# Content-based Image Retrieval Incorporating Models Of Human Perception

M. Emre Celebi
Dept. of Computer Science and Engineering
University of Texas at Arlington
Arlington, TX 76019-0015 U.S.A.
celebi@cse.uta.edu

Y. Alp Aslandogan
Dept. of Computer Science and Engineering
University of Texas at Arlington
Arlington, TX 76019-0015 U.S.A.
alp@cse.uta.edu

## Abstract

*In this work, we develop a system for retrieving medical images with focus objects incorporating models of human perception. The approach is to guide the search for an optimum similarity function using human perception. First, the images are segmented using an automated segmentation tool. Then, 20 shape features are computed from each image to obtain a feature matrix. Principal component analysis is performed on this matrix to reduce the number of dimensions. Principal components obtained from the analysis are used to select a subset of variables that best represents the data. A human perception of similarity experiment is designed to obtain an aggregated human response matrix. Finally, an optimum weighted Manhattan distance function is designed using a genetic algorithm utilizing the Mantel test as a fitness function. The system is tested for content-based retrieval of skin lesion images. The results show significant agreement between the computer assessment and human perception of similarity. Since the features extracted are not specific to skin lesion images, the system can be used to retrieve other types of images.*

## Keywords

Content-based image retrieval, human perception, pattern recognition, multidimensional access methods.

## 1. Introduction

Medical information systems with advanced browsing capabilities play an important role in medical training, research, and diagnostics. Since images represent an essential component of the diagnosis, it is natural to use medical images to support browsing and querying of medical databases [20].

Current content-based retrieval systems use low-level image features based on color, texture, and shape to represent images. However, except in some constrained applications such as human face and fingerprint recognition, these low-level features do not capture the high-level semantic meaning of images [26]. In order to bridge the gap between the low-level features and high-level semantics, researchers have proposed to compute increasingly high level features based on simpler ones [X]. However, another aspect that is as important as the features themselves has been neglected: The processing and interpretation of those features by human cognition. Although the ultimate goal of all image similarity metrics is to be consistent with human perception, little work has been done to systematically examine this consistency. Commonly, the performance of similarity metrics is evaluated based on anecdotal accounts of good and poor matching results [10].

In this work, we develop a system for retrieving medical images with focus objects incorporating models of human perception. The approach is to guide the search for an optimum similarity function using human perception.

Figure 1 shows an overview of the system. First, the images are segmented using an automated segmentation tool. Then, 20 shape features are computed from each image to obtain a feature matrix. Principal component analysis is performed on this matrix to reduce its dimension. The principal components obtained from the analysis are used to select a subset of variables that best represents the data. A human perception of similarity experiment is designed to obtain an aggregated human response matrix. Finally, an optimum weighted Manhattan distance function is designed using a genetic algorithm utilizing the Mantel test as a fitness function.

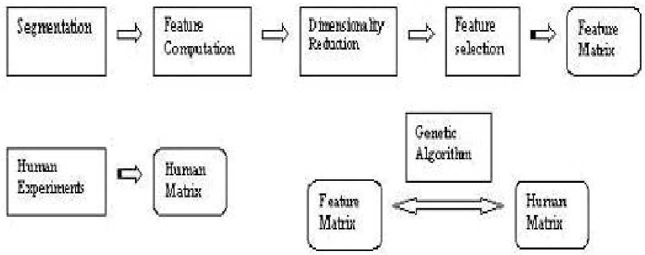

**Figure 1. System Overview**

The system is tested for content-based retrieval of skin lesion images. The results show significant agreement between the computer assessment and human perception of similarity. Since the features extracted are not specific to skin lesion images, the system can be used to retrieve other types of images.

## 2. Segmentation and Feature Extraction

We use 184 skin lesion images obtained from various online image databases [4,5] as our data set. All images have a resolution of about 500 pixels per centimeter.

### 2.1 Segmentation

Segmentation is partitioning of an image into regions that have similar characteristics such as color, texture, etc. It is an extremely important step in image retrieval since accurate computation of shape features depends on good segmentation [26]. For segmentation of lesion images we have used an automated tool, SkinSeg, developed by Goshtasby et al. [29]. Some of the segmented lesion images are shown in figure 2.

