International Journal of Computational Intelligence and Applications
Vol. 2, No. 2 (2002) 173–184

AN INTELLIGENT HYBRID APPROACH FOR CONTENT-BASED IMAGE RETRIEVAL

S. KULKARNI
Department of Mathematics and Computer Science
Nipissing University, North Bay, ON P1B 8L7, Canada
siddhik@nipissingu.ca

B. VERMA
School of Information Technology
Griffith University, Gold Coast Campus, Qld 9726, Australia
B.Verma@gu.edu.au
http://intsun.int.gu.edu.au

Received 5 November 2001
Revised 17 May 2002

The paper presents an intelligent hybrid approach for content-based image retrieval based on texture feature. The proposed approach employs an Auto–Associative Neural Network (AANN) for feature extraction and a Multi–Layer Perceptron (MLP) with a single hidden layer for the classification. Two intelligent approaches such as AANN–MLP and statistical–MLP were investigated. The performance of the proposed approaches was evaluated on a large benchmark database of texture patterns. The results are very promising compared to other existing traditional and intelligent techniques. Some of the experimental results conducted during the investigation, comparative analysis of the results and suggestions to select the appropriate techniques for texture feature extraction and classification are presented in this paper.

Keywords: Auto-associator; texture feature extraction; multi-layer perceptron; content-based image retrieval.

1. Introduction

In recent years, multimedia technology has grown rapidly which allows fast access to large image databases on the world wide web. The technology is being applied to access digital libraries in every field of human life. An efficient content-based image retrieval technique is very useful and necessary to access such large databases. Color, texture and shape are the most important features to retrieve digital images efficiently. Texture is an important characteristic for the analysis of many types of the images including natural scenes, remotely sensed data and biomedical modalities. Texture refers to the visual patterns that have properties of homogeneity that do not result from the presence of only single color intensity. The perception of texture is believed to play an important role in the visual system for recognition and interpretation. In order to retrieve an image, it is necessary to determine the texture features, which are useful in the classification process.

Texture has been studied for more than twenty years, various techniques have been developed for texture segmentation, texture classification, texture synthesis and texture feature extraction.[1] In recent years, many efforts have been made to improve the performance of image retrieval systems by using texture features. Researchers have explored many different directions, trying to use the existing techniques with promising results achieved in other areas.[2]

There are three principal approaches to describe the texture of an image region: they are statistical, structural and spectral. Statistical techniques attempt to describe texels by attributing them with the statistical distribution of pixels, while structural techniques use the arrangement of image primitives and shapes to classify and recognize texels. Spectral techniques, based on the Fourier transform, are used mainly to detect global periodicities. The spectral methods are generally regarded as inferior to the statistical ones in terms of classification.[3] Since the natural textures are not very regular, the structural techniques are not very popular now. The often-used techniques are statistical and spectral approaches. An important aspect of texture is scale. Psychovisual studies indicate that the human visual system processes images in multiscale texture analysis methods. Wavelet theory is a mathematical framework, which provides a unified approach to multiresolution representations.[4] Another texture analysis method is Markov Random Field (MRF) modeling. In this method, a texture is looked at as a realization of a Markov random field. To model a texture is to specify the corresponding conditional probabilities or Gibbs clique potential parameters. In recent years, several authors proposed methods to integrate the filtering theory and stochastic models to represent texture effectively. One instance of MRF model is the simultaneous autoregressive (SAR) model, which has been used in various texture applications. With SAR models, there are major difficulties in selecting the size of the dependent pixel neighborhood and the appropriate window size in which the texture is regarded as homogenous. To overcome these difficulties and to expand the SAR model into the multiresolution image representation paradigm, the multiresolution simultaneous autoregressive (MRSAR) model was presented by Mao and Jain.[5] But MRSAR is computationally a very expensive set of features. Liu and Picard[6] try to use 2-D Wold decomposition to represent textures with "periodicity," "directionality" and "randomness" by decomposing a texture into three mutual orthogonal components in the Fourier transform domain. The wold decomposition model avoids the actual decomposition of images and tolerates a variety of in-homogeneities in natural data, making it suitable for use in large collections of natural patterns.[7] The statistical properties such as mean and variance are extracted from the wavelet subbands as texture representations. To explore the middle band characteristics, the tree structured wavelet transform is used to improve the classification accuracy.[8] Haralick *et al.* proposed the co-occurrence matrix representation of texture feature.[9] This approach explored the gray level spatial dependence of texture. It first constructed a co-occurrence matrix based on the orientation and distance between image pixels and then extracted meaningful statistics from the matrix as the texture representations. In early 90's after the Wavelet transform was introduced and its theoretical framework established, many researchers began to study the use of the wavelet transform in texture representations.[10–12] Smith and Chang[13] used the statistics (mean and variance) extracted from the Wavelet subbands as the texture representations. Wavelets provide a convenient way to obtain a multiresolution representation from which texture features are extracted. These energy signatures have proven to be very powerful in texture analysis.[14,15] Pichler *et al.*[16] compare wavelet transforms with adaptive Gabor filtering feature extraction and report superior results using the Gabor technique. In Ref. 17, Ohanian and Dubes compared and evaluated four types of texture representations, namely: Markov Random Field representation,[18] multichannel filtering representation, fractal based representation,[19] and co-occurrence representation. They found out that the co-occurrence matrix representation performed best in their test sets.

