An Intelligent System
for Aerial Image Retrieval and Classification

Antonios Gasteratos[1], Panagiotis Zafeiridis[2], and Ioannis Andreadis[2]

[1] Laboratory of Robotics and Automation, Section of Production Systems,
Department of Production and Management Engineering, Democritus University of Thrace
Building of University's Library, Kimmeria, GR-671 00 Xanthi, Greece
`agaster@pme.duth.gr`
`http://utopia.duth.gr/~agaster`
[2] Laboratory of Electronics, Section of Electronics and Information Systems Technology,
Department of Electrical and Computer Engineering, Democritus University of Thrace
Vassilisis Sophias 12, GR-671 00 Xanthi, Greece
`{pzafirid,iandread}@ee.duth.gr`

Abstract. Content based image retrieval is an active research area of pattern recognition. A new method of extracting global texture energy descriptors is proposed and it is combined with features describing the color aspect of texture, suitable for image retrieval. The same features are also used for image classification, by its semantic content. An exemplar fuzzy system for aerial image retrieval and classification is proposed. The fuzzy system calculates the degree that a class, such as sea, clouds, desert, forests and plantations, participates in the input image. Target applications include remote sensing, computer vision, forestry, fishery, agricultures, oceanography and weather forecasting.

Keywords: CBIR, Machine intelligence, Fuzzy systems, Data fusion

1 Introduction

The recent improvements in network technologies lead to higher data transmission rates. Consequently this leads to faster internet connections around the globe. On the other hand one might say that the vast number of internet users necessitated the high speed internet connections and pushed the research to faster networks. No matter which comes first, the fast internet connections along with today's powerful computers and the proliferation of the imaging devices (scanners, digital cameras etc) moved forward a relatively new branch of *pattern recognition*; the so-called *content-based image retrieval* (CBIR). This is the retrieval of images on the basis of features automatically derived from the images themselves. The features most widely used are texture [1-3], color [4-6] and shape [7-9]. A plethora of texture features extraction algorithms exists, such as wavelets [10-12], mathematical morphology [13] and stochastic models [14], to mention few. A simple but efficient method to represent textures is using signatures based on texture energy [15, 16]. Energy images result from the convolution of the original image with special kernels representing specific texture properties. An attempt to describe the texture by means of color information was carried out in [17]. This method allows an effective evaluation of texture similarity in terms of color aspect and, therefore, to attribute textures to classes based on their color composition.

A review of the existing image retrieval techniques is presented in [18]. These are categorized into three groups: automatic scene analysis, model-based and statistical approaches and adaptive learning from user feedback. Conclusively, it is said that the CBIR is in its infancy and that, in order to develop truly intelligent CBIR systems, combination of techniques from the image processing and artificial intelligence fields should be tried out. In the present paper such an algorithm is proposed. It combines texture and color features by means of a least mean square (LMS) technique. The texture features of the images are extracted using the Laws convolution method [15, 16]. However, instead of extracting a new image each of its pixels describing the local texture energy, a single descriptor is proposed for the whole image. Each class of scenes corresponds to a certain band in the descriptor space. The color similarity is examined by means of its characteristic colors [17]. The same feature set can be used also for image classification, by its semantic content. The classification is performed by a fuzzy system. The membership functions (mf) of the proposed method are con-structed by statistical analysis of the training features. As an example, a system that classifies aerial images is described. Experiments demonstrate the high efficiency of the proposed system. The use of these particular texture and color texture descriptors is attempted for the first time. The redundancy of texture information decreases the classification uncertainty of the system.

2   Algorithm Description

2.1   Texture Feature Extraction

The texture feature extraction of the proposed system relies on Laws texture measures [15], where the notion of "local texture energy" is introduced. The idea is to convolve the image with 5x5 kernels and then to apply a nonlinear windowing operation over the convolved image. In this way a new image results, each pixel of which represents the local texture energy of the corresponding pixel of the original image. Laws have proposed 25 individual zero-summing kernels, each describing a different aspect of the local texture energy. These kernels are generated by the one-dimensional kernels, shown in Figure 1. As an example of how the 2-dimensional kernels are generated, L5S5 results by multiplying the 1-dimensional kernel L5 with S5. Experiments with all the 25 kernels showed that, as far as our application is concerned, the most potent ones are R5R5, E5S5, L5S5 and E5L5. More specifically, applying each of these four masks to images of a certain class (sea, forest, etc.) the global texture descriptors were more concentrated than with the rest of the masks. These kernels were used to extract the four texture descriptors of the proposed system.