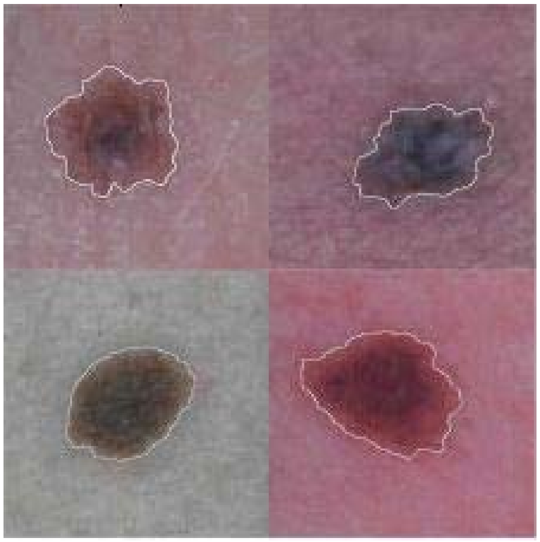

**Figure 2. Segmented skin lesion images**

### 2.2 Feature Computation

The ABCD rule of dermatoscopy [22], recommended by American Cancer Society, summarizes the clinical features of pigmented lesions suggestive of melanoma (a deadly form of skin cancer) by: asymmetry (A), border irregularity (B), color variegation (C) and diameter greater than 6 mm (D). Interestingly, three of these features are shape-based features. For each image in the database, we compute 20 shape features: Diameter, bending energy, contour sequence moments, convex hull area and perimeter, lesion area and perimeter, convexity, solidity, compactness, major and minor axis length, eccentricity, orientation, extent, asymmetry, and border irregularity.

## 3. Dimensionality Reduction and Feature Selection

After the feature computation step, we have 20 dimensional data to be analyzed. Well known problems associated with high dimensionality include (a) high computational cost, (b) classifier accuracy degradation, known as Hughes phenomenon, and (c) difficulty in visualization.

Dimensionality reduction can be achieved in two different ways. One approach, called feature subset selection, is selection of a subset of features that minimizes intra-class variance and maximizes inter-class variance. The other approach, called feature extraction, is linear or non-linear transformation of the original feature vector that optimizes a particular class separability criterion.

### 3.1 Principal Component Analysis of Skin Lesion Data

Principal Component Analysis (PCA) is an unsupervised linear feature extraction technique that transforms a set of variables into a substantially smaller set of uncorrelated variables that represents most of the information in the original set of variables [7].

Our data consists of 184 samples each having 20 features. Since features have arbitrary units, PCA is applied on the correlation matrix. The eigenvalue and the percentage of explained variance for each PC are given in table 1.

**Table 1. Eigenvalues and Explained Variances**

| | | % | % Var. |
|---|---|---|---|
| PC | Eigenvalue | Variance | Cum. |
| 1 | 7.17871 | 35.89 | 35.89 |
| 2 | 4.06138 | 20.31 | 56.2 |
| 3 | 3.17062 | 15.85 | 72.05 |

| | | | |
|---|---|---|---|
| 4 | 1.53324 | 7.67 | 79.72 |
| 5 | 1.03291 | 5.16 | 84.88 |
| 6 | 0.91134 | 4.56 | 89.44 |
| 7 | 0.82177 | 4.11 | 93.55 |
| 8 | 0.44458 | 2.22 | 95.77 |
| 9 | 0.38006 | 1.9 | 97.67 |
| 10 | 0.21225 | 1.06 | 98.73 |
| 11 | 0.09135 | 0.46 | 99.19 |
| 12 | 0.05471 | 0.27 | 99.46 |
| 13 | 0.0427 | 0.21 | 99.68 |
| 14 | 0.03532 | 0.18 | 99.85 |
| 15 | 0.018 | 0.09 | 99.94 |
| 16 | 0.00618 | 0.03 | 99.98 |
| 17 | 0.00298 | 0.01 | 99.99 |
| 18 | 0.00154 | 0.01 | 100 |
| 19 | 0.00027 | 0 | 100 |
| 20 | 0.00012 | 0 | 100 |

There are several rules of thumb for deciding how many PCs to retain [7]:

1) Kaiser [15] recommends discarding PCs of a correlation matrix having eigenvalues less than 1.