Motivated by psychological studies in human visual perception of texture, Tamura *et al.* explored the texture representation from a different angle.[20] They developed computational approximations to the visual texture properties found to be important in psychological studies. The six visual texture properties were coarseness, contrast, directionality, linelikeness, regularity and roughness. One major distinction between the Tamura texture representation and the co-occurrence matrix representation is that all texture properties in Tamura representation are visually meaningful whereas some of the texture properties used in the co-occurrence matrix may not (for example entropy). This characteristics makes the tamura texture representation very attractive in image retrieval, as it can provide a more friendly user interface. The QBIC system[21] and MARS system[22,23] further improved the texture representation. The Gabor transform provides an attractive approach, which is well suited to texture classification and database retrieval. These are far superior compared to co-occurrence features and less sensitive to noise.[24] But there are possibilities for either mistreatment or adaptation to suit specific data. This technique is not generally applicable for segmentation or image analysis.

There are also a number of papers published describing intelligent techniques such as neural based texture feature extraction and classification. However, most research using intelligent techniques was mainly based on Kohonen neural network.[25] The research proposed in this paper focuses on intelligent hybrid techniques to extract and classify texture features. The hybrid technique is based on multilayer perceptron (MLP) and autoassociator neural networks in conjunction with statistical feature extraction technique. Research in other areas such as neural networks and pattern recognition, showed[26] that the neural networks could be good feature extractors and classifiers. In this research we would like to get answers to the following research questions: (1) How good is the Auto–Associator Neural Network (AANN) for texture feature extraction? (2) Which neural network (MLP type, Kohonen type) is good for texture feature extraction and classification? (3) Which hybrid technique (AANN–MLP, statistical–MLP) is most suitable for texture feature extraction and classification? (4) Which technique (intelligent based on neural networks, traditional) is good for texture feature extraction and classification?

The rest of the paper is organised as follows: Section 2 describes in detail the proposed feature extraction process using an AANN and classification using MLP. The statistical techniques, the preparation of training and testing data sets from Brodatz texture patterns and texture classes in detail are presented in Sec. 3. The experimental results are presented in Sec. 4. The analysis and comparisons are discussed in Sec. 5. The conclusion is presented in Sec. 6.

2. Proposed Hybrid Approach

This section describes in detail the proposed approach for feature extraction and classification. The overall block diagram of the proposed approach is presented in Fig. 1 as follows.

The proposed approach is divided into two stages. Stage 1 deals with feature extraction from texture sub-images. An auto-associator was designed to extract features. Stage 2, deals with classification of features into texture classes. A Multi–Layer Perceptron (MLP) was designed to classify texture classes. The auto-associator feature extractor and MLP texture feature classifier are described below in detail.

