We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

Content-based Image Retrieval Using Constrained Independent Component Analysis: Facial Image Retrieval Based on Compound Queries

Tae-Seong Kim[1] and Bilal Ahmed[2]
[1]Department of Biomedical Engineering
[2]Department of Computer Engineering
College of Electronics and Information
Kyung Hee University
Republic of Korea

1. Introduction

Visual information plays a crucial role in various domains, from medical diagnosis, journalism, crime-prevention to surveillance. Whereas domain specific images carry specific semantics, the problem of interpreting visual information becomes more complex when we talk of natural images. The maxim, 'A picture is worth a thousand words' explains this inherent problem very concisely. Indexing large databases of images for efficient retrieval is crucial for various domains such as journalism, biomedicine, forensics etc. Manual indexing of images in such large databases can be highly subjective and time consuming. In contrast, content-based image retrieval (CBIR) focuses on the development of efficient retrieval mechanisms based on image features or meta-data used for image annotation.

Conventional approaches to CBIR represent images in the form of image-based features. These features vary from global image descriptors such as color or intensity histogram to local ones such as shape and texture. These features along with their combinations have been used previously for CBIR. For example, in (Deng et al., 2001), a region-based color-descriptor, modelling the color values along with their percentages in the region, is proposed. Similarly in (Hadjidemetriou et al., 2004), multi-resolution histograms have been employed for the retrieval of textured images. In (Jeong et al., 2004), the extraction of color histograms through Gaussian mixture vector quantization has been proposed. In (Belongie et al., 2002) and (Petrakis et al., 2002) respectively, shape descriptors and shape matching algorithms have been proposed for image retrieval.

The use of low-level image features such as color histograms, shape, and texture attributes introduces a semantic gap (Chen et al., 2004). This semantic gap arises due to the inability of such low-level features to describe the objects and their inter-relations within the image. The use of such low-level features places the responsibility of achieving semantically coherent results on the user-interface. Various techniques of relevance feedback (Rui et al., 1998) have been introduced in this context. Whereas user feedback might be able to lower this gap, the overall procedure becomes subjective and requires a higher degree of user interaction.

Segmentation-based techniques for image retrieval have also been used for obtaining better shape, texture, and color descriptions of the image contents (Datta et al., 2005). The motivation behind this particular approach is that objects within an image can be segmented and used for querying the database to retrieve more semantically similar images. Various segmentation techniques, such as the Normalized Cuts (Shi & Malik, 2000), Mean Shift Procedures, and Expectation Maximization (Carson et al., 2002) algorithms have been used in image retrieval. Machine-learning approaches augmented with segmentation techniques have also been used. In (Wang et al., 2001), segmentation results augmented with fuzzy logic are used to obtain soft similarity measures. The problem of obtaining a semantically coherent segmentation of an image still remains an open research problem and higher dependency on segmentation-results is not desirable for achieving a semantically accurate retrieval performance.

From an image-retrieval point of view, facial images have attracted a lot of attention. Various machine learning and feature extraction techniques have been employed for the efficient retrieval of facial images. Earlier retrieval systems, such as the Photobook (Pentland et al., 1994), use Principal Component Analysis (PCA) for the retrieval of facial images. In (Liu, 2004), feature extraction through Independent Component Analysis (ICA) in a reduced PCA space is used for characterizing query images. The overall system comprises of classifying the input query image based on the nearest-neighbor rule using various similarity measures. A recent work on facial image retrieval by (Basak et al., 2006) has focused on representing facial images as a collection of local independent components. For this purpose, the query images are decomposed into a number of overlapping and non-overlapping windows to compute the independent components.

The use of multiple images as a compound query has not been explored in much detail. Conventional CBIR systems do not provide a mechanism through which a user can specify his search criterion through multiple examples. This is analogous to multi-word queries in search-engines: the specification of a compound query helps the system in retrieving the desired results with better accuracy. Similarly, when a user cannot find a single image which can specify his search criterion, he should be able to use multiple images to formulate his query. Multiple queries have been used in (Tahagogi et al., 1994): the approach taken is to find a combined result of the query by using the retrieved images corresponding to each query image independently. Similarly, (Basak et al., 2006) uses multiple facial images to retrieve images similar to the independent query-images as well as to their combinations.