2) Jolliffe [14] argues that Kaiser's rule generally retains too few variables, that is, throws away too much information and suggests a cut-off of 0.7 instead of 1.

3) Including enough PCs to account for a given percentage of variation, e.g. 80%.

These rules should be used cautiously. For example, in some cases, Jolliffe's criterion results in retaining twice as many components as Kaiser's criterion. The more PCs, relative to the number of variables, retained, the better is the description of the data. Also, smaller PCs are, generally, harder to interpret than larger ones [7].

In our analysis, we choose to retain 8 principal components. From table 3.1 it can be seen that 8 PCs account for 95.77% of variation. On the other hand, Kaiser's criterion would retain 5 PCs explaining 84.88% of variation, whereas Jolliffe's criterion would retain 7 PCs explaining 93.55% of variation.

It is useful to examine the sums of squares of the loadings for each row of the principal component loading matrix because the row sum of squares indicates how much variance for that variable is explained by the retained PCs [7]. The percentage of variance explained for each variable is presented in table 2.

**Table 2. Explained Variances for each Variable**

| Variable | 5 PCs | 7 PCs | 8 PCs |
|---|---|---|---|
| 1 | 99 | 99 | 98 |
| 2 | 95 | 90 | 58 |
| 3 | 96 | 90 | 87 |
| 4 | 98 | 97 | 94 |
| 5 | 89 | 73 | 67 |
| 6 | 98 | 98 | 96 |
| 7 | 98 | 98 | 97 |
| 8 | 100 | 99 | 99 |
| 9 | 98 | 98 | 98 |
| 10 | 99 | 99 | 99 |
| 11 | 96 | 96 | 78 |
| 12 | 86 | 83 | 78 |
| 13 | 99 | 99 | 92 |
| 14 | 89 | 89 | 46 |
| 15 | 99 | 98 | 91 |
| 16 | 99 | 99 | 98 |
| 17 | 98 | 97 | 97 |
| 18 | 95 | 93 | 90 |
| 19 | 100 | 99 | 75 |
| 20 | 85 | 76 | 60 |

An examination of table 3.2 shows that 5 PCs retained using Kaiser's criterion are insufficient in explaining the variation in features 2 (58 %), 5 (67 %), 11 (78 %), 12 (78 %), 14 (46 %), 19 (75 %), and 20 (60 %). 7 PCs retained using Jolliffe's criterion are weak in representing the variation in features 5 (73 %) and 20 (76 %). On the other hand, 8 PCs account for more than 90% of variation in 16 of the features and more than 85% of variation in all features.

## 3.2 Using Principal Components to Select a Subset of Variables

Substantial dimensionality reduction can be achieved using n << m PCs instead of m variables, but usually the values of all m variables are still needed to calculate the PCs, since each PC is a linear combination of all m variables [14]. Some variables may be difficult or expensive to measure therefore, collecting data on them in future studies may be impractical [7]. Furthermore, while the original variables are readily interpretable, the constructed PCs may not be easy to interpret. Therefore, it might be preferable if, instead of using n PCs, we could use n of the original variables, to account for most of the variation in the data [14].

If the correlations among the variables are high, in many cases, we can represent the variation in the original set of variables by a much smaller subset of variables. Dunteman discusses two methods for variable subset selection [7]:

1) Starting with the smallest discarded PC, delete the variable with the highest loading on the relevant PC. If the variable had been already deleted, then the variable with the next highest loading would be deleted. Totally, m-n variables are deleted this way. This is called Jolliffe's B2 Method [14].

2) Starting with the largest retained PC, select the variable with the highest loading on the relevant PC to represent

that component, unless it has been chosen to represent a larger PC. In this way, a total of n variables are retained. This is called Jolliffe's B4 Method [14].

We use Jolliffe's B4 Method to retain the following variables: 10 (Perimeter), 12 (Solidity), 18 (Eccentricity), 6 (4th Moment), 19 (Orientation), 20 (Extent), 14 (Asymmetry), and 5 (3rd Moment) and discard the remaining ones.