2.1. *Auto-associator feature extractor*

The main idea of an auto-associator feature extractor is based on input:hidden:output mapping, where input and outputs are the same patterns. AAFE learns the same patterns and provides a characteristic through its hidden layer as a feature vector. As shown in Fig. 2, we designed an auto-associator feature extractor using a single hidden layer feed-forward neural network. It has $\boldsymbol{n}$ inputs,

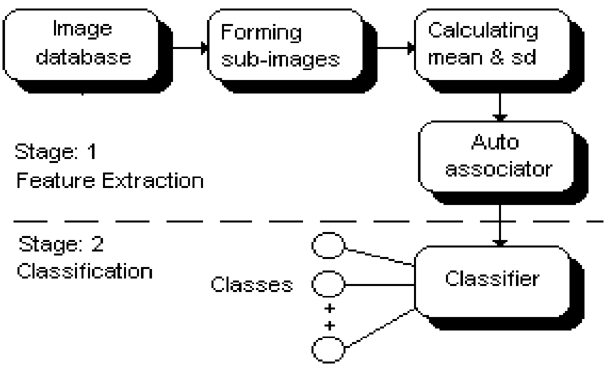


Fig. 1. Block diagram of the proposed approach.

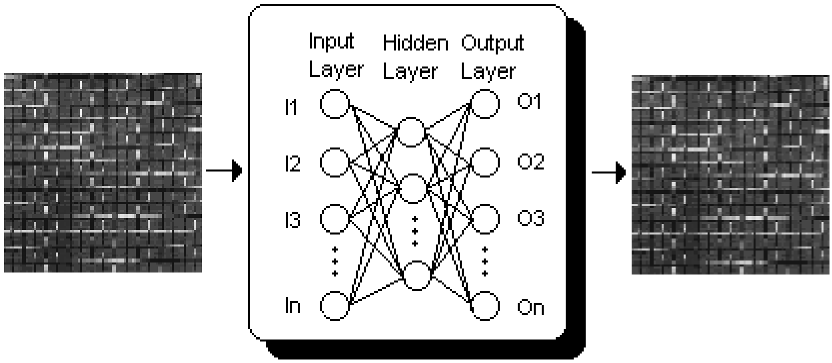


Fig. 2. Auto-associator as a feature extractor.

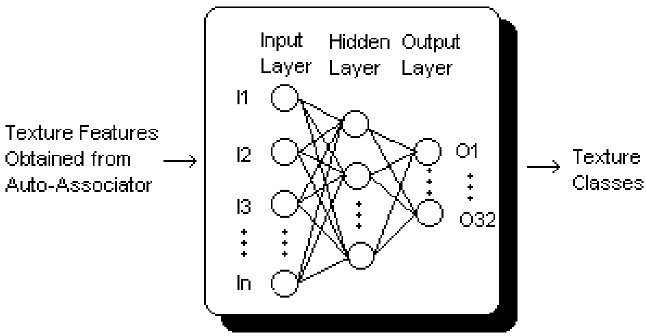


Fig. 3. MLP texture feature classifier.

$\boldsymbol{n}$ outputs and $\boldsymbol{p}$ hidden units. The input and output of the AAFE are the same texture patterns and the network is trained using supervised learning. After training is finished, the values of the hidden units are extracted and taken as a feature vector. The feature vector is fed to the MLP feature classifier which is described in the next section.

2.2. *MLP texture feature classifier*

An MLP texture feature classifier is shown in Fig. 3. It has $\boldsymbol{n}$ inputs which is the same as the number of hidden units in the auto-associator feature extractor. The output of the hidden layer that was obtained from the auto–associator was used as input to the classifier. There were 32 texture classes, so the number of outputs was 32.

