An Optimum 2D Color Space for Pattern Recognition

Victor-Emil Neagoe
Faculty of Electronics, Telecommunications and Information Technology,
"POLITEHNICA" University of Bucharest, Romania, E-mail: victoremil@gmail.com

Abstract **-** *This paper presents an optimum color conversion from the 3D RGB space into a 2D selected space to the purpose of pattern recognition. The method is based on the Karhunen-Loève transform (KLT), also known as Principal Component Analysis (PCA). The resulted 2D space is defined by the two color components (called C_1 and C_2), corresponding to the two largest eigenvalues of the RGB pixel covariance matrix. Using the above color projection technique, we propose a color face recognition system based on feature fusion of the C_1 and C_2 components and a concurrent neural network classifier. The proposed system is experimented for a color face database containing 3520 color images of 151 subjects. We also present a color image segmentation using pixel clustering in the 2D color space by means of a self-organizing neural network. The new 2D color projection model may have wide applications in the areas of color-based pattern recognition.*

Keywords: *optimum 2D color conversion, color-based pattern recognition, color face recognition, color image segmentation*

1 Introduction

Like humans, the artificial intelligence systems use *color*, shape and texture for pattern recognition. There are a lot of systems for pictorial content representation and recognition based on color features.

Color image segmentation is a significant research area, useful in many applications. From the segmentation results, it is possible to identify regions of interest and objects in the scene, which is very beneficial to the subsequent image analysis.

All about the world, governments and private companies are putting *biometric technology* at the heart of ambitious projects, ranging from access control and company security to high-tech passports, ID cards, driving licenses, and company security. One of most important areas of biometric technology is *face recognition***.** A common feature found in almost all technical approaches proposed for *face recognition* is the use of only the luminance associated to the face image. Although the majority of images are recorded in the color format nowadays, most face recognition systems convert the color information to luminance component data and do no use color information. One of the key challenges in face recognition lies in determining the contribution of different cues to the system performance and one of these cues is the *color* attribute.

We further present an approach to improve the color-based pattern recognition performance by optimizing the color conversion. In [8], a neural model is given, for exploiting both *spectral* and also spatial image correlation, to reduce space dimensionality of color pictures. Recently, Jones and Abbott [1] performed a color conversion of the R, G, B components into the optimized monochrome form (instead of luminance) for face recognition, using the Karhunen-Loève transformation (KLT). We extend their approach by proposing and evaluating the transformation of the 3D RGB space into a 2D optimized space.

Then we propose a color face recognition system, where the images belonging to the face data base were projected in the previously mentioned KLT 2D color space (of components C_1 and C_2). For feature extraction, one chooses the Principal Component Analysis (PCA) model for each of the C_1 and C_2 channels. The next stage corresponds to *feature fusion*. The last processing stage means the application of the multiple neural system called CSOM (Concurrent Self-Organizing Maps) [7]. For comparison, we considered two scheme variants of color face recognition based on the 3D RGB color space. The systems are evaluated using the Essex color face database (151 selected subjects).

The application of the new 2D color projection techniques for color image segmentation is also considered. Using a 2D color optimized representation, proposed in this paper, instead of the 3D color space, one can significantly reduce the computational effort, by preserving the information content.

2 Color Conversion from RGB Space into an Optimum 2D Space for Pattern Recognition

Consider the color pixels in a given image as 3D vectors

$$P(x,y) = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix},$$

where R(x, y), G(x, y) and B(x, y) are the red, green and blue components of the pixel of co-ordinates (x, y).

We assume that color images exhibit features that can be useful in the conversion from a 3D full color space representation to the 2D space. For color conversion, we have chosen the Karhunen-Loève transformation (KLT), also known as Principal Component Analysis (PCA), by exploiting the correlation of the R, G, and B color channels. It is an *optimum projection solution*, by minimizing the mean square error for vector dimensionality reduction, when one projects the 3D RGB space into the 2D *KLT color space* with uncorrelated axes.

To deduce the KLT matrix, one firstly computes the covariance matrix of the color pixels (represented as 3D vectors). Then, one computes the eigenvalues of the covariance matrix. Finally, we deduce the two eigenvectors, corresponding to the largest two eigenvalues. Thus, one obtains the KLT matrix K

$$K = \begin{bmatrix} A^T \\ B^T \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix},$$

where

$$A = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \qquad \text{and } B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

(A and B are the eigenvectors of the covariance matrix corresponding to the two largest eigenvalues and T denotes transposition).