The total amount of variation the selected variables account for can be used as a criterion to evaluate the efficiency of a particular subset of variables in representing the original set of variables [7]. The total amount of variation that a subset of variables explains is the sum of the variation they explain in each of the discarded variables plus the sum of the variances for the variables comprising the subset. Each discarded variable is regressed on the retained variables and the corresponding squared multiple correlations are summed. If we add to that the variances of the retained variables, in our case 1 for each variable, we can obtain a measure of the total amount of variation that a subset of variables explains. This can be formulated as:

$$n + \sum_{i=1}^{m-n} R^2 (i)$$

where $R^2$ (i) is the squared multiple correlation of the ith discarded variable with the retained variables.

The percentage of variation that the selected subset of variables explains is 78.98% if we retain n=5 PCs (Kaiser's criterion), 85.67% if we retain n= 7 PCs (Jolliffe's criterion) and 88.29% if we retain n= 8 PCs. Therefore, retaining 8 PCs seems to be the best choice.

## 4. Human Perception of Similarity Experiments

Since the ultimate user of an image retrieval system is human, the study of human perception of image content from a psychophysical level is crucial [26]. However, few content-based image retrieval systems have taken into account the characteristics of human visual perception and the underlying similarities between images it implies [11]. In their paper, Payne and Stonham state that if perceptually derived criteria and rank correlation are used to evaluate textural computational methods, retrieval performances are typically 25% or less, unlike the 80%-90% matches often quoted [23].

We design a human perception of similarity experiment to incorporate human perception into our shape-based image retrieval system. Figure 3 shows a snapshot of the graphical user interface of the experiment. The image on the left is the reference image, and the one on the right is the test image.

Since the image features are based on shape this is a "shape similarity" experiment. To focus the subjects only on shape similarity we binarized the lesion images so that there is no color or texture information in them.

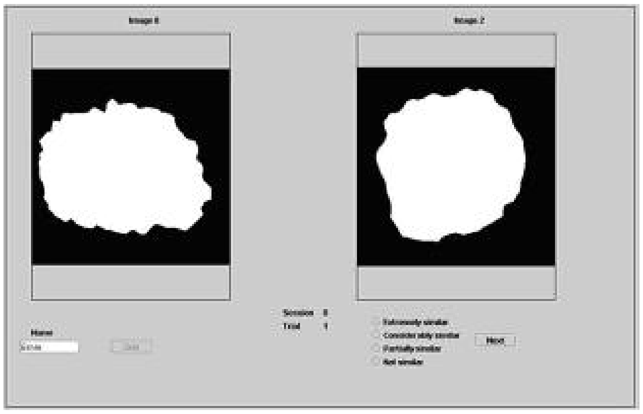

**Figure 3. Experiment GUI**

The subjects are expected to rate the similarity between pairs of images on a scale of four: Extremely similar, considerably similar, partially similar, and not similar. This scale is adapted from a psychophysical study [16]. Note that "Not similar" does not mean absolute dissimilarity; instead the term "not" is akin to "barely" or "hardly" in this context.

Since the number of trials is proportional to the square of the number of images, using all of the 184 images in the database is impractical. Therefore we select 48 images from the database to use in the experiment. This means C (48, 2)= 1128 trials.

The experiment consists of 3 sessions each comprised of approximately 376 trials. Nine subjects participated in the experiments. Each session took about 19 minutes on the average.

We construct the human response matrix following the approach described in Guyader et al. [11]. The overall dissimilarity matrix S is the weighted average of the individual similarity matrices $S_K=\{S_K (i\ j)\}$, one for each level of similarity. (Matrix $S_1$ is for "Extremely similar", $S_2$ for "Considerably similar", $S_3$ for "Partially similar", and $S_4$ for "Not similar"). Each time a subject associates a test image j to a reference i, we increase the corresponding $S_K$ (i j) by 1. We take the weighted average of $S_K$ to obtain the overall dissimilarity matrix S as follows:

$$S (i,j)= (1/64)*S_1 (i,j)+(1/27)*S_2 (i,j)+(1/8)*S_3 (i,j)+S_4 (i,j)$$

Note that weighting is necessary while obtaining S since we want to emphasize dissimilarity more than similarity. Since S is a dissimilarity matrix, a large entry represents a pair of dissimilar images, whereas a small entry corresponds to a pair of similar images.

## 5. Optimization of the Similarity Function Using a Genetic Algorithm

Genetic algorithms (GAs) are stochastic and adaptive optimization methods. Their advantages are:

1) They can be applied to an extremely wide range of problems [6].