In this work, we have devised a system which can cater for both single and multiple exemplar image retrieval. It does not decompose the query images or the database images to windows as in (Basak et al., 2006) or uses PCA for dimension reduction (Liu, 2004): thus the chances of any information loss are minimal. There is also no need to store additional feature information or the need for any offline learning as in (Basak et al., 2006). Our approach is centered on the idea of constrained ICA (cICA) (Lu & Rajapakse, 2005) which has the ability to extract specific independent components conforming to certain prior information (known as reference signals or images). Query images are provided to the constrained ICA algorithm as references, and the output of the constrained ICA algorithm specifies the contribution of each database image to the extracted component. Based on the magnitude of this contribution factor, the database images are ranked for retrieval.

The rest of the chapter gives details about the constrained ICA-based image retrieval system using multiple query images. Section 2 explains the constrained ICA framework and Section 3 describes the whole system in detail. Experimental results are given in Section 4. Finally, we conclude in Section 5.

2. Constrained ICA

Conventional ICA techniques perform blind-source separation (BSS) assuming a linear mixing model of the independent sources. If the observed values at a pixel location from a set of images are represented as $\mathbf{x} = (x_1,...,x_n)^T$ and the original sources as $\mathbf{s} = (s_1,...,s_m)^T$. ICA assumes that each $\mathbf{x}$ is a linear mixture of the original independent sources. Therefore,

$$\mathbf{x} = \mathbf{As} \tag{1}$$

where $\mathbf{A}$ is the mixing matrix of $n \times m$. Conventional ICA algorithms aim at finding a demixing matrix $\mathbf{W}$ to recover all the ICs of the observed image such that $\mathbf{s} = \mathbf{Wx}$. In general, existing ICA algorithms find as many ICs as the number of observations (i.e., $n = m$). The user must manually identify which ICs represent which sources. The primary reason for this manual intervention is the inability of the ICA algorithms to calculate the energies or signs of the ICs. This may also lead to problems where the number of sources is less than the number of observations. Deflation-based ICA techniques (Cichoki et al., 1997); (Hyvärinen & Oja, 1996) have also been developed, but they also suffer from the arbitrary ordering of the extracted independent components.

Constrained ICA (Lu & Rajapakse, 2005) has been developed to find only those independent components which are of interest to the current task at hand. This is achieved by providing some prior knowledge about these ICs to the constrained ICA algorithm. This prior information may not be exact, but it could be the specification of statistical properties of the desired component or just a crude approximation (e.g., template). Therefore, if we have some *a priori* information about the desired sources, we can incorporate this information into constrained ICA. The constrained ICA algorithm uses this *a priori* information about the desired IC, encoded into a set of reference images, $\mathbf{r} = (r_1,...,r_l)^T$ to obtain a set of output IC images, $\mathbf{y} = (y_1,...,y_l)^T$ which contains statistically independent extracted sources. The closeness constraint can be written as,

$$g(\mathbf{w}) = \varepsilon(y_i, r_i) - \xi \le 0 \tag{2}$$

where $\mathbf{w}$ is the weight vector to be learned, ε some closeness measure, and ξ an appropriate closeness threshold parameter. The measure of closeness can take any form, such as the mean squared-error (MSE), correlation, or any other suitable measure. The number of reference signals determines the number of independent components to be extracted from the complete set of observations. The final mathematical model for constrained ICA can be represented as,

$$\text{maximize} \sum_{i=1}^{l} J(y_i) \tag{3}$$
$$\text{subject to } g(\mathbf{W}) \le 0, h(\mathbf{W}) = 0$$

where,

$$J(y_i) \approx \rho[E\{G(y_i)\} - E\{G(v)\}]^2 \tag{4}$$

is the one-point contrast function for ICA introduced in (Hyvärinen et al., 2001). ρ is a positive constant, $G(\cdot)$ a non-quadratic function, and v a zero mean and unit variance Gaussian random variable. $h(\mathbf{W})$ constrains the output component to have unit variance. Equation (4) is a constrained optimization problem and can be solved by the augmented Lagrangian functions.