Then, the projection of the 3D color vector P(x,y) in the 2D space is the vector C(x,y)

$$C(x,y) = \begin{bmatrix} C_1(x,y) \\ C_2(x,y) \end{bmatrix},$$

given by the equation

$$C(x,y) = = K \bullet P(x,y).$$

2.1 Example 1

One assumes the image ("peppers") in Fig. 1 (a), having 256 x 256 pixels with 24 bits/pixel.

The eigenvalues of the covariance matrix are

$$\lambda_1 = 7334.6; \ \lambda_2 = 1803.0 \ ; \ \lambda_3 = 347.8.$$

For the above example, by retaining first two largest eigenvalues, one deduces that the projection error is of 3.66% only!

The corresponding eigenvectors defining color KLT are

A^T = (0.2490 0.8428 0.4772)
B^T = (0.9492 -0.3102 0.0525).

(a) (b)

Fig. 1. (a) Original "peppers". (b) Reconstructed "peppers" from 2D KLT color space.

Thus, one can perform the projection in the 2D space. In Fig. 1(b), the reconstructed version of the image 1(a) from 2D space is given.

2.2 Example 2

We considered the original RGB image in Fig. 2(a) (from Berkeley segmentation data set) and the reconstructed version from its 2D KLT projection (Fig. 2(b)). One can remark that the reconstructed picture is very similar to the original.

(a) (b)

Fig. 2. (a) Original "Berkeley". (b) Reconstructed "Berkeley" from 2D KLT color space.

3 Color Image Segmentation Using 2D Pixel Clustering

We further apply the previous color conversion method for *color image segmentation*. We perform clustering of color pixels represented as 2D vectors (by the corresponding C_1 and C_2 color components).

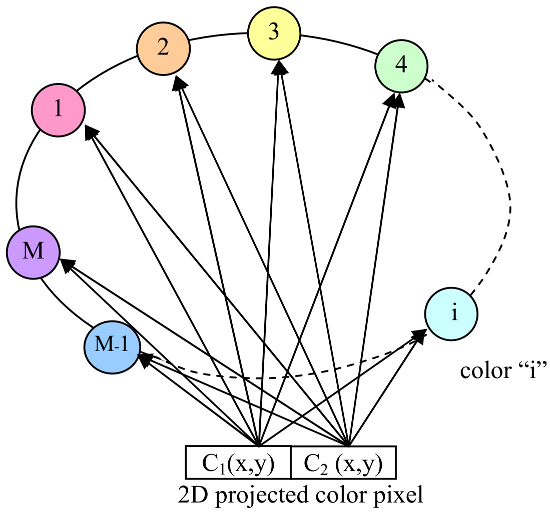

Fig. 3. Circular self-organizing map for color pixel clustering in the 2D KLT color space.

The above vectors are applied to the input of a Self-Organizing Map (SOM), also called Kohonen neural network, having a circular architecture with M output neurons (Fig. 3). Each output neuron is a potential prototype of a color class, so that the maximum number of color classes (M), is given by the number of output neurons. The system stores the correspondences between input pixels (2D vectors) and the index of the corresponding winning neuron, so that we can assigned to each class a *natural* color that is the average of the colors characterizing the pixels assigned to that class. The *pseudo-color* representation can also be used.
We have assumed as input the color image “peppers” given in Fig. 1(a). The result of segmentation (by pixel clustering in the 2D space) is the representation of the considered picture by maximum M color classes (see Fig. 4). The cases of M=10 in Fig. 4(a, c) and M=5 in Fig. 4(b, d) are considered.

For comparison, we experimented the pixel clustering in the 3D RGB space using the same kind of neural network (Fig. 3), but using a 3D input (R, G, B). The results of simulation are given in Fig. 5.

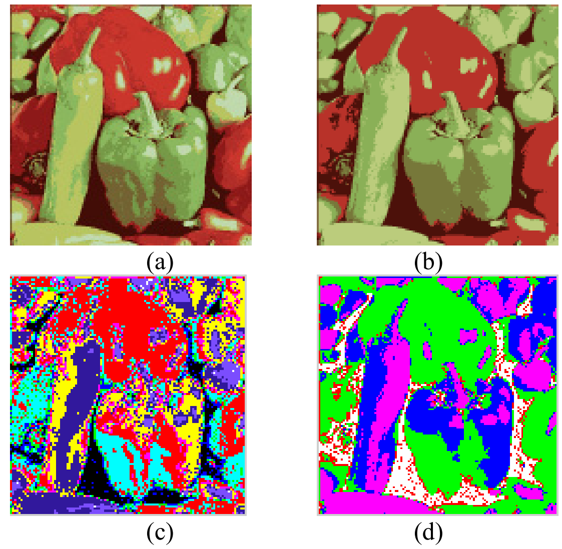

Fig. 4. Segmented “peppers” by 2D pixel clustering in M color classes with a circular SOM. The inputs are 2D vectors of (C_1, C_2) color components. Use M output neurons: (a) natural colors, M=10; (b) natural colors, M=5; (c) pseudo-colors, M=10; (d) pseudo-colors, M=5.

