# Color Texture Classification by Wavelet Energy Correlation Signatures

G. Van de Wouwer*, S. Livens, P. Scheunders, D. Van Dyck

Vision Lab, Department of Physics, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerpen, Belgium

**Abstract.** In the last decade, multiscale techniques for gray-level texture analysis have been intensively studied. In this paper, we aim on extending these techniques to color images. We introduce wavelet energy-correlation signatures and we derive the transformation of these signatures upon linear color space transformations. Classification experiments demonstrate that the wavelet correlation features contain more information than the intensity or the energy features of each color plane separately. The influence of image representation in color space is evaluated.

## 1 Introduction

For image analysis, color and texture are two of the most important properties, especially when one is dealing with real world images. Classical image analysis schemes only take into account the pixel gray-levels, which represents the total amount of visible light at the pixels position. The performance of such schemes can be improved by adding color information [1]. The color of a pixel is typically represented with the RGB tristimulus values, corresponding to the Red, Green and Blue frequency bands of the visible light spectrum. Color is then a feature in the 3-dimensional RGB color space, which contains information regarding the spectral distribution of light complementary to the gray-level information.

An important topic when processing color images is their representation. The RGB representation is frequently being transformed into other color spaces [2] [3]. The performance of an image analysis system can strongly depend on the choice of the color representation [4] [5]. However, there does not appear to be a systematic means of determining an optimum color-coordinate system for a particular task.

In the analysis of color images, the description of image regions has mainly been performed using color histograms [3] [6]. However, they no longer suffice when local spatial correlations are important to characterize a region. The extra information needed to adequately describe the image regions is commonly known as "texture". Texture has been studied extensively and many texture analysis schemes have been proposed [7]. The fundamental property which they all have in common is that they exploit local spatial interactions between pixels.

---

* corresponding author: email: wouwer@ruca.ua.ac.be

A rather limited number of systems use combined information of color and texture, and even when they do, both aspects are mostly dealt with using separate methods [8] [9]. It is only recently that attempts are being made to combine both aspects in a single method, by extending gray-level texture analysis methods to color images [10] [11]. This combination can be made more formal by defining "color-texture" as "the set of local statistical properties of the colors of image regions". Efficient characterization of color texture requires the exploitation of spatial correlations as well as correlations between color bands.

The importance of a joint color-texture characterization is expected to grow rapidly in the near future, e.g. for indexing image databases. At present time, the color extensions of several major texture analysis methods are still unexplored. We will investigate one of them, based on multiresolution decomposition. These techniques give rise to an interesting class of texture analysis methods. Strong arguments for their use can be found in psychovisual research, which offers evidence that the human visual system processes images in a multiscale way [12]. Wavelets provide a convenient way to obtain a multiresolution representation [13], from which texture features are easily extracted [14] [15] [16].

We propose a scheme for the characterization of colored texture images. Feature extraction using wavelet decomposition is described. Wavelet correlation signatures are defined which contain the energies of each color plane and the cross-correlation between different planes. While the first have already successfully been used for texture characterization, the latter represent the coupling between texture and color. We will show that these features transform linearly upon linear color space transformation. The experiments will demonstrate the usefulness of correlation signatures as texture features. The influence of the choice of color space representation on classification performance will be investigated.

## 2　Wavelet Signatures

The (continuous) wavelet transform of a 1-D signal $f(x)$ is defined as

$$(W_a f)(b) = \int f(x) \psi_{a,b}^{\star}(x) dx \quad \text{with} \quad \psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi(\frac{x-b}{a}) \qquad (1)$$

The *mother wavelet* $\psi$ has to satisfy the admissibility criterion to ensure that it is a localized zero-mean function. (1) can be discretized by restraining $a$ and $b$ to a discrete lattice ($a = 2^n$, $b \in \mathcal{Z}$). Typically some more constraints are imposed on $\psi$ to ensure that the transform is non-redundant, complete and constitutes a multiresolution representation of the original signal. This has led to an efficient real-space implementation of the transform using quadrature mirror filters.