3. Constrained ICA-based facial image retrieval

Viewing it from another perspective, the constrained ICA framework can be used for specifying the type of information we would like to extract from huge amounts of data. The reference image(s) can be formulated as the query image(s) specified by the user and as the accuracy of the extracted information depends upon the accuracy of the provided references. In our case, the image(s) provided by the user would serve this purpose, and point the constrained ICA algorithm in the appropriate direction. The overall system architecture is depicted in Fig. 1.

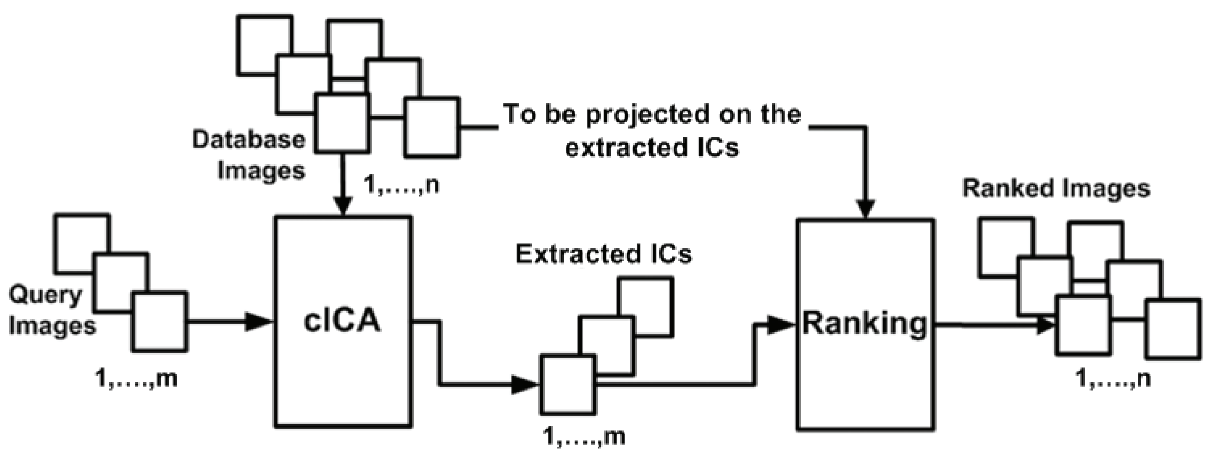


Fig. 1. Overview of the constrained ICA-based CBIR system.

Since constrained ICA extracts components y_i from the given set of observations corresponding to the provided reference image(s), we can ascertain the contribution of each observation by reconstructing it from the extracted component. The reconstruction procedure involves the estimation of the mixing matrix $\mathbf{A}$ and the reconstruction of the entire set of observations. Consider that we have n observations $\mathbf{x} = (x_1,...,x_n)^T$ and m extracted sources $\mathbf{s} = (s_1,...,s_m)^T$ where $m << n$. The mixing matrix pertaining to the extracted sources with respect to the entire set of observations can be estimated using,

$$\mathbf{A} = \mathbf{x}\mathbf{s}^+ \tag{5}$$

where $\mathbf{s}^+$ is the pseudoinverse of the extracted sources. Furthermore the reconstruction of $\mathbf{x}$ can be done using,

$$\mathbf{x}_R = \mathbf{As} \tag{6}$$

where $\mathbf{x}_R$ is the set of reconstructed observations.

Once all images have been reconstructed with the extracted IC images, we need to estimate how well each image has been reconstructed from the extracted sources. This involves the comparison of each reconstructed image with its original image. Simple measures of similarity such as correlation or mutual-information can be used. In our case we have used correlation to determine the similarity between the original and the reconstructed images.

4. Experimental results

We have conducted extensive experiments on a publicly available facial-image database, the ORL face database (Samaria & Harter, 1994). The ORL database contains ten facial images of 40 individuals with varying pose, expressions, and spectacles. The images were scaled to 64×64 with no pre-processing or feature extraction.

For the purpose of evaluating our retrieval system, we have divided the queries into two categories: homogeneous and heterogeneous queries. Homogeneous queries contain the images of the same person with different pose, expression, or occlusion. Heterogeneous queries are composed of images of different subjects. The main motivation behind this formulation is to bring out the essence of fusing information from different independent images and evaluating the semantic coherence of the retrieved images.

