# A multi color space approach for texture classification: experiments with Outex, Vistex and Barktex image databases

A. Porebski $^{(1, 2)}$ , N. Vandenbroucke $^{(1, 2)}$  and L. Macaire $^{(2)}$ 

alice.porebski@eipc.fr, nicolas.vandenbroucke@eipc.fr, ludovic.macaire@univ-lille1.fr

(1) École d'Ingénieurs du Pas-de-Calais Département Automatique Campus de la Malassise 62967 Longuenesse Cedex - FRANCE

## Abstract

The color of pixels can be represented in different color spaces which respect different properties. Many authors have compared the classification performances reached by these color spaces in order to determine the one which would be the well suited to color texture analysis. However, the synthesis of these works shows that the choice of the color space depends on the considered texture images. Moreover, the prior determination of a color space which is well suited to the considered class discrimination is not easy. That is why we propose to consider a multi color space approach designed for color texture classification. It consists in selecting, among a set of color texture features extracted from images coded in different color spaces, those which are the most discriminating for the considered color textures. In this paper, we experimentally study the contribution of this multi color space with three well-known benchmark databases, namely Outex, Vistex and Barktex. Comparison and discussion are then carried out.

## Introduction

Many authors have shown that the use of color improves the results of texture classification compared with grey level image analysis [4, 8, 11, 3]. However, there exists a large number of color spaces which respect different properties and it is well known that the choice of the color space impacts the quality of color texture classification. That is why many authors have compared the performances reached by classification schemes operating in different color spaces in order to determine the best suited one to color texture analysis [4, 8, 11, 18, 6, 20]. The synthesis of these works shows contradictions that do not allow to conclude on the definition of a single color space which is well suited to the discrimination of all the textures. Moreover, the prior determination of the color space which is well suited to the considered texture discrimination, is not easy. That is why we propose to experimentally study in this paper a multi color space approach where the properties of several different color spaces are simultaneous considered in order to analyze any color textures. For this purpose, color texture features are first computed from images coded in several color spaces. The most discriminating color texture features are then selected thanks to a sequential feature selection procedure achieved during a supervised learning scheme. In this paper, we propose to measure the contribution of the multi color space approach by comparing the classification results reached by a single color space approach with those obtained by the proposed multi color space approach. In order to assess these results, this (2) Laboratoire LAGIS - FRE CNRS 3303 Université Lille 1 - Sciences et Technologies Cité Scientifique - Bâtiment P2 59655 Villeneuve d'Ascq - FRANCE

experimental study is carried out on three well-know benchmark databases: Outex, Vistex and Barktex [10, 12, 7].

In the second section of this paper, we present the influence of the choice of color spaces on texture classification. Then, in the third section, we describe the color texture features used in this paper to carry out our experiments. The fourth section details the used texture classification method, which is based on the sequential selection of a low dimensional feature space. In the fifth section, the set of images coming from the OuTex, VisTex and BarkTex image databases and used in our experiments are presented. Finally, experimental results show the contribution of the multi color space approach for color texture classification.

## **Color spaces**

The color of pixels can be represented in different color spaces which respect different physical, physiologic, and psychovisual properties [2]. Figure 1 shows several spaces commonly used in color image analysis. They can be classified into four families: the primary spaces, the luminance-chrominance spaces, the perceptual spaces and the independent color component spaces [2].

Many authors have tried to compare the classification performances obtained with different color spaces [4, 8, 11, 18, 6, 20]. Their goal is to determine the color space which would be the best suited to their color texture classification application. However, the synthesis of these works shows that there is no color space which is adapted to the discrimination of all the textures. For example, the conclusions of Mäenpää draw that the choice of color space choice depends on the considered texture images. Indeed, by applying the same color texture classification method to the images of the VisTex and OuTex databases, he shows that the  $(L^*, a^*, b^*)$  perceptually uniform color space allows to obtain better results than the (R, G, B) acquisition color space when the Vis-Tex texture images are classified, and inversely when the OuTex ones are considered [8]. By applying other classification methods and by using other texture features, the best classification results are obtained with the (I4, S3, H2) perceptual color space for the Outex database and with the (I1, I2, I3) independent component color space for the Vistex Database.