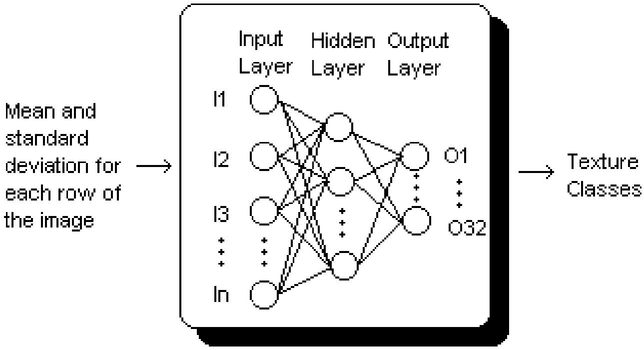


Fig. 4. Statistical-MLP texture feature classifier.

2.3. *Statistical-MLP texture feature classifier*

As shown in Fig. 4, statistical techniques were used to extract the texture features from the images. The mean and standard deviation, which form the properties of texture, were calculated for each row of the image. These were treated as the features of the texture patterns and applied to the texture feature classifier, as input.

3. Extraction of Training and Testing Sets From Brodatz Texture Database

The Brodatz texture database[27] was used to evaluate the performance of the proposed techniques detailed in the previous section for texture features extraction and classification. The collection of Brodatz textures consists of textures of both statistical and structural natures. Structural textures are considered to be consists of texture primitives which are repeated systematically within the texture. In statistical textures usually no repetitive texture can be identified. The database contains 96, 512×512 texture images. In order to create a number of small images which belong to the same class, we partition each of the 512×512 images into 128×128 sub-images, thus forming 16 sub-images from each image. To reduce the size of input vector to the neural network, the mean and standard deviation was calculated for each row (128 pixels).

First 12 sub-images were used for the training of the auto-associator and the last 4 sub-images were used as the testing data set. These images were normalized in the range of 0 and 1.

There were a total of 96 texture patterns in the database which were grouped into 32 similar clusters, each of them containing 1–6 texture classes, these are gray scale texture images that contain the texture of brick wall, wood grain, woolen cloth, beach sand, lizard skin etc. All the texture sub-images belonging to the same similarity cluster are visually similar. This classification was done manually[25] and

Table 1. Texture clusters used for classifier.

Cluster	Texture Class	Cluster	Texture Class	Cluster	Texture Class
1	D1, D6, D14, D20, D49	12	D62, D88, D89	23	D19, D82, D83, D85
2	D8, D56, D64, D65	13	D24, D80, D81	24	D66, D67, D74, D75
3	D34, D52	14	D50, D51, D68, D70, D76	25	D2
4	D18, D46, D47	15	D25, D26, D96	26	D86
5	D11, D16, D17	16	D94, D95	27	D37, D38
6	D21, D55, D84	17	D69, D71, D72, D93	28	D9
7	D53, D77, D78, D79	18	D4, D29, D57, D92	29	D12, D13
8	D5, D33	19	D39, D40, D41, D42	30	D15
9	D23, D27, D28, D30, D54	20	D3, D10, D22, D35, D36, D87	31	D31
10	D7, D58, D60	21	D48, D90, D91	32	D32
11	D59, D61, D63	22	D43, D44, D45		

Table 1 shows these various similarity clusters and the corresponding textures. Cluster 1 contained a total of five images. Cluster 20 contained the maximum number of images which was six. Some of the clusters have only one image in their cluster.

4. Experimental Results

The proposed techniques were implemented in C on the UNIX platform in conjunction with MATLAB's neural network package. All experiments included in this section were carried out on SP2 supercomputer, the SP2 was selected mainly because of its speed, larger disk storage and virtual memory space. The experiments were conducted separately for the auto-associator-classifier technique and statistical-neural technique. The following Secs. 4.1 and 4.2 explain the results obtained using the proposed techniques on the Brodatz texture patterns.[27]

4.1. *Auto-associator and classifier*

The experiments were conducted in two stages, firstly the training of the auto-associator and then the training of the classifier. The auto-associator was trained for the different number of hidden units and iterations to improve the feature extraction. Table 2 shows the RMS error for some of the experiments during the training of the auto-associator. The number of inputs and outputs were 256. The values of momentum and learning rate used for the experiments were 0.7 and 0.8 respectively. The number of pairs for training and testing were 640 and 160 respectively. The experiments were conducted by varying the number of hidden units and iterations to minimize RMS error and achieve better features.