2) They are simple to understand and easy to code [6].

3) They can be applied to general NP-complete problems such as job-shop scheduling, timetabling, traveling salesman, and facility layout problems [18].

4) They are inherently parallel, since various regions of the search space are explored simultaneously [9].

5) They are effective in situations where the search space is mathematically uncharacterized and not fully understood [12].

6) They can handle high-dimensional, nonlinear optimization problems [4].

We have used Parallel Genetic Algorithm Library (PGAPack) developed by David Levine [17] to implement our algorithm. The parameters of the algorithm are given in table 3. These parameters are determined empirically.

**Table 3. GA Parameters**

| Parameter | Value |
|---|---|
| Representation | Integer |
| Chromosome length | 8 genes |
| Population size | 2000 |
| Termination criterion | 2000 iterations |
| Parent selection strategy | Tournament selection |
| Crossover type | 2-point crossover |
| Crossover probability | 0.85 |
| Mutation type | Permutation |
| Population replacement strategy | Steady-state replacement |

### 5.1 Mantel Test

The Mantel test is a statistical technique for comparing two n x n dissimilarity (or similarity) matrices. It involves a measure of the association between the elements in two matrices by a statistic r, and then evaluates the significance of this measure by comparing it with the distribution of the values found by randomly reallocating the order of the elements in one of the matrices [1].

The statistic used for measuring the correlation between two matrices A and B is the classical Pearson coefficient:

$$r = \frac{1}{N-1}\sum_{i=1}^{N}\sum_{j=1}^{N}\left[\frac{(A_{ij}-\overline{A})}{s_A}\right]\left[\frac{(B_{ij}-\overline{B})}{s_B}\right] \quad \text{(Equation 1)}$$

where N is the number of elements in the lower or upper triangular part of the matrix, $\bar{A}$ and $S_A$ are the mean and standard deviation of elements of A, respectively.

Suppose we have two symmetric dissimilarity matrices A and B of size n x n. The testing procedure for the simple Mantel test is as follows [3]:

1) Compute the Pearson correlation coefficient $r_{AB}$ using eq. 1.

2) Permute randomly rows and the corresponding columns of one of the matrices, creating a new matrix A'.

3) Compute the $r_{A'B}$ statistic between matrix A' and matrix B using equation 1.

4) Repeat steps 2 and 3 a great number of times (>5000). The number of repeats determines the overall precision of the test.

We have used a GNU general public license software "zt" for performing the Mantel test [3]. As mentioned before, we use the Mantel test as the fitness function of our genetic algorithm. In other words, the output of the Mantel test, which is the correlation between two matrices, is used as a fitness value indicating the goodness of a particular set of weights. Since the Mantel test works with symmetric dissimilarity (or similarity) matrices we need to transform these matrices to symmetric dissimilarity ones. In each generation, we can transform the feature matrix to a dissimilarity matrix by taking the weighted Manhattan distances between each row using the gene values of the fittest individual in that generation as weights. Since Manhattan distance is a symmetric function, this matrix is guaranteed to be symmetric. The human response matrix is already a dissimilarity matrix. But, it is not necessarily symmetric. To symmetrize this matrix, we can take the average of symmetric entries (i.e., [i,j] and [j,i]).

### 5.2 Results of the Optimization

Initially (i.e., with all weights equal to 1.0), the correlation between the aggregated human matrix and the feature matrix is 0.56. After optimization the correlation becomes 0.73. This means using optimization we achieve 31% improvement in the agreement between the computer assessment and human perception. Figure 4 shows a selection of matching results.