4.1 Homogeneous queries

In order to retrieve the facial images of a single person from the database under varying pose and occlusion conditions (i.e., wearing spectacles), a single example might not be enough. The same is also true if the database has different expressions and scale. Fig. 2 (*left*) shows the results of using a single query. The database contains ten images of each individual: with a single query image, our system has been able to retrieve eight images in the top ten retrieved images. The images, which have been left out of the top ten: the image at (3rd row, 1st column) and image at (3,2) in Fig. 2 (*left*), have the same individual but with his head tilted to the right side. The query image given in Fig. 2 (*left*) was unable to describe the features present in these left-out images.

Fig. 2. Constrained ICA-based CBIR applied to the ORL database. (*Left*) A result with a single query: here the system acts as a face recognition system. (*Right)* A result using two images of the same individual with different pose.

In Fig. 2 (*right*), we have used two query images: one depicts the individual with a left tilt whereas the other depicts the same pose but in the opposite direction. All the ten relevant images have been retrieved from the database and have the highest ranking, as can be seen from the results. This particular case shows the fact that the constrained ICA-based retrieval technique is able to fuse features from two independent images and retrieve images in which the subject has a straight pose.

4.2 Heterogeneous queries

Fig. 3 shows the results obtained for two heterogeneous queries. In the first query depicted in Fig. 3 (*left*), two images of two different individuals have been used. One of them is wearing spectacles whereas the other has none. In the retrieved images, we see that the initial nine images correspond to the two individuals, where images of the second subject wearing spectacles are also given a higher rank. Similarly, after the two top rows, the system has retrieved images of individuals with and without spectacles and bearing some facial similarity to the individuals depicted in the query images.

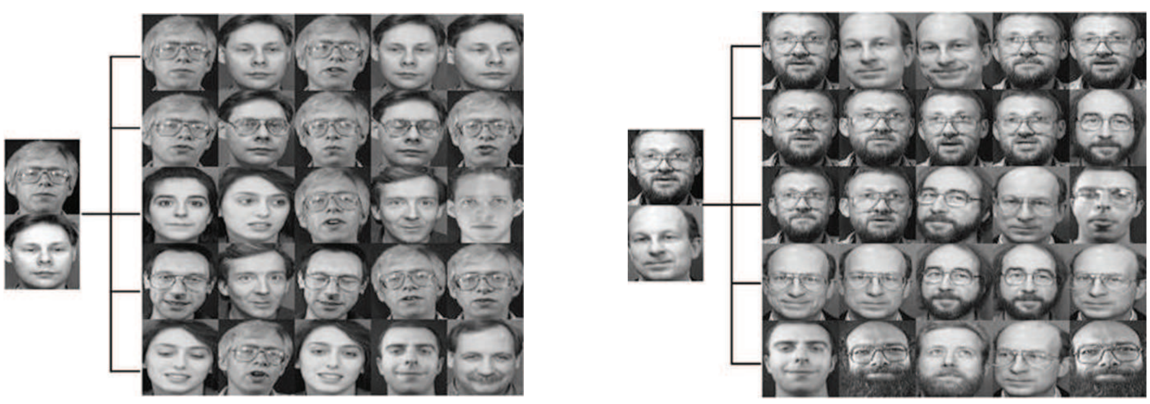


Fig. 3. Results for heterogeneous queries.

In the second case shown in Fig. 3 (*right*), again images of two different individuals are used. This time, the top query image has an individual who has a beard and spectacles. Whereas the other individual is clean shaven and has no spectacles. The images retrieved by the system not only contain the individuals present in the query but also their various combinations: persons having both beard and spectacles (the same as the individual in the top query images), persons having only beard, persons wearing only spectacles with no beard, and persons having none of these (corresponding to the individual depicted in the lower query image).