On the one hand, when one considers several available color spaces, it is difficult to select those which are well suited to all the color image analysis applications. On the other hand, the performance of an image classification procedure depends on the choice of the color space.

To solve this problem, few works propose to combine differ-



Figure 1. Color space families.

ent color spaces or color components in order to improve color image classification performance. Chindaro uses a color texture classification system based on a set of independent classifiers. Each of these classifiers is assigned to a specific color space. In order to classify images, he fuses the classification decisions of the classifiers. He concludes that the association of informations coming from different color spaces improves performances [3].

Porebski proposes to code color texture images into several different color spaces and to compute some texture features from the so-coded images. A feature selection procedure selects the most discriminating color texture features for the texture classification purpose. She proposes an approach that selects the most discriminating Haralick features extracted from chromatic co-occurrence matrices of color images coded in 28 different color spaces thanks to a sequential forward selection [14].

In order to take into account all the properties of these spaces, we propose to extract color texture features from images coded in these  $N_S = 28$  spaces and to select the most discriminating ones in order to classify textures images.

In this paper, we propose to experimentally compare the performance of a multi color space approach where the  $N_S = 28$  color spaces of figure 1 are used with the performance reached by a single color space approach.

### Color texture features

For the texture classification purposes, a color texture is described by a set of d features and is represented by a point in a d-dimensional feature space in order to achieve the classification.

There exists a large number of color texture descriptors and it is well known that the performance of the classifier depends on the choice of the color texture feature but also on the dimension of the feature space. Indeed, a too high dimensional feature space can decrease the rate of well-classified textures and increase the computation time.

Different definitions of "color texture" have been proposed in the literature depending on color and texture informations are considered jointly or separately [4, 8]. In order to combine both the color and the texture informations, three kinds of features are used [11, 1]:

- "luminance based texture features" mixed with color statistical features,
- "within color component texture features" which take into account only the spatial relationships within a single color component,
- "between color component texture features" which consider spatial relationships within and between different color components.

In order to represent both the color and texture informations, we need to use color texture features which take into account both the color distribution in the color space and the spatial interaction between pixels in the image plane. Moreover, because the color information is projected in a three-dimensional color space, we need features which consider spatial relationships within and between different color components in order to completely describe the information represented by the whole color texture.

Even if many color texture features can be used, we have shown that the well known Haralick features extracted from Reduced Size Chromatic Co-occurrence Matrices (RSCCM) are relevant for this purpose [17]. That is why, we choose these features to be easily implemented in order to carry out our experimental study.

RSCCMs are  $N \times N$  co-occurrence matrices where the quantization level N of the color components is reduced in order to decrease the memory storage cost and so, the time required to extract features from these matrices. In our experiments, N is set to 16. It has been proved that the choice of the neighborhood shape used to compute RSCCM depends on the analyzed textures [15]. When textures do not present any privileged direction, it is necessary to compute features which take into account all the possible directions. That is the reason why we choose to use the  $3 \times 3$ isotropic 8-neighborhood in order to take into account all directions.

Each color image I, coded in a  $(C_1, C_2, C_3)$  color space of the figure 1, is characterized by the  $N_M = 6$  following RSCCMs:

- three within color component matrices:  $m^{C_1,C_1}[\mathbf{I}], m^{C_2,C_2}[\mathbf{I}] \text{ and } m^{C_3,C_3}[\mathbf{I}],$
- three between color component ones:  $m^{C_1,C_2}[\mathbf{I}], m^{C_1,C_3}[\mathbf{I}] \text{ and } m^{C_2,C_3}[\mathbf{I}].$

Usually, RSCCMs are not directly exploited because they contain a large amount of information. To reduce it while preserving the relevance of these descriptors, we use the first  $N_H = 13$  Haralick features denoted  $f_1$  to  $f_{13}$  and extracted from these matrices [5].

In order to exploit all the properties of the  $N_S$  color spaces, each image is firstly coded in each of these color spaces. Then,

for each color space, the  $N_M = 6$  RSCCMs are computed and the  $N_H = 13$  Haralick features are extracted from each matrix. A color texture is firstly represented by  $N_f = N_H \times N_M \times N_S = 13 \times 6 \times 28 = 2184$  candidate color texture features (see. figure 2).



Figure 2. Candidate color texture features.