The classifier was trained after obtaining the output from the hidden layer of the auto-associator. The experiments for the classifier were conducted for various numbers of hidden units and iterations and the results are shown in Table 3.

Table 2. Auto-associator feature extractor.

Hidden Units	Iterations	RMS Error
10	1000	0.005698
10	10000	0.003166
15	10000	0.002160
22	10000	0.001471
32	10000	0.000956

Table 3. MLP texture feature classifier.

Hidden Units	Iterations	RMS Error	Classification Rate [%]
16	10000	0.0053	94.37
16	5000	0.0061	90.62
16	2000	0.0075	88.12
32	50000	0.0088	86.19

Table 4. Statistical-MLP texture feature classifier.

Hidden Units	Iterations	RMS Error	Classification Rate [%]
18	100000	0.00517	96.06
18	50000	0.00520	95.57
18	40000	0.00532	95.31
18	30000	0.00556	94.79
18	20000	0.00592	92.44
18	10000	0.00693	88.02
15	10000	0.00810	85.15

The following graphs (Fig. 5) show the difference in the amplitude of the features extracted from different classes.

4.2. *Statistical–MLP texture feature classifier*

The 16 non-overlapping sub-images were formed from the image of size 512×512. The size of the sub-images was 128×128. The mean and standard deviation were calculated for each row of the image, i.e. for 128 pixels, thus forming 256 as the number of inputs. The number of classes was 32, which is same as it was used in auto-associator and classifier experiments. The results obtained from the statistical-texture feature classifier are presented in Table 4. The values of momentum and learning rate used for the experiments were 0.7 and 0.8 respectively. The number of training pairs was 1152 for training and 384 for testing. The experiments were conducted for different numbers of hidden units. As the number of hidden units was increased, the error was decreased and the classification rate was improved.

5. Analysis and Comparison of the Results

The classification results obtained by the proposed techniques were compared with other existing techniques. In Ref. 25, Ma and Majunath compared a few statistical techniques such as Gabor, MR–SAR, PWT, and intelligent neural based technique such as Kohonen-LVQ. They listed very low classification rates for Gabor (74%), MR–SAR (73%) and PWT (68.7%). The classification rate for Kohonen–LVQ is not very clear, we are trying to guess the classification rate for this technique which could be around 93%. In Ref. 28 Jones and Jackway used the granold technique for texture classification. A granold is a texture representation technique that uses two parameterized monotonic mappings to transform an input image into a two and a half-dimensional surface. The granold spectrum for each image was then calculated and the gray level and size marginals formed. The confusion matrix was used to calculate the classification rate, which was 76.9% on the Brodatz texture database. In Ref. 29, Wang and Liu compared the Nearest Linear Combination (NLC) with Nearest Neighbor (NN) Classification. The testing set was formed from selecting random images from the training set with various sizes (100×100 to 200×200) to verify the classification rate. The highest classification rate obtained

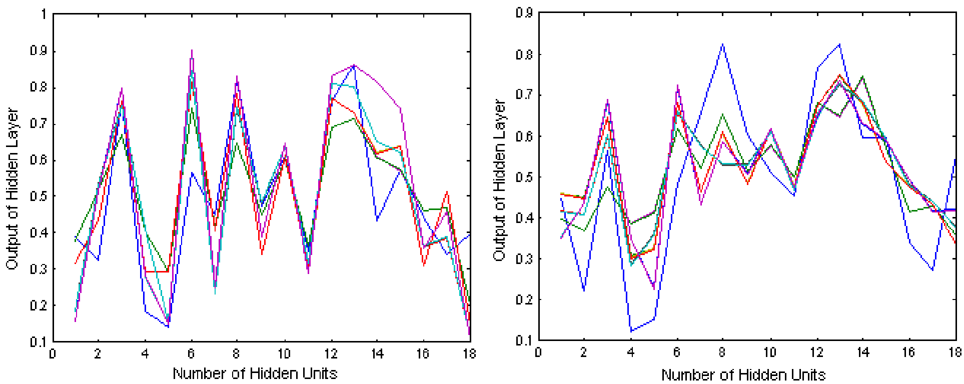
Table 5. Comparison with other techniques.