4.3 Image retrieval of covered faces

In the case of image retrieval, query images have a profound effect on the output of the system. It could be the case that the images available at the time of query formulation contain only partial information about the target image(s). As an example, consider the case of querying the database when the available face images are covered with some objects such as sunglasses or muffler. Fig. 4 *(left)* depicts a case of such a query in which the lower-half of the subject's face is covered with muffler. As the results show, the retrieved images are those in which all the subjects have their face covered in the same manner.
This situation can be alleviated by fusing information from another face image. The constrained-ICA based image retrieval framework allows for such information fusion

through the use of multiple-query images. Fig. 4 *(right)* shows the results for such a query: where along with the covered face, we now provide the system with another face image without occlusion. Based on this partial information, the system is able to retrieve the desired face image at (3rd row, 4th column). The system also retrieves the same subject, wearing sunglasses at (4,3).

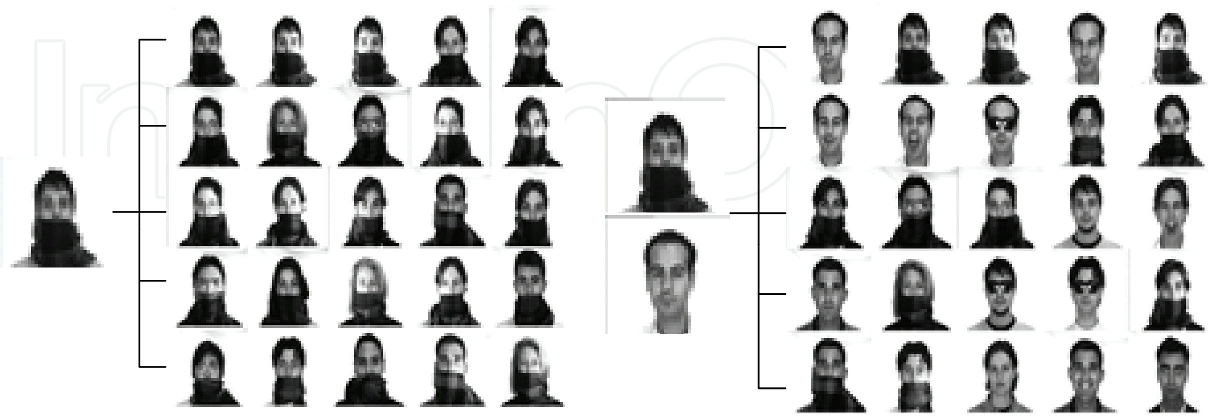


Fig. 4. Constrained ICA-based CBIR applied to the Aleix Face database (Martinez & Benavente, 1998). Two results are shown with a partially covered query image (*left*) and augmented query images with a normal face (*right*).

4.4 Performance analysis

We conducted one hundred simulations of the system with random query formulations for the homogeneous query case. A hard similarity evaluation was used: only the retrieved images pertaining to the same individual as depicted in the query were considered relevant as opposed to (Basak et al., 2006) where it was assumed that user feedback is available. Simple measures of *precision* (Baeza-Yates & Ribeiro-Neto, 1999) and *recall* (Baeza-Yates & Ribeiro-Neto, 1999) have been used to evaluate the efficiency of the system:

$$precision = {}^{N_{RL}}\!\big/_{N_R} \tag{7}$$

$$recall = {}^{N_{RL}}\!\big/_{N_{RD}} \tag{8}$$

where N_{RL} is the number of relevant images in the retrieved images, N_R the total number of retrieved images, and N_{RD} the total number of relevant images in the database. In the case of Fig. 2 (a), Precision = 10/25 and Recall = 10/10. Note that, when N_R equals N_{RD} the two measures become equal. This is the break-even point of the system and indicates its overall accuracy.

The evaluation measures for queries consisting of one, two, and three images are shown in Fig. 5. In the figure, T1, T2, and T3 represent the break-even points of the system for the queries formulated from one, two, and three images respectively. In the case of single-image queries, the system has achieved an accuracy of 76%. Whereas, in the case of compound queries composed of two and three images, this accuracy increases to 80% and 90% respectively. In contrast to the conventional systems, the constrained ICA-based retrieval system achieves this higher level of performance without any feature-extraction and offline-learning.

