoient à des couleurs jaunâtres. De même que, V est la mesure rouge-verte tels que les valeurs positives représentent des couleurs rougeâtres et les couleurs négatives renvoient à des couleurs verdâtres. Cette représentation vient du fait que le cerveau humain pense au rouge et au bleu comme des composantes opposées au vert et au jaune. Une autre alternative de YUV est le YIQ qui est employé dans la norme de NTSC TV. Dans YIQ, le composant de Y est le même que qui de YUV, alors que I et Q est une paire d'axe a tourné 33° à partir de la paire d'axes UV.

Dans une autre optique, Ohta, Kanade, et Sakai [6] ont choisi un ensemble de couleurs caractéristiques « efficaces » après analyse de 100 caractéristiques couleurs différentes. Ces caractéristiques ont été employées pour segmenter huit types d'images couleurs. Ceux sont ces caractéristiques couleurs sélectionnés que nous nommerons modèle couleur « Ohta » dans le reste de ce papier, cet espace couleur est calculé linéairement à partir de l'espace RVB. Ainsi, XYZ, YUV, et Ohta sont des transformations linéaires de l'espace de RVB. Par conséquent, ils sont des espaces couleurs non-uniformes, La CIE a recommandé le Cie-Lab et le Cie-Luv en tant qu'espaces de couleur uniformes, car ils sont des transformations non linéaires de l'espace de RVB [7].

Dans [8], des caractéristiques couleurs provenant de différents espaces couleurs sont choisis pour construire un EHD (espace hybride décorrélé). Une analyse en composante principale (ACP) est exécutée sur la matrice de covariance composée du nombre total de primaires couleurs disponible. Les 3 axes les plus significatifs sont choisis pour réduire le taux de corrélation entre les composantes couleurs. Cette technique nous amène à décrire plus amplement les méthodes de sélection de caractéristiques.

3 La sélection de caractéristiques

La sélection de caractéristiques (feature selection) est un domaine très actif depuis quelques années, en particulier dans le cadre du data-mining. En effet, la "fouille de données" dans de très grandes bases devient un enjeu crucial pour des applications telles que le traitement d'image, la finance, etc.. Il s'agit en fait de résumer et d'extraire intelligemment de la "connaissance" à partir de données brutes. La fouille de données est un domaine basé sur la statistique, l'apprentissage automatique et la théorie des bases de données. La sélection de variables joue un rôle important dans le data-mining, en particulier dans la

préparation des données avant leur traitement. En effet, les intérêts de la sélection de variables sont les suivants :

- Lorsque le nombre de variables est vraiment trop grand l'algorithme d'apprentissage ne peut pas se terminer en un temps convenable. La sélection réduit la dimension de l'espace des caractéristiques.

- D'un point de vue intelligence artificielle, créer un classifieur revient à créer un modèle pour les données. Or, une attente légitime pour un modèle est d'être le plus simple possible (principe du Rasoir d'Occam [9]). La réduction de la dimension de l'espace des caractéristiques permet alors de réduire le nombre de paramètres nécessaires à la description de ce modèle évitant également le phénomène de sur apprentissage et accentuant la synthèse des informations.

- Elle améliore la performance de la classification, sa vitesse et son pouvoir de généralisation.

- Elle augmente la compréhensibilité des données : on voit mieux quels sont les processus qui leur ont donné naissance. Cette sélection consiste en :

- L'élimination de variables indépendantes de la classe,

- L'élimination de variables redondantes.

3.1 Concept global

Une structure générale des algorithmes de sélection de caractéristiques peut être proposée de la façon de la figure 2 ([10]). Jusqu'à ce qu'un certain critère soit satisfait, des sous-ensembles Y sont générés en parcourant l'espace des sous ensembles, et ils sont évalués. Nous d'écrivons successivement chaque phase.

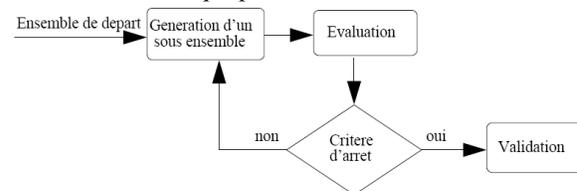


Figure 2 : Architecture de la procédure de sélection de caractéristiques

La génération de sous-ensembles est une procédure de recherche dans l'espace des sous-ensembles de cardinal 2^N . Toutes les méthodes de parcours classiques sont utilisables. Par exemple [11] propose les méthodes de forward addition et backward élimination, [12] et [13] utilisent des algorithmes évolutionnistes.

3.2 Recherche et Evaluation des caractéristiques

L'évaluation d'un sous-ensemble est traitée de façons très diverses. Il existe deux grandes classes d'algorithmes :

1. Les wrappers : qui utilisent l'algorithme de classification pour évaluer les sous-ensembles générés.
2. Les filters : qui sont complètement indépendants de cet algorithme, mais se basent sur des considérations statistiques, entropiques, de cohérence, de distance, etc.