L5 = [	1	4	6	4	1	]
E5 = [	-1	-2	0	2	1	]
S5 = [	-1	0	2	0	-1	]
W5 = [	-1	2	0	-2	1	]
R5 = [	1	-4	6	-4	1	]

Fig. 1. 1-dimensional kernels; the mnemonics stand for Level, Edge, Spot, Wave and Ripple, respectively.

The first texture descriptor of the image is extracted by convolving it with the first kernel (R5R5). The descriptor is the absolute average of the convolved image pixels. Thus, instead of measuring local texture descriptors, by averaging over local windows (typically 15x15), as it is proposed in Laws original work, we keep one global texture descriptor by averaging over the whole image. This descriptor is normalized by the maximum one, found among a database of 150 training images. If, for a sought image, the absolute average of the convolved image is greater than maximum value, then the descriptor is 1:

$$d_1 = \begin{cases} \dfrac{\left(\dfrac{1}{m \times n} \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} \left| I(i,j) * R5R5 \right| \right)}{d_{1\max}} & \text{if } \left(\dfrac{1}{m \times n} \sum\limits_{i=1}^{m} \sum\limits_{j=1}^{n} \left| I(i,j) * R5R5 \right| \right) \le d_{1\max} \\ 1 & \text{otherwise} \end{cases} \tag{1}$$

The same procedure is followed to extract the other three texture descriptors d_2, d_3 and d_4, by replacing in eqn (1) kernel R5R5 with the kernels E5S5, L5S5 and E5L5, respectively.

2.2   Color Feature Extraction

According to [17], in order to extract the characteristic colors of an image the following steps are followed:

1. On each color appearing in the image, its frequency of appearance is assigned.
2. Colors are sorted in descending order according to their frequency of appearance.
3. Given a color and a certain radius, a spherical volume is constructed in the RGB color space. The first color in the descending order comprises the first characteristic color of the image. Starting with the second color it is examined whether it lies within the volume of any color above it. If so, then the examined color is merged with the color in the volume where it lies. Otherwise it comprises a new characteristic color of the image.

Considering the set of the characteristic colors as a vector, the color similarity of two images is computed by means of the angle between these two vectors. More specifically, the ratio of the inner product to the product of the measures of the two vectors corresponds to the cosine of the angle of these two vectors. The greater the value of the cosine, the smaller the angle and the more similar the two images (in terms of their color prospect). Therefore, the cosine could be used as the color descriptor of similarity. However, because of the fact that the angle is the absolute descriptor and the cosine is a nonlinear function, the descriptor used in the proposed system is:

$$d_5 = \frac{2}{\pi} \arccos \frac{\overline{C}_1 \cdot \overline{C}_2}{\left\| \overline{C}_1 \right\| \cdot \left\| \overline{C}_2 \right\|} \tag{2}$$

where: $\overline{C}_1$ and $\overline{C}_2$ the set of the characteristic colors of images 1 and 2, respectively.

2.3   Image Retrieval

After extracting the descriptors both for the input and the sought images, the retrieval is performed by minimizing the following distance:

$$m = \frac{1}{\sum\limits_{i=1}^{5} w_i} \sqrt{\sum_{i=1}^{4} w_i \left(din_i - ds_i \right)^2 + w_5 (d_5)^2} \tag{3}$$

where: din_i (i=1,…4) are the four texture descriptors of the input image, resulting according to eqn (1) and ds_i is the corresponding texture descriptor of the sought image; d_5 is the color descriptor according to eqn (2) and w_i is a weight tuning the retrieval process according to the importance of each descriptor. By comparing eqns (1) and (2) it can be observed that though d_5 is a differential descriptor, i.e. it presents the difference of two images by means of their color aspect, $d_1,…d_4$ are absolute ones. This is the reason why the difference of the last ones appears in eqn (3).