Gabor[25]	Granold[28]	NLC[29]	NN[29]	Kohonen–LVQ[25]	AANN–MLP (Proposed)	Statistical–MLP (Proposed)
74.00%	76.9 %	95.48%	93.47%	∼ 93.00%	94.37%	96.06%

was 95.48% with NLC and 93.47% with NN. As shown in Tables 3 and 4, the proposed techniques achieved 94.37% and 96.06% classification rates, which are higher than other existing techniques. Table 5 summarizes the results obtained by a number of existing and proposed techniques. The results in the above table clearly show that the proposed statistical-neural technique outperformed other techniques.

6. Conclusions

In this paper, we have proposed and investigated an auto-associator texture feature extractor and two hybrid intelligent techniques such as an auto–associator–MLP, and statistical–MLP for texture feature extraction and classification. The auto-associator seems to be a promising feature extractor, a number of feature graphs have been presented in Fig. 5, which show that the auto-associator is capable of separating texture classes very well and without any feedback from the user. The feature extraction and classification techniques were tested on a large database of texture patterns namely the Brodatz texture database. The results obtained by our techniques were analysed and compared with other intelligent and conventional techniques. The highest classification rate was obtained using our statistical-MLP technique. The proposed hybrid neural technique outperformed other statistical techniques such as MRF, NN, etc. and neural network based techniques such as Kohonen neural network and learning vector quantization. We can conclude by



(a) (b)

Fig. 5. Extracted features for (a) cluster 2 and (b) cluster 32.

addressing our research questions mentioned in the introductionary section as follows: (1) AANN is a good texture feature extractor, however it did not perform as well as statistical–MLP. (2) The MLP type neural network is better than the Kohonen type neural network for texture feature extraction and classification. (3) The statistical–MLP is the best technique for texture feature extraction and classification. As shown in Table 5, it achieved the highest classification rate. (4) Intelligent techniques based on neural networks are better than the traditional techniques. Based on our research, we conclude that a hybrid intelligent approach such as statistical–MLP is best suited for texture feature extraction and classification. In the future, we would like to conduct more experiments on a different benchmark database and different input matrix to our auto-associator feature extractor.

References

1. B. S. Manjunath and W. Y. Ma, Texture Features for Browsing and Retrieval of Image Data, *IEEE Trans. Pattern Analysis and Machine Intelligence* **8**, 18 (1996) 837–842.
2. W. Niblack, The QBIC Project: Querying Images by Content using Color, Texture and Shape, *SPIE Proc. Storage and Retrieval for Color and Image Video Databases* (1993) 173–187.
3. A. K. Jain and F. Farrokhnia, Unsupervised Texture Segmentation Using Gabor Filters, *J. Pattern Recognition* **24**, 12 (1991) 1167–1186.
4. Y. Rubner and C. Tomasi, Texture Matrices, *Proc. IEEE Int. Conf. Systems, Man and Cybernetics* (San-Diego, USA, 1998) 4601–4607.
5. J. Mao and A. Jain, Texture Classification and Segmentation using Multiresolution Simultaneous Autoregressive Models, *J. Pattern Recognition* **25**, 2 (1992) 173–188.
6. F. Lui and R. Picard, Periodicity, Directionality and Randomness: Wold Features for Image Modelling and Retrieval, *IEEE Trans. Pattern Analysis and Machine Intelligence* **18**, 7 (1996) 722–733.
7. A. Rao and G. Lohse, Towards a Texture Naming System: Identifying Relevant Dimensions of Texture, *Proc. IEEE Conf. Visualization* (San Jose, USA, 1993) 220–227.
8. I. Daubechies, The Wavelet Transform, Time-Frequency Localisation and Signal Analysis, *IEEE Trans. Information Theory* **9**, 36 (1990) 961–1005.
9. R. Haralick and K. Shanmugam, Texture Features for Image Classifications, *IEEE Trans. System, Man and Cybernetics* **8**, 6 (1978) 460–473.
10. T. Chang and C. Kuo, Texture Analysis and Classification with Tree Structured Wavelet Transform, *IEEE Trans. Image Processing* **2**, 4 (1993) 429–441.
11. A. Laine and J. Fan, Texture Classification by Wavelet Packet Signatures, *IEEE Transaction Pattern Recognition and Machine Intelligence* **15**, 11 (1993) 1186–1191.
12. M. Gross, R. Koch, L. Lippert and A. Dreger, Multiscale Image Texture Analysis in Wavelet Spaces, *Proc. IEEE Int. Conf. Image Proc.* **3** (1994) 412–416.
13. J. Smith and S. Chang, Transform Features for Texture Classification and Discrimination in Large Image Databases, *Proc. IEEE Int. Conf. Image Proc.* (1994) 407–411.
14. M. Unser, Texture Classification and Segmentation using Wavelet Frames, *IEEE Trans. Image Proc.* **4**, 11 (1995) 1549–1560.
15. A. Laine and J. Fan, Texture Classification by Wavelet Packet Signatures, *IEEE Trans. Pattern Analysis and Machine Intelligence* **15**, 11 (1993) 1186–1190.