Because the dimension of the feature space is too high, it is necessary to select the most discriminating color texture features during a supervised learning stage. For this purpose, we use a sequential forward selection (SFS) procedure [14]. The next section details the supervised classification scheme used in our experimental study.

## **Texture classification scheme**

In order to carry out our experimental study, we propose to apply a supervised classification scheme divided into two successive stages:

- A supervised learning stage selects a low number of discriminating texture features among a set of candidate ones in order to build a feature space thanks to a feature selection procedure. During this stage, the classifier is trained to partition this feature space.
- A decision stage where the examined color texture is represented by a point in the so-selected feature space.

For this purpose, each set of images used in experiments is divided into two subsets: the training subset and the test subset. The training and the test subsets are built according to the Holdout method: half of the images are used to build the training subset, while the other images are used to test the performance reached by the classifier.

#### Feature selection

The determination of the most discriminating feature space is achieved thanks to a sequential feature selection procedure based on a supervised learning scheme. This non-exhaustive search procedure has given very encouraging results for color image classification of the Outex, Vistex and Barktex databases [17, 14].

At each step *s* of this procedure, an informational criterion  $J_s$  is evaluated in order to measure the discriminating power of each candidate feature space. At the beginning of this procedure (s = 1), the  $N_f$  one-dimensional candidate feature spaces, defined by each of the  $N_f$  available color texture features, are examined. The candidate feature which maximizes (or minimizes)  $J_1$  is selected at the first step. It is associated in the second step of the procedure (s = 2) with each of the  $(N_f - 1)$  remaining candidate color texture features in order to constitute  $(N_f - 1)$  two-dimensional candidate feature spaces. We consider that the two-dimensional

space which maximizes (or minimizes)  $J_2$ , is the best space for discriminating the texture classes...

In order to only select color texture features which are not correlated, we measure, at each step  $s \ge 2$  of the procedure, the correlation level between each of the available color texture features and each of the (s-1) other color texture features constituting the previously selected (s-1)-dimensional space. The considered features will be selected as candidate ones only if their correlation level with the already selected color texture features is lower than a threshold fixed by the user [19].

We assume that the more the clusters associated with the different texture classes are well separated and compact in the candidate feature space, the higher the discriminating power of the selected color texture features is. That leads us to choose measures of class separability and class compactness as measures of the discriminating power.

The measure of compactness of each texture class  $T_j$  is defined by the within-class dispersion matrix  $\Sigma_C$ . The measure of the class separability is defined by the between-class dispersion matrix  $\Sigma_S$ . The most discriminating feature space minimizes the Wilks's criterion:

$$J_{s} = \frac{|\Sigma_{C}|}{|\Sigma_{C} + \Sigma_{S}|}.$$

At each step of the sequential selection procedure, a feature is added and the dimension d of the selected feature space increases. This procedure is iterated  $d_{max}$  times in order to build  $d_{max} d$ -dimensional pre-selected color texture feature spaces  $(1 \le d \le d_{max})$ . In order to determine the dimension  $\hat{d}$  of the most discriminating feature space, we propose to measure the rate  $T_d$ of well-classified images obtained with each d-dimensional preselected space. The selected color texture feature space is the pre-selected space for which  $T_d$  is maximum:

$$\hat{d} = \operatorname*{argmax}_{d=1}^{d_{max}} T_d.$$

Usually, the images of the training subset are used to determine the rate of well-classified images. However, when an image is examined, the partition of the feature space can be modified, for example when a k-nearest neighbor classifier is used. That is why, we propose a different approach where another set of images is used to measure the rate of well-classified images. Here, the test subset is used for this purpose.

#### Classification

During the decision stage, we propose to classify the images of the test subset thanks to the nearest neighbor (1-NN) classifier which is a simple and widely used classifier. This classifier operates in the selected  $\hat{d}$ -dimensional feature space.

## Experimented texture databases

In this section, we present three well-known and largely used benchmark color texture databases examined by our experiments and more precisely, three different benchmark texture sets, coming from OuTex, VisTex and BarkTex databases, respectively [10, 12, 7].