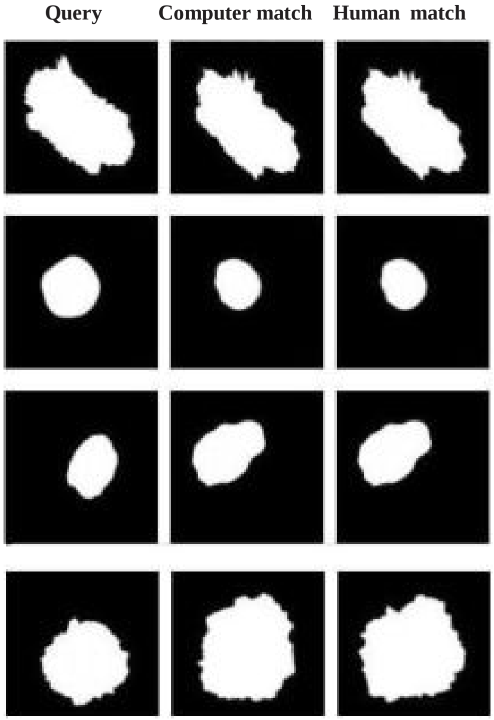

**Figure 4. A selection of matching results**

## 6. Related Work

To the authors' knowledge relatively little work has been done to incorporate human perception of similarity in CBIR systems in a systematic manner.

For the specific case of shape similarity Scassellati et al. [27] have used shape similarity judgments from human subjects to evaluate the performance of several shape distance metrics.

Frese et al. [10] have developed a methodology for designing similarity metrics based on human visual system. The metric they propose is based on color and spatial attributes. Therefore, their method cannot be used for general similarity based image retrieval.

Rogowitz et al. [25] have studied how human observers judge image similarity. They conduct two psychophysical scaling experiments aimed at uncovering the dimensions human observers use in rating the similarity of photographic images and compare the results with two algorithmic image similarity methods. Although their work provides insight into how humans judge similarity, this information is not used to improve the performance of a CBIR system.

Mojsilovic et al. [19] have developed a perceptually based image retrieval system based on color and texture attributes. They perform subjective experiments and analyze the results using multidimensional scaling to extract relevant dimensions. They also design distance metrics for color and texture matching. However, these metrics are not optimized for maximum agreement with human perception of similarity.

Guyader et al. [11] have developed a natural scene classification system incorporating human perception based on Gabor features. However, their model is suitable for only global similarity retrieval. For example, classes such as people or animals cannot be discriminated.

## 7. Conclusion

Content-based image retrieval has been an active research area during the last two decades. Since the early 90s numerous image retrieval systems, both research and commercial, have been developed [21, 24, 28, 13]. The main contribution of this work is the incorporation of human perception into this task in a systematic and generalizable manner.

In this work, we used aggregated human perception as a guide in optimizing a visual similarity function in a content-based image retrieval system. A psychophysical experiment was designed to measure the perceived similarity of each image with every other image in the database. The weights of the similarity function were optimized by means of a genetic algorithm using the dissimilarity matrix obtained from the human experiments. As a result of optimization, agreement between the computer assessment and human perception is improved by 31 percent.

In this system we focus on shape similarity. However, the same approach can be used to develop similarity functions based on other low-level features such as color or texture. Also, for general similarity based retrieval, another content-based image retrieval system powerful in color or texture aspects can be combined with our system.

Currently the human experiment requires approximately one hour to be completed. Further work needs to be done to minimize the experiment duration in order to include experts in the experiments.

In this system, indexing aspect necessary for efficient retrieval has not been considered. For this purpose, efficient multidimensional indexing techniques such as SS-trees, SR-trees or X-trees can be used.

## 8. References

[1] Besancon R. and Rajman M. (2002) "Evaluation of a Vector Space Similarity Measure in a Multilingual Framework." Proceedings of the Third International Conference on Language Resource and Evaluation

[2] Bezzant L. "Dermatology Image Bank." University of Utah Health Sciences Library, Available at http://www-medlib.med.utah.edu/kw/derm/

[3] Bonnet E. and Peer Y. (2002) "zt: A Software Tool for Simple and Partial Mantel Tests." Journal of Statistical Software, 7: 1-12.

[4] Chambers L., ed. (2001) "The Practical Handbook of Genetic Algorithms, Applications." Chapman & Hall/CRC.

[5] Cohen B. and Lehmann C. U. "DermAtlas" Johns Hopkins University of Medicine, Available at http://dermatlas.med.jhmi.edu/derm

[6] Coley D. A. (1999) "An Introduction to Genetic Algorithms for Scientists and Engineers." World Scientific Publishing Company.

[7] Dunteman G. H. (1989) "Principal Component Analysis." Sage Publications.

[8] Eiben, A.E. (2002) "Evolutionary computing: the most powerful problem solver in the universe?" Dutch Mathematical Archive, 5: 126-131.