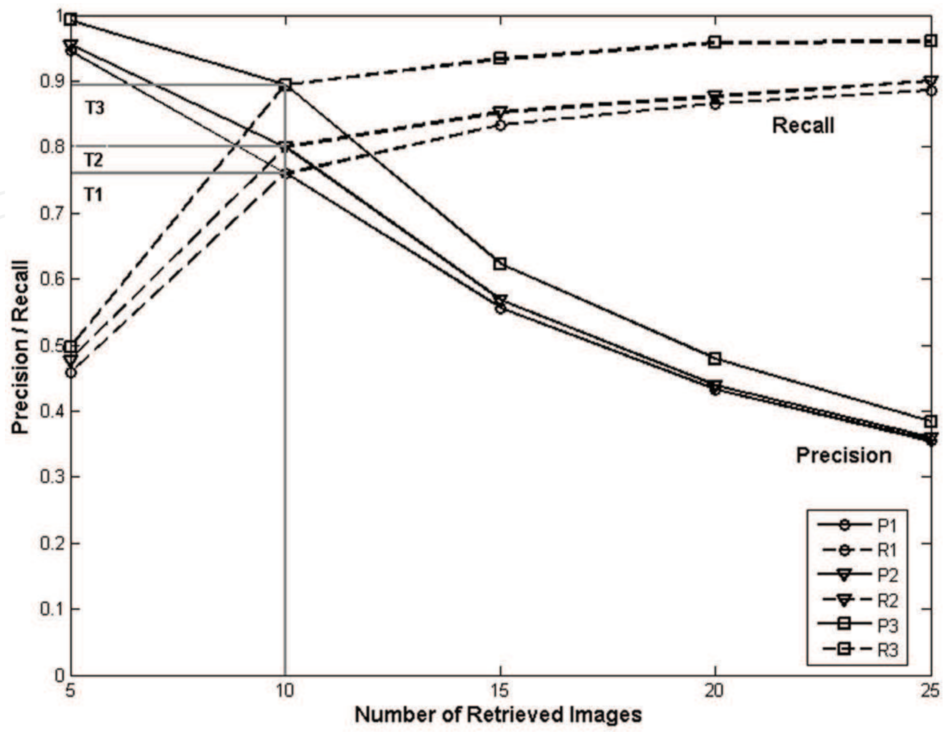


Fig. 5. Performance evaluation on homogeneous queries. The graph shows the precision (P) and recall (R) values for queries formulated with 1, 2, and 3 images.

5. Conclusion

In this work, we have proposed a new technique of facial image retrieval based on constrained ICA. Our technique requires no offline learning, pre-processing, and feature extraction. The system has been designed so that none of the user-provided information is lost, and in turn more semantically accurate images can be retrieved. As our future work, we would like to test the system in other domains such as the retrieval of chest x-rays and CT images.

6. Acknowledgement

This research was supported by the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Advancement) (IITA-2008-(C1090-0801-0002)).

7. References

Baeza-Yates, R. A. & Ribeiro-Neto, B. A. (1999). *Modern Information Retrieval*. ISBN: 020139829X, Addison-Wesley.

Basak, J.; Bhattacharya, K. & Chaudhury, S. (2006). Multiple Exemplar-Based Facial Image Retrieval Using Independent Component Analysis. *IEEE Transactions on Image Processing*, 15, 12, (December 2006) 3773–3783, ISSN: 1057-7149.

Belongie, S.; Malik, J. & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 4, (April 2002) 509–522, ISSN: 0162-8828.

Carson, C.; Belongie, S., Greenspan, H. & Malik, J. (2002). Blobworld: Image segmentation using expectation-maximization and its applications to image querying. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 8, (August 2002) 1026–1038, ISSN: 0162-8828.

Chen, Y.; Li, J. & Wang, J. Z. (2004). *Machine Learning and Statistical Modeling Approaches to Image Retrieval*. ISBN: 1402080344, Kluwer Academic Publishers.

Cichoki, A.; Thawonmas, R. & Amari, S. (1997). Sequential blind signal extraction in order specified by stochastic properties. *Electronics Letters*, 33, 1, (Janunary 1997) 64–65, ISSN: 0013-5194.

Datta, R.; Li, J. & Wang, J. Z. (2005). Content-Based image retrieval - approaches and trends of the new age. *Proceedings of the 7th ACM SIGMM International Workshop on Multimedia Information Retrieval*. Singapore, November 2005.