#### **OuTex database**

Outex contains a very large number of surface textures acquired under controlled conditions by a 3-CCD digital color camera. To build the Outex set, each of 68 textures of this database is split up from an original image into disjoint sub-images whose size is  $128 \times 128$  pixels. Since the original image size is  $746 \times 538$ pixels, there are 20 sub-images by a texture. Figure 3 illustrates each class of the used Outex texture by one image.



Figure 3. Example of OuTex color textures: each image represents a class of texture.

The examined dataset is available at the OuTex web site (http://www.outex.oulu.fi/temp/) as test suite Contrib-TC-00013. Among the 1360 images of the Outex set, 680 images constitute the training subset and 680 images constitute the test subset [13, 8, 1, 6, 20].

## VisTex database

Vistex is a large set of natural color textures acquired under non-controlled conditions by different color cameras. To build the Vistex set, each of 54 textures of this database is split up from an original image into disjoint sub-images whose size is  $128 \times 128$ pixels. Since the original image size is  $512 \times 512$  pixels, there are 16 sub-images by a texture (see figure 4).



Figure 4. Example of VisTex color textures: each image represents a class of texture

This set is available at the OuTex web site (http://www.outex.oulu.fi/temp/) as test suite Contrib-TC-00006. Among the 864 images of the Vistex set, 432 images are used for the training subset and 432 for the test subset [13, 8, 1, 6].

## BarkTex database

Color images of the BarkTex database are equally divided into six tree bark classes, with 68 images by class. The size of each image is  $256 \times 384$  pixels.

To build the BarkTex set, a region of interest, centered on the bark and whose size is  $128 \times 128$  pixels, is firstly defined. Then, four sub-images whose size is  $64 \times 64$  pixels are extracted from each region. We thus obtain a set of  $68 \times 4 = 272$  sub-images by a class (see figure 5).



Figure 5. Example of BarkTex color textures: each column represents a class of texture.

Among the 1632 images of the Barktex set, 816 images are used for the training subset and 816 for the test subset.

## **Experimental results**

In order to examine the contribution of the multi color space approach on the quality of texture classification, the proposed classification method is separately applied to the texture sets which have been previously presented.

For this purpose, we propose to characterize the textures according to different ways:

- Haralick features are computed from grey level cooccurrence matrices. In these cases, only the luminance computed from (R, G, B) color is analyzed. So, only 13 features are considered.
- Haralick features are extracted from RSCCMs computed from images coded in a single color space. In these cases, d features are selected among  $13 \times 6$  available ones. We measure the classification results obtained by examining images coded in the (R, G, B) acquisition color space. We also evaluate the classification results obtained by considering images coded in the color space for which the best rate of well-classified images is obtained by other authors.
- Haralick features are extracted from RSCCMs computed from images coded in the  $N_S = 28$  different color spaces of figure 1. In these cases, the feature selection scheme determines *d* features among  $13 \times 6 \times 28$  available ones.

The best rates of well-classified images for the Outex and Vistex sets are obtained by Mäenpää with the (I4,S3,H2) and (I1,I2,I3) color spaces respectively [8]. To obtain these result,

the 3D color histograms are directly used as features. For the Barktex set, Münzenmayer obtains a rate of well-classified images of 87 % with the (R, G, B) color space and by using features extracted from interplane sum- and difference-histograms [9].

Figures 6, 7 and 8 illustrate the classification results obtained by our approach applied to OuTex, VisTex and BarkTex color texture image databases, respectively. These figures display, for the four different ways to characterize color textures, the rates T of well-classified images of the test subset, with respect to the dimension d of the feature space built by the feature selection procedure. Let us notice that when the color textures are characterized by grey level features, the number  $N_f$  of candidate features is not higher than 13.



**Figure 6.** Rates T(%) of well-classified images obtained with the OuTex test subset according to the dimension *d* of the selected feature space.







**Figure 8.** Rates T(%) of well-classified images obtained with the BarkTex test subset according to the dimension *d* of the selected feature space.

Globally, we can notice that the classification results tend to stabilize when d is higher than 10.

The classification results obtained with features extracted from images coded in a single color space outperform those obtained with grey level features. So, the use of color significantly increases the quality of texture classification.

On the other hand, the comparison of the classification results reached with the Outex and Vistex sets by considering the (R,G,B) color space with those selected by Mäenpää, shows that choosing a relevant color space to discriminate color texture classes significantly improves the classification results. However, the prior choice of such a space is difficult and the feature selection procedure is thus an efficient way to determine the most discriminating feature space.