[9] Filho J., Treleaven P., and Alippi C. (1994) "Genetic Algorithm Programming Environments." IEEE Computer 27: 28-43.

[10] Frese T., Bouman C. A., and Allebach J.P (1997) "A Methodology for Designing Image Similarity Metrics Based on Human Visual System Models." Proceedings of the SPIE Conference on Human Vision and Electronic Imaging II, 3016: 472-483.

[11] Guyader N., Herve L. B., Herault J., and Guerin A. (2002) "Towards the Introduction of Human Perception in a Natural Scene Classification System." IEEE International Workshop on Neural Network for Signal Processing.

[12] Hussein F., Kharma N., Ward R. (2001) "Genetic Algorithms for Feature Selection and Weighting, a Review and Study." Sixth International Conference on Document Analysis and Recognition, pp. 1240-1244.

[13] Iqbal Q. and Aggarwal J. K (2002) "CIRES: A System for Content-based Retrieval in Digital Image Libraries." ICARCV, pp. 205-210.

[14] Jolliffe I. T. (1986) "Principal Component Analysis." Springer-Verlag New York Inc.

[15] Kaiser H. F. (1960) "The Application of Electronic Computers to Factor Analysis." Educational and Psychological Measurement, 20: 141-151

[16] Kurtz D. B., White T. L., and Hayes M. (2000) "The labeled dissimilarity scale: A metric of perceptual dissimilarity." Perception & Psychophysics, 62: 152-161

[17] Levine D. (1996) "Users Guide to the PGAPack Parallel Genetic Algorithm." Available at ftp://ftp.mcs.anl.gov/pub/pgapack/user_guide.ps

[18] Michalewicz Z., Deb K., Schimdt M., and Stidsen T. (1999) "Evolutionary Algorithms for Engineering Applications." Proceedings of EUROGEN'99, pp. 73-94

[19] Mojsilovic A., Kovacevic J., Hu J., and Safranek R. J., Ganapathy S. K. (2000) "Matching and Retrieval Based on the Vocabulary and Grammar of Color Patterns." IEEE Transactions on Image Processing, 9: 38-54

[20] Mojsilovic A. and Gomes J. (2002) "Semantic based categorization, browsing and retrieval in medical image databases." IEEE International Conference on Image Processing, 3: 24-28

[21] Niblack W., Barber R., Equitz W., Flickner M., Glasman E. H., Petkovic D., Yanker P., Faloutsos C., Taubin G. (1993) "The QBIC Project: Querying Images by Content, Using Color, Texture, and Shape." Proceedings of the SPIE Conference on Storage and Retrieval for Image and Video Databases, 173- 187

[22] NIH Consensus Conference (1992) "Diagnosis and treatment of early melanoma." JAMA pp. 1314- 1319

[23] Payne J. S. and Stonham T. J. (2001) "Can Texture and Image Content Methods Match Human Perception?" Proceedings of 2001 International Symposium on Intelligent Multimedia, Video, and Speech Processing

[24] Pentland A., Picard R. W., and Sclaroff S. (1996) "Photobook: Content-based manipulation of image databases." Int. Journal of Computer Vision, 18: 233-254

[25] Rogowitz B. E., Frese T., Smith J. R., Bouman C. A., and Kalin E. (1998) "Perceptual Image Similarity Experiments." Proceedings of the SPIE Conference on Human Vision and Electronic Imaging 3299: 26-29

[26] Rui Y., Huang T. S., and Chang S. (1999) "Image Retrieval: Current Techniques, Promising Directions and Open Issues" Journal of Visual Communication and Image Representation, 10: 39-62

[27] Scassellati B., Alexopoulos S., and Flickner M. (1994) "Retrieving images by 2D shape: a comparison of computation methods with human perceptual judgments." in Niblack W. and Jain R. (Eds.), Proceedings of 2nd Storage and Retrieval for Image and Video Databases Conference

[28] Wang J. Z., Li J., and Wiederhold G. (2001) "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture Libraries." IEEE Transactions on Pattern Analysis and Machine Intelligence, 23: 947-963

[29] Xu L, Jackowski M., Goshtasby A., Roseman D., Bines S., Yu C., Dhawan A., and Huntley A. (1999) "Segmentation of Skin Cancer Images." Image and Vision Computing, 17: 65-74