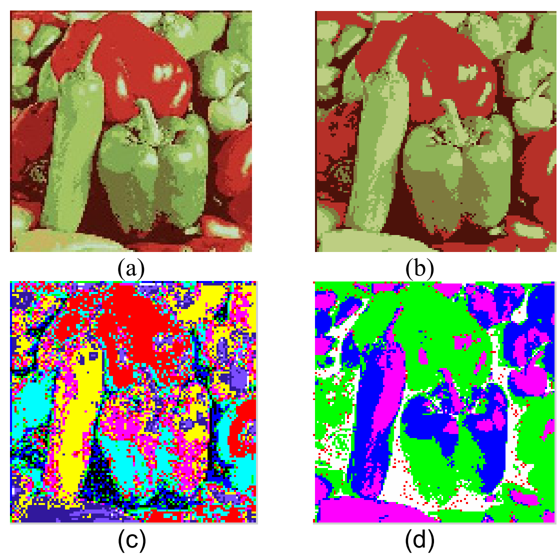

Fig. 5. Segmented “peppers” by 3D pixel clustering in M color classes with a circular SOM. The inputs are 3D vectors of (R, G, B) color components. Use M output neurons: (a) natural colors, M=10; (b) natural colors, M=5; (c) pseudo-colors, M=10; (d) pseudo-colors, M=5.

The advantage of 2D representation over RGB is that by performing a color image segmentation in a space with 2 dimensions (instead of 3), one can obtain an equivalent clustering quality with a reduced computational effort.

4 Face Recognition in the 2D Color Space

4.1 Feature Fusion Model

Using the proposed color projection model, a new system of color face recognition is proposed (Fig. 6). It contains the following processing stages:

1) Color conversion of the R, G, and B components into the two optimized new components C_1 and C_2, according to the KLT
2) Principal Component Analysis (PCA) for each of the two color channels (C_1 and C_2)
3) Feature fusion (amalgamation of the eigen-components of the two channels)
4) Neural network classification. The final processing stage consists of a set of Concurrent Self Organizing Maps (CSOM) [7] shown in Fig. 7.

Concurrent Self-Organizing Maps (CSOM) is a collection of small SOMs, which use a global *winner-takes-all* strategy. Each network is used to correctly classify the patterns of one class only and the number of networks equals the number of classes.

The CSOM training technique is a supervised one, but for any individual net the SOM specific training algorithm is used. We built "n" training patterns sets and we used the SOM training algorithm independently for each of the "n" SOMs. The CSOM models *for training and classification* are shown in Figs. 7 (a) and (b).

For comparison, we further consider two scheme variants of color face recognition, where their inputs are 3D vectors (RGB pixels). In Fig. 8, one can see a model based on the independent processing of the R, G, and B channels. After PCA and neural classification, we can follow one of the color decision or we can perform a decision fusion (for example, by vote). The system in Fig. 9 uses the fusion of the eigen-features corresponding to the R, G, and B color components, followed by the neural classifier.

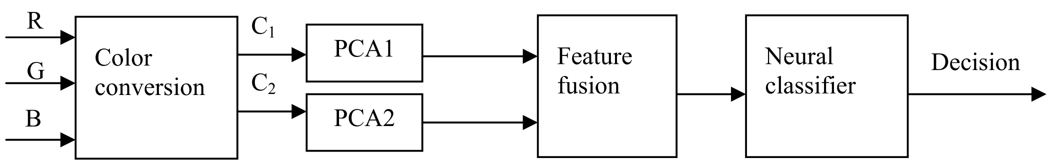

Fig. 6. Color face recognition with color conversion and feature fusion (*using as inputs 2D projected pixels*).