Deng, Y.; Manjunath, B. S., Kenney, C., Moore, M. S. & Shin, H. (2001). An efficient color representation for image retrieval. *IEEE Transactions on Image Processing*, 10, 1, (January 2001) 140–147, ISSN: 1057-7149.

Hadjidemetriou, E.; Grossberg, M. D. & Nayar, S. K. (2004). Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26, 7, (July 2004) 831–847, ISSN: 0162-8828.

Hyvärinen, A.; Karhunen, J. & Oja, E. (2001). *Independent Component Analysis*. ISBN: 047140540X, Wiley-Interscience.

Hyvärinen, A. & Oja, E. (1996). Simple neuron models for independent component analysis. *International Journal of Neural Systems*, 7, 6, (February 1996) 671–687, ISSN: 0129-0657.

Jeong, S.; Won, C.S. & Gray, R. M. (2004). Image retrieval using color histograms generated by Gaussian mixture vector quantization. *Computer Vision and Image Understanding*, 9,1-3, (April 2004) 44–66, ISSN: 1077-3142.

Liu, C. (2004). Enhanced independent component analysis and its application to content based face image retrieval. *IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics*, 34, 2, (April 2004) 1117–1127, ISSN: 1083-4419.

Lu, W. & Rajapakse, J. C. (2005). Approach and applications of constrained ICA. *IEEE Transactions on Neural Networks*, 16, 1, (January 2005) 203–212, ISSN: 1045-9227.

Martinez, A. M. & Benavente, R. (1998). The AR Face Database. *CVC Technical Report #24*, June 1998.

Pentland, A.; Picard, R. W. & Sclaroff, S. (1994). Photobook: tools for content-based manipulation of image databases. *In Proc. SPIE Storage and Retrieval for Image and Video Databases II*, San Jose, CA, USA, February 1994.

Petrakis, E. G. M.; Diplaros, A. & Milios, E. (2002). Matching and retrieval of distorted and occluded shapes using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 11, (November 2002) 1501 – 1516, ISSN: 0162-8828.

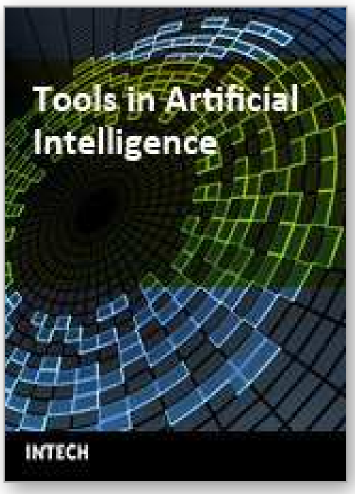
Rui, Y.; Huang, T. S., Ortega, M. & Mehrotra, S. (1998). Relevance feedback: A power tool for interactive content-based image retrieval. *IEEE Transactions on Circuits and Systems for Video Technology*, 8, 5, (September 1998) 644–655, ISSN: 1051-8215.

Samaria, F. & Harter, A. (1994). Parameterization of a stochastic model for human face identification. *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota FL, December 1994.

Shi, J. & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22, 8, (August 2000) 888–905, ISSN: 0162-8828.

Tahaghoghi, S. M. M.; Thom, J. A. & Williams, H. E. (2001). Are two pictures better than one?. *Proc. of the 12th Australasian Database Conference*, pp. 138–144, Gold Coast, Queensland, Australia, January 2001.

Wang, J. Z.; Li, J. & Wiederhold, G. (2001). SIMPLIcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 9, (September 2001) 947–963, ISSN: 0162-8828.



Tools in Artificial Intelligence
Edited by Paula Fritzsche

ISBN 978-953-7619-03-9
Hard cover, 488 pages
Publisher InTech
Published online 01, August, 2008
Published in print edition August, 2008

This book offers in 27 chapters a collection of all the technical aspects of specifying, developing, and evaluating the theoretical underpinnings and applied mechanisms of AI tools. Topics covered include neural networks, fuzzy controls, decision trees, rule-based systems, data mining, genetic algorithm and agent systems, among many others. The goal of this book is to show some potential applications and give a partial picture of the current state-of-the-art of AI. Also, it is useful to inspire some future research ideas by identifying potential research directions. It is dedicated to students, researchers and practitioners in this area or in related fields.

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821