The extension to the 2-D case is usually performed by applying a separable filter bank to the image:

$$L_n(\mathbf{b}) = [H_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \qquad (2)$$

$$D_{n1}(\mathbf{b}) = [H_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \qquad (3)$$

$$D_{n2}(\mathbf{b}) = [G_x * [H_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{4}$$

$$D_{n3}(\mathbf{b}) = [G_x * [G_y * L_{n-1}]_{\downarrow 2,1}]_{\downarrow 1,2}(\mathbf{b}) \tag{5}$$

where $\mathbf{b} \in \mathbb{R}^2$, $*$ denotes the convolution operator, $\downarrow 2,1$ ($\downarrow 1,2$) sub-sampling along the rows (columns) and $L_0 = I(\mathbf{x})$ is the original image. $H$ and $G$ are a low and bandpass filter respectively. $L_n$ is obtained by low pass filtering and is therefore referred to as the low resolution image at scale $n$. The $D_{ni}$ are obtained by bandpass filtering in a specific direction and thus contain directional detail information at scale $n$; they are referred to as the *detail images*. The original image $I$ is thus represented by a set of subimages at several scales; $\{L_d, D_{ni}\}_{i=1,2,3}^{n=0,\dots,d-1}$ which is a *multiscale representation of depth $d$* of the image $I$.

The energy of a subimage $D_{ni}$ is defined as

$$E_{ni} = \int (D_{ni}(\mathbf{b}))^2 d\mathbf{b} \tag{6}$$

The *wavelet energy signatures* $\{E_{ni}\}_{n=0,\dots,d-1,i=1,2,3}$ reflect the distribution of energy along the frequency axis over scale and orientation and have proven to be very useful for gray-level texture characterization. Since most relevant texture information has been removed by iterative low pass filtering, the energy of the low resolution image $L_d$ is generally not considered a texture feature.

The most straightforward extension of the wavelet energy signatures to color images is to transform each color plane separately and extract the energies of each transformed plane; i.e. replace $I$ by the R,G and B-plane consecutively in (2)-(6). We denote such an energy by $E_{ni}^{X_j}$ where the $X_j$ indicates the color plane. This triples the amount of features w.r.t. the gray-level case.

Let us define

$$C_{ni}^{X_j X_k} = \int D_{ni}^{X_j}(\mathbf{b}) D_{ni}^{X_k}(\mathbf{b}) d\mathbf{b} \tag{7}$$

and call the set $\{C_{ni}^{X_j X_k}\}_{n=0,\dots,d-1,i=1,2,3}^{j,k=1,2,3,\,j \le k}$ the *wavelet covariance signatures*. They include the energies for $j = k$; the others represent the covariance between different color planes and consequently the coupling between the color and texture properties of the image.

The covariance signatures, however, are by definition proportional to the energies. They are normalized to remove this redundant information:

$$\tilde{C}_{ni}^{X_j X_k} = \begin{cases} E_{ni}^{X_j} & j = k \\ \dfrac{C_{ni}^{X_j X_k}}{E_{ni}^{X_j} E_{ni}^{X_k}} & j \ne k \end{cases} \tag{8}$$

The features $\{\tilde{C}_{ni}^{X_j X_k}\}_{n=0,\dots,d-1,i=1,2,3}^{j,k=1,2,3,\,j \le k}$ are the *wavelet correlation signatures*.

# 3   Color Space Transforms

For compression purposes, transformations to different color spaces are often employed to achieve image bandwidth reduction without significantly degrading image quality. However, since our goal is to efficiently characterize texture, the choice of color space should enable extraction of useful features rather than visual image representation. Non-linear transforms are mainly employed to obtain a color space in which the 3 coordinates have an intuitive meaning (mostly a luminance, a saturation and a hue component) [17]. They typically introduce some non-removable singularities, which is very impractical for further processing. We will therefore limit ourselves to linear color space transforms, i.e.

$$X' = MX \tag{9}$$

where $X = (X_1(\mathbf{x})\ X_2(\mathbf{x})\ X_3(\mathbf{x}))^\tau$ contains the original components of the signal ($^\tau$ means transpose), $M$ is a 3 by 3 invertible transformation matrix and $X'$ contains the transformed signal.