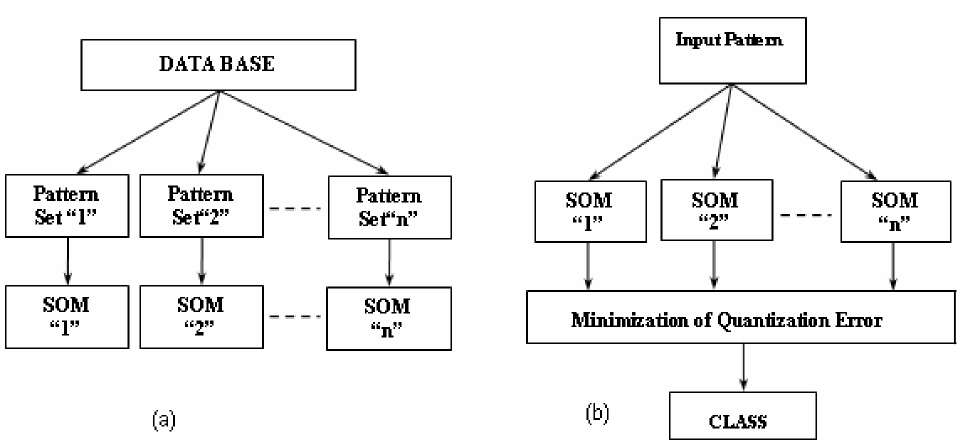

Fig. 7. (a) The CSOM model (training phase). (b) The CSOM model (classification phase).

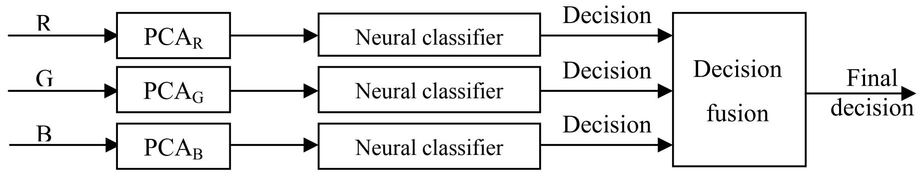

Fig. 8. Color face recognition using the R, G, B components and a decision fusion.

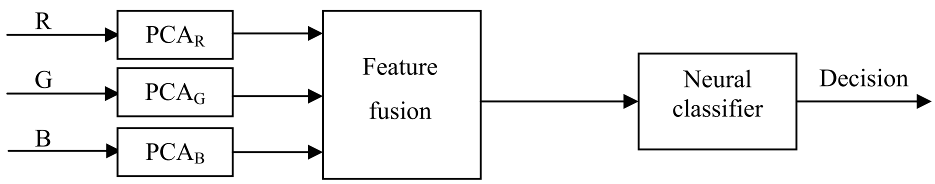

Fig. 9. Color face recognition using feature fusion of the R, G, B channels.

4.2 Experimental Results

We have used the color face database provided by Dr. Libor Spacek, Depart. of Computer Science, University of Essex, U.K. We considered 3020 images from this database, corresponding to 151 subjects, where each subject is represented by 20 pictures (10 images being chosen for training and the other 20 for test). Any picture has 200 x 180 pixels, in RGB format (with 24 bits/pel).

The face database contains images of people of various racial origins, most of them being of 18-20 year old, but some older individuals are also present (Fig. 10).

We have considered both the original images selected from data base and also the corresponding intentionally degraded ones (Fig. 12). The experimental results are given in Tables 1-2 and Figs. 14-17.

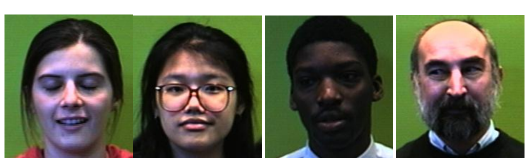

Fig. 10. Several images belonging to the Essex database.

The eigenvalues of the color pixel covariance matrix for the training set of 1510 face images are

$$\lambda_1 = 8140.67;\ \lambda_2 = 984.34\ ;\ \lambda_3 = 223.35.$$

One deduces that the projection error (corresponding to least eigenvalue) is of 2.39% only!

The corresponding eigenvectors defining the color KLT are

$$A^T = (0.6411\ 0.5568\ 0.5282)$$

$$B^T = (0.1273\ -0.7558\ 0.6423).$$

In Fig. 11 (b) one can see the reconstruction of image (11.a) from the 2D KLT color space.

(a) (b) (c) (d)

Fig. 11. (a) Original "Ekavaz". (b) Reconstruction of (a) from 2D KLT color space. (c) Reconstruction of (a) from 1D KLT color space. (d) Luminance component of (a).