2.4   Image Classification

The same feature set described above and used for CBIM may be used to classify images according to their texture and color properties. In this section a fuzzy system for the data fusion of the different descriptors is proposed. The system is tailored to meet the needs of the target application, i.e. the categorization of aerial images into five different classes. However, with slight variations it might be applied to other applications of image classification as well. The inputs of the fuzzy system are the five descriptors presented in the previous paragraphs. In order to construct the mfs for the inputs a statistical analysis was carried out. More specifically, there were used five different classes of photographs named: sea, clouds, desert, forests and plantations. As training data, 100 images of each class were used. For each image the four texture descriptors were extracted. In Figure 2 the histograms of the distribution of the four descriptors for the class of the sea, are presented. As it can be seen, the distribution can be approximated by a trapezoidal or even a triangular mf. However, a Gaussian function is also a good approximation, far better than the two latter. The reason is that its curve is not as steep as these of a triangular or a trapezoidal one and, therefore, it includes also the sided values. Experiments with several mfs proved this intuition. For each descriptor and for each image class the mean value and the standard deviation were calculated. The mf were computed as the normal distribution, for the previous values (see Figure 2). In Figure 3 the membership functions for the descriptor d_1 are depicted, as an example of the four first inputs of the fuzzy system.

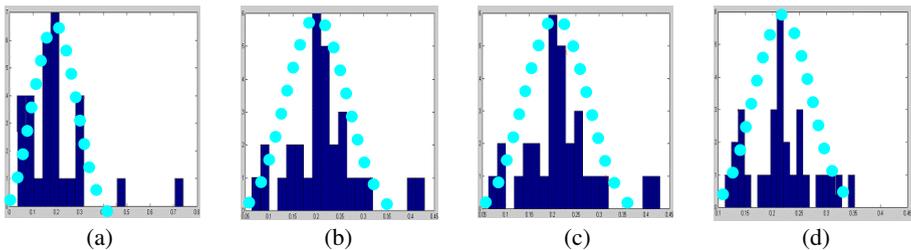| (a) | (b) | (c) | (d) |

Fig. 2. The histogram of the global texture energy distribution for the training images belonging to the class of the sea. Graphs (a), (b), (c) and (d) show the histograms of descriptor d_1,… d_4, respectively.
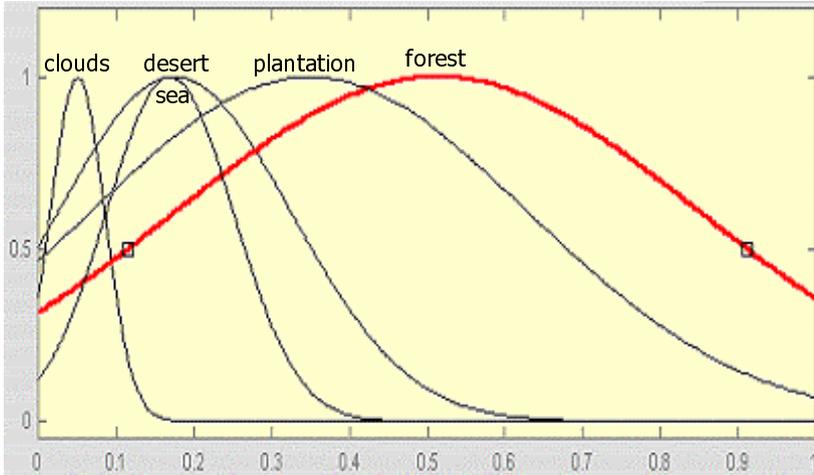
Fig. 3. The first input of the fuzzy system is the descriptor d_1. The membership functions from left to right are: Clouds, desert, sea, plantations and forests.