Les wrappers : Bien que conceptuellement plus simples que les filtres, les wrappers ont été introduits plus récemment par John, Kohavi et Pfleger en 1994. Leur principe est de générer des sous-ensembles candidats et de les évaluer grâce à l'algorithme de classification. Le score d'un ensemble sera par exemple un compromis entre le nombre de variables éliminées et le taux de réussite de la classification sur un fichier de test. Ainsi la phase "évaluation" du cycle de sélection est constituée par un appel à l'algorithme de classification. En fait, celui-ci est appelé plusieurs fois à chaque évaluation car un mécanisme de validation croisée est fréquemment utilisé. De par son principe même, cette méthode génère des sous-ensembles bien adaptés à l'algorithme de classification. Les taux de reconnaissance sont hauts car la sélection prend en compte le biais intrinsèque de l'algorithme de classification. Un autre avantage est sa simplicité conceptuelle : il n'y a nul besoin de comprendre comment l'induction est affectée par la sélection de variables, il suffit de générer et de tester.

Cependant, trois raisons font que les wrappers ne constituent pas une solution parfaite. D'abord, ils n'apportent pas vraiment de justification théorique à la sélection et ils ne nous permettent pas de comprendre les relations de dépendances conditionnelles qu'il peut y avoir entre les variables. D'autre part la procédure de sélection est spécifique à un algorithme de classification particulier et les sous-ensembles trouvés ne sont pas forcément valides si on change de méthode d'induction. Finalement, et c'est le défaut principal de la méthode, les calculs deviennent vite très longs, voir irréalisables lorsque le nombre de variable croît.

Les filtres : Les filtres n'ont pas les défauts des wrappers. Ils sont beaucoup plus rapides, ils reposent sur des considérations plus théoriques, ils nous permettent de mieux comprendre les relations de dépendance entre variables. Mais, comme ils ne prennent pas en compte les biais de l'algorithme de classification, les sous-ensembles de variables générés donnent un taux de reconnaissance plus faible. Pour donner un score à un sous-ensemble, une première solution est de donner un score à chaque variable indépendamment des autres et de faire la somme de ces scores [OneR Selection]. Pour évaluer une variable, l'idée est de déterminer sa corrélation avec la variable de classe. Mais [11] propose des exemples simples montrant que cette approche nommée feature ranking pose des problèmes dans le cas général. En effet, cette approche n'élimine pas les variables redondantes, d'autre part il est possible que des variables peu corrélées avec la classe deviennent utiles lorsqu'on les considère dans le contexte des autres variables. L'autre solution est d'évaluer un sous-ensemble dans sa globalité [12]. On se rapproche ici de l'apprentissage de la structure de réseau bayésien. Il existe un intermédiaire entre feature ranking et subset ranking basé sur une idée de Ghiselli et utilise avec de bons résultats dans le cadre de la CFS (correlation based feature selection) par M. Hall [10]. Le score d'un

sous-ensemble est construit en fonction des corrélations variable-classe et des corrélations variable-variable selon la formule suivante (Eq 3) :

Équation 3 : Score de corrélation associé à chaque caractéristique dans la méthode CFS

$$r_{\theta S} = \frac{k \overline{r_{\theta i}}}{\sqrt{k + k(k-1) \overline{r_{ij}}}}$$

Avec $r_{\theta S}$ le score du sous-ensemble de cardinal k , $\overline{r_{\theta i}}$ la moyenne arithmétique des corrélations entre θ et les variables i , et $\overline{r_{ij}}$ la moyenne des k^2 intercorrélations entre variables. Cette équation exprime que le score du sous-ensemble augmente si les variables sont fortement corrélées avec θ et diminue si elles sont fortement corrélées entre elles. De plus cette formule est valable dans le cadre de scores normalisés (*i.e.* de variance unitaire). L'idée est de dire qu'un bon sous-ensemble est constitué de variables hautement corrélées avec la classe (pour ne pas garder les indépendantes), et peu corrélées entre elles (pour éviter la redondance). Il s'agit d'une approximation car on ne prend en compte que les interactions d'ordre 1. La corrélation ou dépendance entre deux variables peut être définie de plusieurs façons. Utiliser le coefficient de corrélation statistique est trop restrictif car il ne capture que les dépendances linéaires. On peut, en revanche, utiliser un test d'indépendance statistique comme le test du χ^2 . Il est aussi possible de combiner des méthodes *wrapper* et *filter* comme présentées dans [13].

3.3 Critère d'arrêt

Le critère d'arrêt peut être de diverses natures : un temps de calcul, un nombre de générations (pour un algorithme génétique), un nombre de variables sélectionnées, une évaluation heuristique de la "valeur" du sous-ensemble.