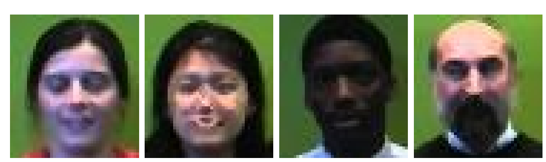

Fig. 12. Intentionally degraded images.

The subjective effect of retaining a various number of eigen-features from the color image can be evaluated in Fig. 13.

Fig. 13. (a) Original image. (b) Reconstructed image from 50 eigen-features/each (R, G, B).
(c) 100 features. (d) 500 features.

Table 1. Recognition score for the test lot of 1510 original color images

Number of features/color component		10	30	50	70	90	100	150	200	300	500	1000
RGB	Feature fusion	97.22	98.34	98.68	98.81	98.68	98.74	98.87	98.94	99.34	99.87	99.87
	Red	94.7	98.08	98.15	98.54	98.48	98.61	98.54	98.87	98.94	99.27	99.34
	Green	95.3	97.88	98.68	98.34	98.48	98.54	98.48	98.74	99.14	99.06	99.67
	Blue	94.1	97.75	98.15	98.21	98.01	97.88	98.01	98.21	98.34	98.74	98.74
	Decision fusion	95.23	98.08	98.54	98.68	98.48	98.48	98.48	98.81	98.87	99.4	99.47
(C_1,C_2)	Feature fusion	97.28	98.94	99.00	99.00	99.27	99.14	99.34	99.47	99.54	99.8	99.8
	C_1	95.1	98.21	98.61	98.68	98.74	98.81	98.81	99	99.4	99.4	99.4
	C_2	95.89	98.08	98.15	98.21	98.34	98.28	98.34	98.34	98.68	98.81	98.87

Table 2. Recognition score for the test lot of 1510 degraded color images

Number of features/color component		10	30	50	70	90	100	150	200	300	500	1000
Luminance		96.25	98.61	98.94	98.94	99	99	99	99.14	99.27	99.54	99.54
RGB	Feature fusion	97.4	98.48	99.00	99.00	98.94	98.94	99.14	99.47	99.54	99.87	99.87
	Red	95.36	98.61	98.87	98.81	98.87	98.94	98.94	99.07	99.27	99.6	99.54
	Green	95.7	98.34	99.47	99.07	99.07	99.14	99.47	99.54	99.87	99.87	99.87
	Blue	95.7	98.15	98.61	98.61	98.34	98.34	98.34	98.61	98.81	99.00	99.07
	Decision fusion	95.96	98.34	99.07	99.00	98.81	98.87	99.07	99.21	99.47	99.8	99.74
(C_1,C_2)	Feature fusion	97.75	99.21	99.21	99.34	99.21	99.47	99.47	99.67	99.8	99.8	99.8
	C_1	95.96	98.61	99.00	99.00	98.94	99.00	99.00	99.21	99.47	99.74	99.74
	C_2	96.29	98.21	98.48	98.48	98.61	98.61	98.68	98.68	98.87	99.07	99.21

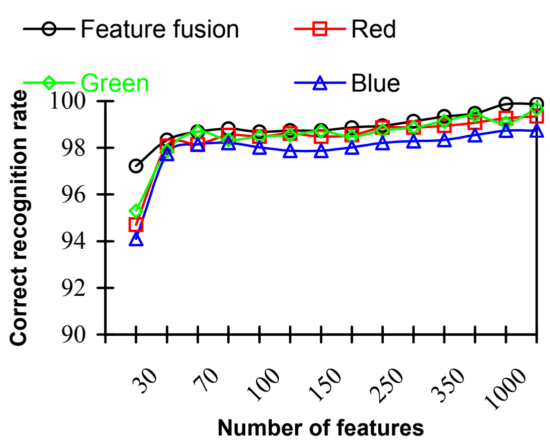

Fig. 14. Recognition score for the systems given in Figs. 8- 9.

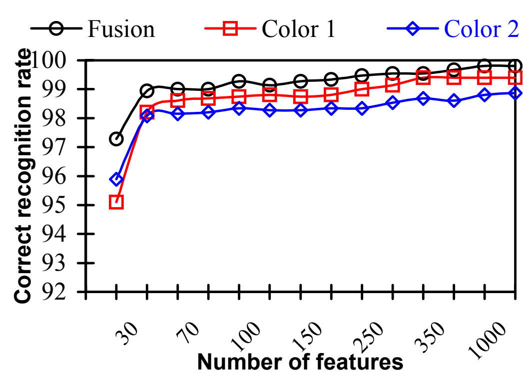

Fig. 15. Recognition score for the system given in Fig. 6.