For the color descriptor five different inputs were used. The characteristic colors of the 100 training images of each class were merged in the same way as described in section 2.2 for a single image. The result is a color codebook [17] containing the characteristic colors of the whole image class. Eqn (2) is used to compute the similarity between the characteristic colors of the input image and the codebook of each of the classes. The result of each of the color similarity values is used as an input to the fuzzy system (inputs from five to ten). Similarly, five sigmoid mfs outputs, one for each class, were used. Having defined the inputs and the output of the system, the following set of if-then rules was used:

1. **If** (IN1 is clouds) **and** (IN2 is clouds) **and** (IN3 is clouds) **and** (IN4 is clouds) **and** (IN5 is clouds) **then** (OUT1 is clouds)

2. **If** (IN1 is plantation) **and** (IN2 is plantation) **and** (IN3 is plantation) **and** (IN4 is plantation) **and** (IN6 is plantation) **then** (OUT2 is plantation)

3. **If** (IN1 is desert) **and** (IN2 is desert) **and** (IN3 is desert) **and** (IN4 is desert) **and** (IN7 is desert) **then** (OUT3 is desert)

4. **If** (IN1 is sea) **and** (IN2 is sea) **and** (IN3 is sea) **and** (IN4 is sea) **and** (IN8 is sea) **then** (OUT4 is sea)

5. **If** (IN1 is forest) **and** (IN2 is forest) **and** (IN3 is forest) **and** (IN4 is forest) **and** (IN9 is forest) **then** (OUT5 is forest)

As far as the other implementation parameters, experiments showed better results when the "*and*" method is the algebraic product and the "*or*" is the minimum. For the implication the minimum is also used and for the aggregation the maximum. Finally, for the defuzzification the *som* (smallest of maximum) method was used. An example of how the system operates is shown in Figure 4. The input image in Figure 4a does not clearly belong to a certain class. The system perceives this fact and gives output for almost each of the classes.

3   Experiments

In order to evaluate the performance of both the retrieval and the classification systems, several experiments were carried out:

3.1   Image Retrieval

The first experiments were carried out, in order to assign the weights of eqn (3), that gives optimum results. Each time the six more relevant images were asked to be retrieved. The precision, i.e. the ratio of the correctly retrieved images over the total retrieved images, was used to measure the efficiency of the retrieval system. It has been observed that the best results occurred when $w_1=w_2=w_3=w_4=1$ and $w_5=11$. In particular, the retrieval precision was measured in the range of 35% to 100%, whilst no other combination of weights had ever resulted to precision 100%. This is to say that the color information plays dominant role in the image retrieval process, as the ratio of color to texture coefficients in the minimization of eqn (3) is 11/4.
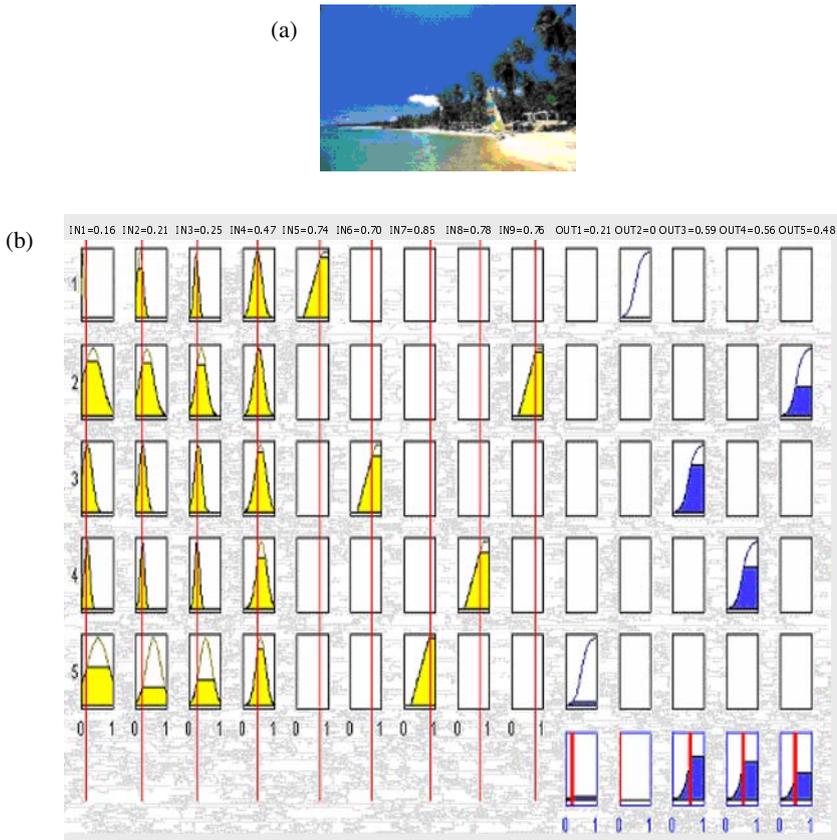
(a)