Tableau 1: Résumé des méthodes de sélections de caractéristiques utilisées

<i>Nom</i>	<i>Type</i>	<i>Evaluation</i>	<i>Recherche</i>
CFS	Filter	CFS	Greedy stepwise
EHD	Filter	ACP	Ranker
GACS	Wrapper	Classification	Algorithme Génétique
OneRS	Wrapper	Classification	Ranker

3.4 Projection multi espaces

Chaque pixel RVB est projeté dans neuf espaces couleurs afin de construire un vecteur de caractéristiques associé à chaque pixel. Soit un ensemble $C = \{C_i\}_{i=1}^N = \{R,G,B, I1,I2,I3, L^*, u^*, v^*, \dots\}$ avec $\text{Card}(C) = 25$.

3.5 Espace hybride en utilisant les algorithmes génétiques (GACS)

Dans un article précédent [14], nous avons défini un algorithme de recherche d'espace couleur. Ce dernier se démarque des autres méthodes évolutionniste par le codage de ses individus sous forme de vecteur limité à trois composantes spécifiquement dédié à la représentation de la couleur. Nous rappelons ici les principes généraux de notre algorithme. La première étape est d'initialiser la population, chaque individu est construit en tirant aléatoirement trois éléments de l'ensemble C. En ce qui concerne l'opérateur de croisement, deux individus $h1$ et $h2$ partagent leur matériel génétique, en permutant une de leur composante, Fig3. Ensuite, pour mener à bien une mutation sur un individu, une composante est choisie et est remplacée par un élément issu de C. Alors que finalement, la phase d'évaluation exploite un classifieur SVM (Séparateur à Vaste Marge) avec une validation croisée sur la base d'apprentissage.

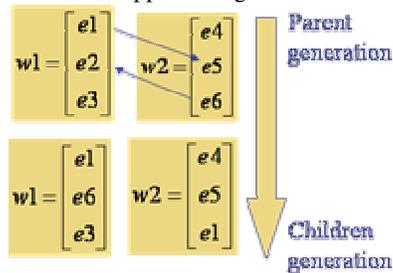


Figure 3 : Opérateur de croisement

4 Résultats expérimentaux

Notre approche est mise en œuvre sur trois différents types d'images. Une scène naturelle, une image de document et une image de synthétique. Chaque espace couleur est évalué sur la base de test à l'aide d'une classification SVM.

Tableau 2 : Description des images de test

<i>Id</i>	<i>Image</i>	<i>Type</i>	<i>Nombre de clusters</i>
Im1	Lenna	Scène naturelle	18
Im2	SatSnake	Image synthétique discriminée par la saturation	23
Im3	Image de document	Plan cadastral	9



Figure 4 : Images traitées par notre procédé de sélection d'espace couleur.

Tableau 3 : Bases de test et bases d'apprentissage

	$ X_{training} $ pixels	$ X_{test} $ pixels
IM1	13107	13107
IM2	10951	10951
IM3	11424	11424

Tableau 4 : Espaces hybrides trouvés sur l'image Im2

<i>Attributs</i>	<i>CFS</i>	<i>GACS</i>	<i>EHD</i>	<i>OneRs</i>
R	0	0	0	0
G	1	0	1	0
B	1	0	0	0
I1	0	0	0	0
I2	1	0	0	0
I3		0	0	0
T	1	0	0	1
S	1	2	0	1
I	0	0	0	0
L*	1	0	0	0
a*	1	0	0	0
b*	1	0	0	0
L*	0	0	0	0
u*	1	1	0	0
v*	1	0	0	0
A	0	0	0	0
C1	1	0	0	0
C2	1	0	0	0
X	0	0	0	0
Y	0	0	1	0
Z	1	0	1	0
Y	0	0	0	0
I	1	0	0	1
Q	1	0	0	0
Y	0	0	0	0
U	1	0	0	0
V	0	0	0	0
Nombre d'attributs	16	3	3	3

Tableau 5 : Nombre d'attributs sélectionnés

	<i>Nombre d'attributs sélectionnés</i>			
	CFS	GACS	EHD	OneRS
IM1	16	3	3	3
IM2	16	3	3	3
IM3	12	3	3	3

Tableau 6 : Taux de confusion sur l'image 1.

IM1			
<i>Espaces Couleurs</i>	<i>Erreur</i>	<i>Espaces Couleurs</i>	<i>Erreur</i>
RGB	0.3608	TSI	0.3917
I1I2I3	0.3814	La*b*	0.4329
XYZ	0.3814	L*u*v*	0.4948
YIQ	0.4742	EHD	0.3917
YUV	0.3195	CFS	0.0615
AC1C2	0.4123	GACS	0.2680
ACP	0.3711	OnRS	0.3608

Tableau 7 : Taux de confusion sur l'image 2.