Finally, the difference between the classification results reached by considering the relevant color space and those obtained with the  $N_S = 28$  color spaces shows the advantage of considering the properties of different color spaces for texture discrimination.

We propose to compare in depth results obtained with the Outex database (see table 1).

Comparison between different rates of well-classified images reached with the Outex test subset.

Texture features	Color space	â	Classifier	T(%)
Haralick features from	Multi color	38	1-NN	95.4
RSCCM ( $N = 16$ )	space			
Haralick features from	(I4, S3, H2)	20	1-NN	91.9
RSCCM ( $N = 16$ )				
Haralick features from	(R,G,B)	45	1-NN	88.1
RSCCM ( $N = 16$ )				
Haralick features from	Luminance	13	1-NN	72.9
<b>RSCCM (</b> <i>N</i> = 16)				
3D color histogram [8]	( <i>I</i> 4, <i>S</i> 3, <i>H</i> 2)	4096	1-NN	95.4
3D color histogram [13]	(R,G,B)	4096	3-NN	94.7
Haralick features from	(R,G,B)	30	5-NN	94.9
RSCCM (N = 32) [1]				
LBP histogram [6]	(I4, S3, H2)	4608	SVM	93.5
Wavelet coefficients [20]	(R,G,B)	5	7-NN	85.2

This table firstly shows that the multi color space approach allows to take advantage of the properties of each color space and so, to significantly improve the classification results compared to the use of the properties of a single beforehand chosen color space. It also constitutes an unified approach to classify the color textures.

The results obtained with the proposed multi color space approach (*T* reaches 95.4%) are equal to those obtained by using 3D color histogram [8]. We can notice that the dimension  $\hat{d}$  of the feature space extracted by our multi color space approach ( $\hat{d}$  is equal to 38) is lower than the dimension of the 3D color histogram used by Mäenpää with a single color space ( $\hat{d}$  is equal to 4096). This property can be very interesting for industrial applications which present strong real-time constraints about on-line texture classification.

However, 3D color histogram is not considered as a texture feature because it does not take into account the spatial relationships between pixels in the image plane. That is the reason why, texture features coupled with statistical features about color distribution seem to be efficient to classify images of the Outex database.

## Conclusion

In this paper, we have shown the interest of considering several color spaces for color texture classification and have measured the contribution of this multi color space approach with regard to texture classification results.

For this purpose, textures have been characterized by Haralick features extracted from RSCCMs computed from images coded in  $N_S = 28$  different color spaces. The most discriminating color texture features have then been selected thanks to a sequential feature selection procedure performed during a supervised learning.

Experimental results achieved with OuTex, VisTex and BarkTex databases have shown that the multi color space approach allows to increase significantly the results of color texture classification compared with a single relevant color space.

Taking into account the properties of different color spaces allows to follow an unified approach to classify the textures. However, this approach is more time consuming than the single color space approach. In order to generalize our approach, we are presently working on coupling Haralick features with other wellknown features like LBP (Local Binary Pattern) features [16].

## Acknowledgements

This research is funded by "Pôle de Compétitivité Maud" and "Région Nord-Pas de Calais".

### References

- V. Arvis, C. Debain, M. Berducat and A. Benassi, Generalization of the cooccurrence matrix for colour images: application to colour texture classification, Image Analysis and Stereology, vol. 23, pp. 63-72, 2004.
- [2] L. Busin, N. Vandenbroucke, and L. Macaire, Color spaces and image segmentation, Advances in Imaging and Electron Physics, vol. 151, ch. 2, pp. 65-168, 2008.
- [3] S. Chindaro, K. Sirlantzis and F. Deravi, Texture classification system using colour space fusion, Electronics Letters, vol. 41, nb. 21, pp. 589-590, 2005.
- [4] A. Drimbarean and P.-F. Whelan, Experiments in colour texture analysis, Pattern Recognition Letters, vol. 22, nb. 10, pp. 1161-1167, 2001.
- [5] R. Haralick, K. Shanmugan and I. Dinstein, Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics, vol. 3, nb. 6, pp. 610-621, 1973.
- [6] D. Iakovidis, D. Maroulis and S. Karkanis, A comparative study of color-texture image features, In Proceedings of 12th International Workshop on Systems, Signals & Image Processing (IWSSIP'05), Chalkida, Greece, pp. 203-207, 2005.
- [7] R. Lakmann, BarkTex benchmark database of color textured images, Koblenz-Landau University, ftp://ftphost.unikoblenz.de/outgoing/vision/Lakmann/BarkTex.
- [8] T. Mäenpää and M. Pietikäinen, Classification with color and texture: jointly or separately?, Pattern Recognition, vol. 37, nb. 8, pp. 1629-1640, 2004.
- [9] C. Münzenmayer, H. Volk, C. Küblbeck, K. Spinnler, et T. Wittenberg. Multispectral texture analysis using interplane sum- and