(b)

Fig. 4. An example of system's operation: (a) An input image and (b) system's outputs, corresponding to forest, clouds, desert, sea and plantation, respectively.

An example of how weights affect the retrieval is shown in Figure 5. In Figure 5a the input image is show. The six best matches of the retrieval procedure with $w_1=w_2=w_3=w_4=w_5=1$ are presented in Figure 5b. Figure 5c illustrates the six best matches, after having altered w_5 to 11. Comparing Figure 5b to Figure 5c, one can see that the retrieved images belonging to a different class than plantations where now reduced to one. Therefore, the improvement in performance is obvious when the weight corresponding to color descriptor is high.

(a)
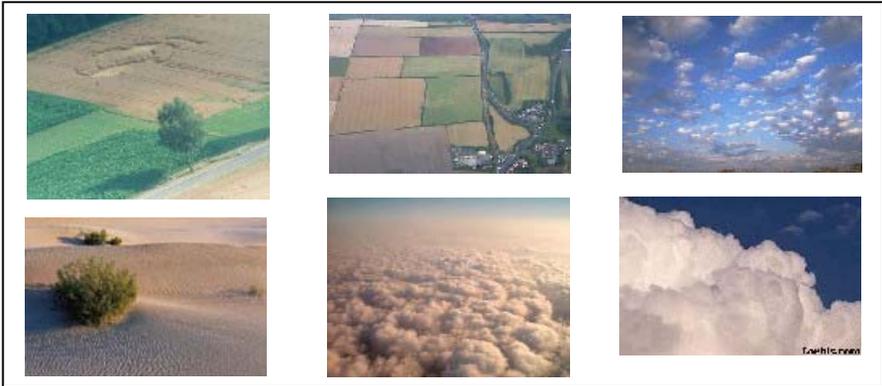
(b)

(c)

Fig. 5. (a) Input image to the image retrieval system; (b) the six best matches when $w_1=w_2=w_3=w_4=w_5=1$ and (c) the six best matches when $w_1=w_2=w_3=w_4=1$ and $w_5=11$.

3.2   Image Classification

As independent test data, a set of 500 landscapes were used. In many of them, that exhibit a mixture of the 5 different classes, the corresponding outputs were activated. We considered a classification to be correct if the output corresponding to the dominant class has the largest value. In this way, a winner-takes-it-all procedure can result to a single output. As an example for the image of dessert in Figure 6 the outputs are: forest: 0.07, clouds: 0, desert: 0.37, sea: 0.16, and plantation: 0.34, which is correct, as the desert output results the largest value. The high values in other outputs are due to the existence of a range of bushes in the picture. The bushes have the same characteristic colors with the class of forests and plantations. The quite high value of the output corresponding to the sea class is justified as the desert and the sea comprise the same texture characteristics, but their characteristic colors are extremely dissimilar. Therefore, to conclude neither the texture itself nor the color itself is adequate for reliable classification, but the combination of these two could lead to efficient and reliable classification. The Empirical Error Rate (EER) of the classification system, which is the number of errors made on independent test data over the number of classifications attempted, was measured over 500 independent test data and it was found 18.02%.