Three particular color space transforms (for which $X = (R\ G\ B)^\tau$) are:

$$M_1 = \begin{pmatrix} 0.405 & 0.116 & 0.133 \\ 0.299 & 0.587 & 0.114 \\ 0.145 & 0.827 & 0.627 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & 0.312 \end{pmatrix} \quad M_3 = \begin{pmatrix} 0.333 & 0.333 & 0.333 \\ 0.500 & 0.000 & -0.500 \\ -0.500 & 1.000 & -0.500 \end{pmatrix}$$

The first transforms RGB to the UVW-space (V=Y=luminance). This is a "perceptually uniform" space constructed so that equal changes in the space are experienced as equal changes in color by human perception. $M_2$ represents the YIQ-space. The Y signal is the image luminance and the I and Q signals carry the chrominance information. The last one ($M_3$) represents the K-L space (Karhunen-Loève transform), which transforms an image to an orthogonal basis in which the axes are statistically uncorrelated, and in that sense decorrelates the information present in RGB space.

*Effect of linear color transform on the wavelet signatures*
We now investigate how the wavelet covariance signatures transform under a linear color space transform. Let us fix $n$ and $i$ and rewrite (1) in vector notation for a color image $X = (X_1(\mathbf{x})\ X_2(\mathbf{x})\ X_3(\mathbf{x}))^\tau$:

$$(W_{2^n,i}X)(\mathbf{b}) = \int X(\mathbf{x})\varphi_{2^n,\mathbf{b}}^{i\,\star}(\mathbf{x})d\mathbf{x} \tag{10}$$

Define $C_{ni}$ as a (symmtric) matrix with the wavelet covariance signatures as elements:

$$C_{ni} = \int (W_{2^n,i}X)(\mathbf{b})\left((W_{2^n,i}X)(\mathbf{b})\right)^\tau d\mathbf{b} \tag{11}$$

After a color space transformation $X' = MX$ the covariance signatures become:

$$C'_{ni} = \int (W_{2^n,i}X')(\mathbf{b}) \left((W_{2^n,i}X')(\mathbf{b})\right)^\tau d\mathbf{b}$$

$$= \int \left( \int X'(\mathbf{x})\varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})d\mathbf{x} \right) \left( \int X'(\mathbf{x})\varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})d\mathbf{x} \right)^\tau d\mathbf{b}$$

$$= \int ( \int MX(\mathbf{x})\varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})d\mathbf{x})( \int MX(\mathbf{x})\varphi^{i*}_{2^n,\mathbf{b}}(\mathbf{x})d\mathbf{x})^\tau d\mathbf{b}$$

$$= MC_{ni}M^\tau \qquad (12)$$

Or, explicitly:

$$C^{X'_j,X'_k}_{ni} = \sum_{r,s=1}^{3} m_{jr}m_{ks}C^{X_r X_s}_{ni} \qquad (13)$$

For the energies this means (taking the RGB-space for $X$):

$$E^{X'_j}_{ni} = C^{X'_j X'_j}_{ni} = \sum_{r,s=1}^{3} m_{jr}m_{js}C^{X_r X_s}_{ni}$$

$$= m^2_{j1}E^R_{ni} + m^2_{j2}E^G_{ni} + m^2_{j3}E^B_{ni} + 2m_{j1}m_{j2}C^{RG}_{ni}$$

$$+ 2m_{j1}m_{j3}C^{RB}_{ni} + 2m_{j2}m_{j3}C^{GB}_{ni} \qquad (14)$$

These formulas offer an interesting insight in the effect of linear color space transform on the wavelet signatures. (13) shows that a linear color space transform implies a linear transform of the covariance signatures. However, from (14) it follows that this is not true for the energy signatures. The first 3 terms reveal that the "new" energy features are linearly obtained from the "old" ones; the next 3 terms however depend on the covariances between the R, G and B planes for the same subimage. There is no clear connection between the energies in the original and transformed color space; to compute the latter the wavelet covariance signatures are required.

(14) also shows that performing a simple linear transform from RGB space to another color space results in a clearly different feature set. Hence, the quality of the features (i.e. their ability to characterize and discriminate between color textures) shall be heavily dependent on the choice of color space. This shall be demonstrated in the experimental section.

When one experiments using several color transforms, a practical advantage of the relation (13) comes into play. It is sufficient to perform the wavelet transform once (for the R,G and B planes) and to compute the covariance signatures. The new wavelet signatures are then obtained using (13) without the need of performing several wavelet transforms.

For the correlation signatures, the simple relation (13) does not hold. To transform correlation signatures into other color spaces, it is therefore convenient to transform the covariance signatures first and to normalize them afterwards.