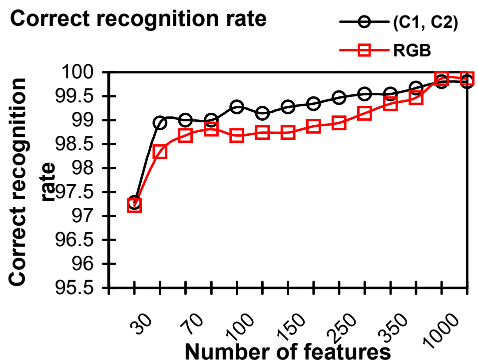

Fig. 16. Comparison of (R, G, B) and (C_1, C_2) best results.

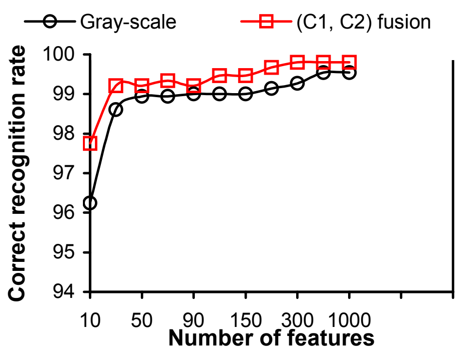

Fig. 17. Comparison of gray and 2D color image recognition performance for degraded images.

5 Concluding Remarks

1. We present a model of 2D color image representation for pattern recognition, using the KLT to project the 3D RGB space into an optimum color plane.
2. The mean square error of color dimensionality reduction (from 3 to 2) is about 3% only, for the considered applications.
3. Using the above 2D color optimized representation, instead of the 3D color space, one can significantly reduce the computational effort for color image processing, by preserving almost all information content.
4. The model has exciting applications for color face recognition and color image segmentation.
5. One proposes a color face recognition system in the 2D color space, using feature fusion and a multiple neural module classifier. We compare this model with two schemes, based on 3D color input vectors.
6. Best results of color face recognition correspond to the new model (feature fusion of the C_1 and C_2 color components and concurrent neural classifier, shown in Fig. 6). This variant is superior both to the feature fusion of R, G, B components and also to the decision fusion of the same color channels.
7. One can remark the role of *color* for face recognition in the case of degraded images (see comparison between gray-scale (luminance) images and (C_1,C_2) color component fusion in Table 2 and Fig. 17).
8. In the case of degraded images, by retaining only C_1 component, one obtains better results than using the luminance.
9. An application of color image segmentation in the 2D color space using neural pixel clustering (with a circular SOM) is also given.
10. The proposed 2D color conversion model may have wide applications in the areas of color-based pattern recognition.

6 References

[1] C. F. Jones III and A. L. Abbott, “Optimization of color conversion for face recognition”, *EURASIP Journal on Applied Signal Processing*, nr. 4, 2004, pp. 522-529.

[2] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, “Color by correlation: a simple, unifying framework for color constancy”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1209-1221, Nov. 2001.

[3] A. Yip, P. Sinha, “Role of color for face recognition”, AI Memo 2001-035, Dec. 2001, Massachusetts Institute of Technology, Cambridge, USA.

[4] L. Torres, J. Y. Reutter, L. Lorente, "The importance of the color information in face recognition”, *Proceedings of the IEEE Conference on Image Processing, ICIP’99*, October 25-28, 1999, Kobe, Japan, vol. 3, pp. 627-631.

[5] J. Aguilar, “A color pattern recognition problem based on the multiple classes random neural network model”, *Neurocomputing 61* (2004), pp. 71-83.

[6] P. Gonzales and R. Woods, *Digital image processing,* Addison Wesley, 1992.

[7] V. Neagoe, A. Ropot, ”Concurrent self-organizing maps for pattern classification”, *Proc. First IEEE International Conference on Cognitive Informatics, ICCI 2002*, August 19-20, 2002, Calgary, Alberta, Canada, pp. 304-312.

[8] V. Neagoe, ”A neural approach to compression of hyperspectral remote sensing imagery”, in: Reusch, B. (ed.): *Computational Intelligence. Theory and Applications* (*Proceedings of International Conference on Computational Intelligence “7th Fuzzy Days in Dortmund”,* October 1-3, 2001, Dortmund, Germany), Springer, Berlin (2001), pp. 436-449.