difference-histograms, In German Association for Pattern Recognition Symposium, Springer-Verlag ed., pp. 42-49, 2002.

- [10] T. Ojala, T. Mäenpää, M. Pietikäinen, J. Viertola, J. Kyllönen and S. Huovinen, Outex new framework for empirical evaluation of texture analysis algorithms, In Proceedings of the 16th International Conference on Pattern Recognition (ICPR'02), Quebec, Canada, vol. 1, pp. 701-706, 2002, http://www.outex.oulu.fi/temp/.
- [11] C. Palm, Color texture classification by integrative co-occurrence matrices, Pattern Recognition, vol. 37, nb. 5, pp. 965-976, 2004.
- [12] R. Picard, C. Graczyk, S. Mann, J. Wachman, L. Picard and L. Campbell, VisTex benchmark database of color textured images, Media Laboratory, Massachusetts Institute of Technology (MIT), Cambridge, http://vismod.media.mit.edu/pub/VisTex/VisTex.tar.gz.
- [13] M. Pietikäinen, T. Mäenpää and J. Viertola, Color texture classification with color histograms and local binary patterns, In Proceedings of the 2nd International Workshop on Texture Analysis and Synthesis (Texture'02), Copenhague, Denmark, pp. 109-112, 2002.
- [14] A. Porebski, N. Vandenbroucke and L. Macaire, Iterative feature selection for color texture classification, Proceedings of the 14th IEEE International Conference on Image Processing (ICIP'07), San Antonio, Texas, USA, pp. 509-512, 2007.
- [15] A. Porebski, N. Vandenbroucke and L. Macaire, Neighborhood and Haralick feature extraction for color texture analysis, In Proceedings of the 4th European Conference on Colour in Graphics, Imaging and Vision (CGIV'08), Terrassa, Spain, pp. 316-321, 2008.
- [16] A. Porebski, N. Vandenbroucke, and L. Macaire, Haralick feature extraction from LBP images for color texture classification, In First international Workshops on Image Processing Theory, Tools and Applications (IPTA'08), Sousse, Tunisia, pp. 1-8, 2008.
- [17] A. Porebski, N. Vandenbroucke and L. Macaire, Selection of color texture features from Reduced Size Chromatic Co-occurrence Matrices, In Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA'09), Kuala Lumpur, Malaysia, 2009.
- [18] E.L. Van Den Broek and E.M. Van Rikxoort, Evaluation of color representation for texture analysis, In Proceedings of the Belgium-Dutch Conference on Artificial Intelligence (BNAIC'04), Groningen, The Netherlands, pp. 35-42, 2004.
- [19] 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, vol. 90, nb. 2, pp. 190-216, 2003.
- [20] Q. Xu, J. Yang and S. Ding, Color texture analysis using the waveletbased hidden Markov model, Pattern Recognition Letters, vol. 26, nb. 11, pp. 1710-1719, 2005.

## Author Biography

Alice Porebski received the PhD from the University of Lille in 2009. Since, she is an associate professor in the Ecole d'Ingénieurs du Pasde-Calais. Her research topics concern color texture classification and quality control for industrial applications.

Nicolas Vandenbroucke received the PhD in computer science and control from the University of Lille 1 in 2000. Since 2002, he is an associate professor in the Ecole d'Ingénieurs du Pas-de-Calais. His research interests include color representation and color image analysis applied to segmentation and classification.

Ludovic Macaire is full professor at the University of Lille 1. His research interests deal with color imaging.