# 4    Classification Methods

A $k$-nearest neighbor classifier ($k$-nn) [18] is used to estimate recognition performance. Since the emphasis in this work is on the feature extraction stage, $k$-nn provides an efficient and robust classification scheme for evaluation of recognition rates and comparison of feature sets.

Recognition rate is estimated by the leave-one-out method. This method sequentially picks each available data sample and classifies it (by the $k$-nn rule) using the remaining samples. Each available sample is thus employed once as a test sample. The *recognition rate* is estimated by counting the total number of samples classified correctly.

Well known in pattern recognition literature is the curse of dimensionality phenomenon, which dictates that classification performance not necessarily increases with an increasing number of features (given a fixed amount of data samples). Therefore, given a feature extraction scheme and a finite number of training instances, there exists an optimal number of features for a particular task. This becomes inherently important when dealing with colored images, since the number of extracted features is much larger compared to the gray-level case. Therefore, it is crucial to adopt a feature selection (or extraction) scheme to find a (sub-)optimal set of features. In this work the Floating Forward Feature Selection scheme (FFFS) [19] is adopted. This algorithm is initialized by taking the best feature ("best" is defined here as giving the best recognition performance). The selection then continues by iteratively adding (or deleting) a feature in each step to obtain a subset of all available features which gives the highest classification performance.

# 5    Experiments and Conclusion

30 real-world (512x512) RGB color-images from different natural scenes [20] were selected: Bark0, Bark4, Bark6, Bark8, Bark9, Brick1, Brick4, Brick5, Fabric0, Fabric4, Fabric7, Fabric9, Fabric11, Fabric13, Fabric16, Fabric17, Fabric18, Food0, Food2, Food5, Food8, Grass1, Sand0, Stone4, Tile1, Tile3, Tile7, Water6, Wood1 and Wood2. A database of 1920 color image regions of 30 texture classes was constructed by dividing each image into 64 non-overlapping 64x64 subimages. The following classification experiments were conducted:

1. Intensity (gray-level) images were generated by computing the luminance, hereby discarding color information. A wavelet transform of depth 4 was performed and energy signatures were computed for each of the 12 detail images. (total: 12 features)
2. Each R, G and B component was wavelet transformed (depth 4) and energy signatures were computed from each detail image. (total: 36 features)
3. Each R, G and B component was wavelet transformed (depth 4) and correlation signatures were computed from each detail image. (total: 72 features)
4. 72 correlation signatures were computed using (13) for the 3 color spaces mentioned in section 3: a) UVW space, b) YIQ space, c) K-L space
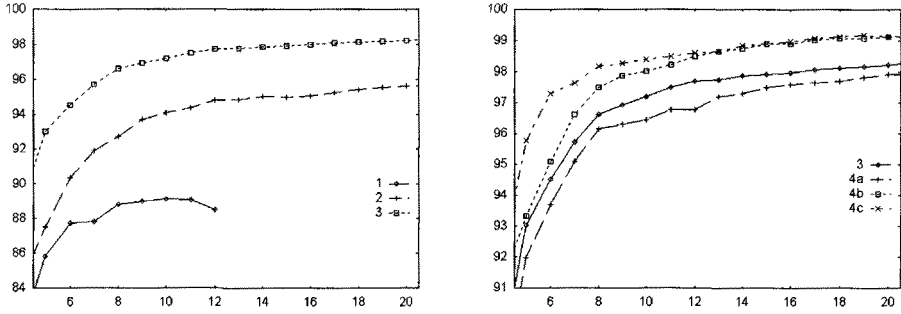
**Fig. 1.** recognition rate (%) versus feature set dimensionality graphs. a) 1. intensity 2. energy RGB 3. correlation RGB. b) correlation signatures in different color spaces: RGB (3), UVW (4a), YIQ (4b), K-L(4c).

Classification results are depicted in Fig. 1. One observes that the recognition rates saturate about a dimensionality of 10, at which point the error rates for each texture class were investigated. We found that recognition for Fabric0-7-17-18, Food5, Sand0, Tile7 and Wood2 was 100% for all classifiers. Recognition rates for Fabric9-16 and Water6 were 97-100% and did not differ much between classifiers. This shows that intensity alone contains sufficient information to characterize some textures, but fails to do so on others. Fig. 1,a) shows that adding color information does significantly increase recognition performance. Comparing curves 2 and 3 in Fig. 1,a) shows that the correlation signatures offer a clear advantage over the energies.