16. O. Pichler, A. Teuner and B. Hosticka, A Comparison of Texture Feature Extraction using Adaptive Gabor Filter, Pyramidal and Tree Structured Wavelet Transforms, *J. Pattern Recognition* **29**, 5 (1996) 733–742.
17. P. Ohanian and R. Dubes, Performance Evaluation for Four Classes of Texture Features, *J. Pattern Recognition* **25**, 8 (1992) 819–833.
18. G. Cross and A. Jain, Markov Random Field Texture Models, *IEEE Trans. Pattern Recognition and Machine Intelligence* **5** (1983) 25–39.
19. A. Pentland, Fractal-based Decomposition of Natural Scenes, *IEEE Trans. Pattern Recognition and Machine Intelligence* **6**, 6 (1984) 661–674.
20. H. Tamura, S. Mori and T. Yamawaki, Texture Features Corresponding to Visual Perception, *IEEE Trans. Systems, Man and Cybernetics* **8**, 6 (1978) 460–473.
21. W. Equitz and W. Niblack, Retrieving Images from a Database using Texture-Algorithms from the QBIC System, Technical Report RJ 9805, Computer Science, IBM Research Report, 1994.
22. T. Huang, S. Mehrotra and K. Ramachandran, Multimedia Analysis and Retrieval System (MARS) Project, Proc. 33rd Annual Clinic on Library Application of Data Processing-Digital Image Access and Retrieval, 1996.
23. M. Ortega, Y. Rui, K. Chakrabarti, S. Mehrotra and T. Huang, Supporting Similarity Queries in MARS, *Proc. ACM Conf. Multimedia* (1997) 38–42.
24. J. Smith and S. Chang, Automated Binary Texture Feature Sets for Image Retrieval, *Proc. IEEE Int. Conf. Accost, Speech, and Signal Processing* **4** (May 1996) 2239–2242.
25. W. Y. Ma and B. S. Manjunath, Texture Features and Learning Similarity, *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition* (San Francisco, USA, 1996).
26. B. Lerner, H. Guterman, M. Aladjem and H. Dinstein, A Comparative Study of Neural Network based Feature Extraction Paradigms, *J. Pattern Recognition Letters* **20** (1999) 7–14.
27. P. Brodatz, *Textures: A Photographic Album for Artists and Designers* (Dover Publications, New York, 1996).
28. D. Jones and P. Jackway, Using Granold for Texture Classification, *Fifth International Conference on Digital Image Computing, Techniques and Applications*, DICT 99, Perth, Australia (1999) 270–274.
29. L. Wang and J. Liu, Texture Classification using Multiresolution Markov Random Field Models, *J. Pattern Recognition Letters* **20** (1999) 171–182.