IM2			
<i>Espaces Couleurs</i>	<i>Erreur</i>	<i>Espaces Couleurs</i>	<i>Erreur</i>
RGB	0.22	TSI	0.51
I1I2I3	0.24	La*b*	0.41
XYZ	0.23	L*u*v*	0.35
YIQ	0.43	EHD	0.39
YUV	0.35	CFS	0.14
AC1C2	0.29	GACS	0.19
OnRS	0.39	ACP	0.35

Tableau 8 : Tableau 5 : Taux de confusion sur l'image 3.

IM3			
<i>Espaces Couleurs</i>	<i>Erreur</i>	<i>Espaces Couleurs</i>	<i>Erreur</i>
RGB	0.5444	TSI	0.3666
I1I2I3	0.2222	La*b*	0.2666
XYZ	0.5777	L*u*v*	0.3333
YIQ	0.3111	EHD	0.36
YUV	0.3777	CFS	0.0333
AC1C2	0.3	GACS	0.1888
OnRS	0.4111	ACP	0.2444

Analyse des résultats :

L'espace couleur minimisant le taux d'erreur de classification est l'espace discriminant le mieux les pixels de notre image RVB de départ. L'espace engendrant le

moins d'erreur sera donc gardé pour continuer les traitements sur l'image. L'espace choisi minimise donc la distance intra-classe, à l'intérieur d'un même ensemble chromatique tout en maximisant la distance inter-classes, c-a-d en augmentant la distinction des couleurs. De telles propriétés sont favorables à des post-traitements tel que la segmentation ou l'extraction d'objet graphique.

5 Conclusion

Dans ce papier, nous avons présenté un framework de sélection d'espace couleur. Notre contribution porte sur un système clef-en-main permettant de trouver un espace couleur adapté. Notre outil se place en amont de tout processus s'appuyant sur des données couleur pour extraire de l'information. Partant du constat qu'un réel besoin existe en matière de représentation de la couleur, les espaces de la CIE sont particulièrement bien adaptés aux images « standards » (Scènes naturelles), mais trouvent leurs limites lorsqu'il s'agit d'images très spécifiques, tel que les images médicales ou des images de plans cadastraux, alors un espace dédié construit par notre approche est nécessaire.

Bibliographie

- [1] G. Wyszecki and W. S. Stiles. Color Science: Concepts and Methods, Quantitative Data and Formulae. John Wiley & sons, second edition, 1982.
- [2] N. Vandenbroucke, L. Macaire, and J.G. Postaire. Color pixels classification in an hybrid color space. In Proceedings of the IEEE International Conference on Image Processing - ICIP'98, volume 1, pages 176-180, Chicago, 1998.
- [3] N. Vandenbroucke, L. Macaire, and J. G. Postaire. Color image segmentation by pixel classification in an adapted hybrid color space: application to soccer image analysis. Computer Vision and Image Understanding, 90(2):190-216, 2003.
- [4] J. M. Tenenbaum, T. D. Garvey, S.Weyl, and H. C.Wolf. An interactive facility for scene analysis research. Technical Report 87, Artificial Intelligent Center, Stanford Research Institute, Menlo Park, CA, 1974.
- [5] <http://www.cie.co.at/cie/index.html>.
- [6] Y. I. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. Computer Graphics and Image Processing, 13:222-241, 1980.
- [7] H. Palus. Colour spaces. In S.J. Sangwine and R.E.N. Home, editors, The Colour Image Processing Handbook, pages 67-90. Chapman & Hall, Cambridge, Great Britain, 1998.
- [8] J. D. Rugna, P. Colantoni, and N. Boukala, "Hybrid color spaces applied to image database".
- [9] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K.Warmuth. Occam's razor. Information Processing Letters, 24(6) :377-380, 1987.

- [10] M. Hall. Correlation-based feature selection for machine learning, 1998. Thesis IN Computer Science at the University of Waikato.
- [11] Daphne Koller and Mehran Sahami. Toward optimal feature selection. In International Conference on Machine Learning, pages 284–292, 1996.
- [12] P. Dangauthier. Feature Selection For Self-Supervised Learning, AAAI Spring Symposium Series, AAAI (American Association for Artificial Intelligence), 445 Burgess Drive, Menlo Park, California 94025-3442 USA, March 2005
- [13] Jihoon Yang and Vasant Honavar. Feature subset selection using a genetic algorithm. IEEE Intelligent Systems, 13 :44–49, 1998.
- [14] Romain Raveaux, Jean-Christophe Burie, Jean-Marc Ogier. A colour document interpretation: Application to ancient cadastral maps. The 9th International Conference On Document Analysis(ICDAR 2007).