Fig. 1,b) compares the performance of the wavelet correlation signatures in different colour spaces, which demonstrates that recognition performace is color space dependent. It is apparent that the recognition rate for the UVW space is lower than in the K-L and YIQ spaces and also lower than for the RGB space. Recognition performance is thus indeed color space dependent. Overall, the best results are obtained with the K-L transform.

The conducted experiments demonstrate that color texture can adequately be described by the wavelet correlation signatures. These features are not only suited for image classification, but can easily be employed for other color texture analysis tasks. For instance, for segmentation the wavelet signatures are computed over a (small) local window centered on each pixel of the image, resulting in one feature vector per pixel. Each pixel is then assigned to a particular image region, e.g. by (unsupervised) clustering techniques in the space of feature vectors.

# References

1. G.J. Klinker, S.A. Shafer, and T. Kanade. A physical approach to color image understanding. *Int. J. Comput. Vision*, 4:7–38, 1990.
2. G. Wyszecki and W. S. Stiles. *Color Science, Concepts and Methods, Quantitative Data and Formulas, 2nd ed.* J. Wiley and Sons, New York, 1982.
3. Q-T Luong. Color in computer vision. In C.H. Chen, L.F. Pau, and P.S.P. Wang, editors, *Handbook of Pattern Recognition & Computer Vision*, chapter 2.3, pages 311–368. World Scientific, Singapore, 1993.
4. Y.I. Ohta, T. Kanade, and T. Sakai. Color information for region segmentation. *Computer graphics and image processing*, 13:222–241, 1980.
5. W.K. Pratt. Spatial transform coding of color images. *IEEE Trans. Comm. Tech.*, 19(6):980–992, 1971.
6. C.L. Novak and S.A. Shafer. Methods for estimating scene parameters from color histograms. *J. Opt. Soc. Am. A.*, 11:3020–3036, 1994.
7. T.R. Reed and J.M.H. du Buf. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372, 1993.
8. D. Lee, R. Barber, W. Niblack, M. Flickner, J. Hafner, and D. Petkovic. Indexing for complex queries on a query-by-content image database. In *Proc. of the 12nd IAPR Int. Conf. on Pattern Recognition*, volume I, pages 142–146, Jerusalem, Israel, 1994.
9. J. R. Smith and S. Chang. Local color and texture extraction and spatial query. In *IEEE Proc. Int. Conf. on Im. Proc.*, volume III, pages 1011–1014, 1996.
10. T. Caelli and D. Reye. On the classification of image regions by colour texture and shape. *Pattern Recognition*, 26(4):461–470, 1993.
11. D.K. Panjwani and G. Healey. Markov random field models for unsupervised segmentation of textured color images. *IEEE Trans. Pattern Anal. Machine Intell.*, 17(10):939–954, 1995.
12. T.S. Lee. Image representation using 2d gabor wavelets. *IEEE Trans. Patt. Anal. Machine Intell.*, 18(10):959–971, 1996.
13. S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Patt. Anal. Machine Intell.*, 11(7):674–693, 1989.
14. T. Chang and C.-C.J. Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Im. Proc.P*, 2(4), 1993.
15. G. Van de Wouwer, P. Scheunders, D. Van Dyck, M. De Bodt, F. Wuyts, and P.H. Van de Heyning. Wavelet-filvq classifier for speech analysis. In *Proc. Int. Conf. Patt. Rec., Vienna*, pages 214–218, 1996.
16. O. Pichler, A. Teuner, and B.J. Hosticka. A comparison of texture feature extraction using adaptive gabor filtering, pyramidal and tree structured wavelet transforms. *Pattern Recognition*, 29(5):733–742, 1996.
17. H. Levkowitz and G. T. Herman. Glhs: A generalized lightness, hue and saturation model. *CVGIP*, 11(11):3011–3019, 1994.
18. P.A. Devijver and J. Kittler. *Pattern recognition: a statitical approach.* Prentice/Hall, Englewood Cliffs, New Jersey, 1982.
19. P. Pudil, J. Novovicova, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.
20. VisTex. Color image database. http://www-white.media.mit.edu/vismod/-imagery/VisionTexture, 1995. MIT Media Lab.