Fig. 6. The outputs of the fuzzy classification system for this dessert image are: forest: 0.07, clouds: 0, desert: 0.37, sea: 0.16, and plantation: 0.34.

4   Conclusions

The extraction of global texture and color descriptors was presented in this paper. Laws' local texture energy method was modified to extract four different global texture descriptors. The color descriptor extraction was based on a method for describing the color aspect of texture images. It has been shown that the same descriptors can be combined for the construction of an image-retrieval and an image classification system, providing good results. Target applications of such a system include remote sensing, computer vision, forestry, fishery, agricultures, oceanography and weather forecasting.

References

1. Manjunath, B.S. and Ma, W.Y.: Texture Features for Browsing and Retrieval of Large Image Data. *IEEE Trans. Pattern Anal. Mach. Intell.*, **18**, (1996) , 837-842.
2. Gimelfarb, G.L. and Jain, A.K.: On Retrieving Textured Images From an Image Database. *Pattern Recognition*, **29**, (1996), 1461-1483.
3. Carkacloglu, A and Yarman-Vural, F.: SASI: a generic texture descriptor for image retrieval. *Pattern Recognition*, **36**, (2003), 2615-2633.
4. Gevers, T. and Smeulders, A.W.M.: Color-based Object Recognition. *Pattern Recognition*, **32**, (1999), 453-464.
5. Konstandinidis, K. and Andreadis, I.: On the Use of Color Histograms for Content Based Image Retrieval in Various Color Spaces. *Int. Conf. of Computational Methods in Sciences and Engineering*, Kastoria, Greece, (2003).
6. Eftekhari-Moghadam, A.M., Shanbehzadeh, J., Mahmoudi, F. and Soltanian-Zadeh, H.: Image Retrieval Based on Index Compressed Vector Quantization. *Pattern Recognition*, **36**, (2003), 2635-2647.
7. Mechrotra, R. and Gary, J.E.: Similar Shape Retrieval in Shape Data Management. *IEEE Computer*, **28**, (1995) , 57-62.
8. Del Bimbo, A., Pala, P. and Santini, S.: Image Retrieval by Elastic Matching of Shapes and Image Patterns. *Proc. IEEE Int. Conf. Multimedia Systems and Computing*, Hiroshima, Japan, (1996), 215-218.
9. Oonincx, P.J. and de Zeeuw, P. M.: Adaptive Lifting for Shape-based Image Retrieval. *Pattern Recognition*, **36**, (2003), 2663-2672.
10. Zhang, D. S. and Lu, G.: Content-based Image Retrieval Using Gabor Texture Features. *Proc. First IEEE Pacific-Rim Conference on Multimedia*, Sydney, Australia, (2000), 392-395.
11. Kam, A.H., Ng, T.T., Kingsbury, N.G. and Fitzgerald, W.J.: Content Based Image Retrieval Through Object Extraction and Querying. *Proc. IEEE Workshop on Content-based Access of Image and Video Libraries*, Hilton Head Island, S. Carolina, (2000), 91-95.
12. Wang, J.Z. and Li, J. and Wiederhold, G.: SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on PAMI,* **23**, (2001), 947-963.
13. Soille, P.: Morphological Texture Analysis: a survey. *Workshop on Texture Analysis 1998*, Albert-Ludwigs-Universitat Freiburg, Germany, (1998), 193-207.
14. Cross, G.R. and Jain, A.K.: Markov Random Field Texture Models. *IEEE Trans. Pattern Anal. Mach. Intell.*, **18**, (1983) , 25-39.
15. Laws, K.: *Textured Image Segmentation*, Ph.D. Dissertation, University of South California, (1980).
16. Laws, K.: Rapid Texture Identification. *SPIE Vol 238, Image Processing for Missile Guidance*, (1980), 376-380.
17. Scharcanski, J., Hovis, J.K. and Shen, H.C.: Representing the Color Aspect of Texture Images. *Pattern Recognition Letters*, **15**, (1994) , 191-197.
18. Eakins, J. P.: Towards Intelligent Image Retrieval. *Pattern Recognition*, **35**, (2002) , 3-14.