An improved method for choosing effective Independent Component Filters for CBIR

Nabeel Mohammed
Faculty Of Information Technology
Monash University. Clayton Campus.
Wellington Road. Victoria 3800, Australia.
Email: Nabeel.Mohammed@monash.edu

David McG. Squire
Faculty Of Information Technology
Monash University. Clayton Campus.
Wellington Road. Victoria 3800, Australia.
Email: David.Squire@monash.edu

***Abstract*—Our main aim is to find methods to improve collecition-specific Content Based Image Retrieval (CBIR). We have used Independent Component Analysis for this purpose in a previous study. One of the issues with using ICA is the large number of filters that are extracted. Choosing a smaller subset of the extracted filters is challenging as it is desirable to choose filters which provide the most useful features without redundancy. A published technique uses measures of variance for this purpose. However this method has the drawback of choosing a set which may provide redundant features. In this study we propose a new method, using normalised cross-correlation followed by clustering. We have carried out a comparison of the effectiveness of the clustering based method with the variance based method. On our test data the filters chosen by clustering seem to represent underlying textures better and also perform better for CBIR.**

I. Introduction

This paper documents parts of our work in aiming to find more effective techniques for collection-specific Content Based Image Retrieval (CBIR). Our work in [1] documents the use of Independent Component Analysis (ICA) for this purpose. One of the challenges we found during this work is the issue of the large number of components that ICA extracts, which can be potentially used as filters. To make CBIR more practical we need to choose a smaller set of filters. It is desirable to choose a set of filters where each filter would give us different but useful image features. A documented method for choosing such a set of filters is the use of normalized variance [2]. However while reviewing the method we decided that it may not be suitable for the purposes of CBIR due to an issue pointed out by [3], where the selected filters might give us very similar features. For [1], we used image cross-correlation followed by clustering to select a smaller set of filter. However no comparison was done to ensure that the our method did indeed perform better than the use of variance. The work outlined in this paper bridges that gap. We perform a comparison of effectiveness of the filter selection methods by analysing the effectiveness of the selected filters for CBIR. If we can correctly identify the problem with the method discussed in [2] and show that our proposed method solves that problem, then it would further our initial aim of making collection-specific CBIR more effective. The next section provides a brief background of the current work. We then describe the experiments and the different filter selection methods. Following this, the results are discussed and then we conclude the document.

II. Background

A. Independent Component Analysis (ICA)

ICA is defined as "a method for finding underlying factors or components from multivariate (multidimensional) statistical data" [4]. The initial motivation behind ICA was to perform Blind Source Separation (BSS), which refers to the task of discovering the source signals from some observed linear mixture of the sources [4]. In fact, BSS is a good example to use when describing ICA as a mathematical problem. Here a simple version of ICA is presented, where we assume that the number of observed signals and the number of source signals are equal. Let $x_1(t)$, $x_2(t)$ and $x_3(t)$ represent the observed signals of some source signals $s_1(t)$, $s_2(t)$ and $s_3(t)$ at time t. Based on this information, it can be said that for $i \in [1, 3]$,

$$x_i(t) = a_{i1}s_1(t) + a_{i2}s_2(t) + a_{i3}s_3(t). \tag{1}$$

The source signals $s_i(t)$ and the mixing weights a_{ij} are unknown. The only known values are the observed signals x_{ij}. The problem of BSS is to find the original signals $s_i(t)$ from the observed mixtures $x_i(t)$. The assumption is that there is an invertible matrix A formed from the mixing coefficients a_{ij}. The inference then is that there is a matrix W, with w_{ij} as coefficients, which allows the separation of the s_i, according to

$$s_i(t) = w_{i1}x_1(t) + w_{i2}x_2(t) + w_{i3}x_3(t). \tag{2}$$

That is, $W = A^{-1}$. This is the basic mathematical problem. ICA provides a solution to this seemingly hard problem by assuming that the signals are statistically independent. If two random variables v_1 and v_2 are independent, then, for any non-linear transformations f and g, $f(v_1)$ and $g(v_2)$ will also be uncorrelated [4]. The task of ICA is to find the components such that the components themselves are uncorrelated and also remain uncorrelated under non-linear transformations f and g.

An important principle for estimating independent components is the maximisation of non-Gaussianity. The central limit theorem states that the distribution of the sum of non-Gaussian random variables will be closer to Gaussian than that of the original variables. Finding maxima in non-Gaussianity in a linear combination of the mixture variables $y = \sum_i b_i x_i$ thus gives a means to estimate the independent components [4].

B. ICA in CBIR

Several studies have attempted to use ICA for CBIR. Khaparde et al. compare ICs extracted from a query image with those from images in a database to determine the query results [5]. The paper mentions the use of a ICA filter bank but does not clarify how the filters were designed. It seems as if the process proposed extracts ICs from the query image, uses them as filters and collects filter responses from the database. If this is indeed the case, then there are certain issues with the approach. Firstly, it requires repetitive execution of ICA on the query image. This can be quite an expensive process. Also, ICA can extract a large number of components and it is important to reduce this number for practical CBIR. Sun et al. use ICA in conjunction with Generalized Gaussian Density for the purposes of CBIR [6]. Their results are very encouraging, however their method also suffers from the use of ICA in the image feature extraction process. Bai et al. use Probabilistic ICA to extract image features and uses the z-values of ICs to find a component-wise similarity bipartite [7]. Wang and Dai use ICA features and other low-level features, along with a learning algorithm for image retrieval, and also show very promising results [8]. Although not directly related to CBIR, there has been other work done in applying ICA for image features for a variety of tasks, including segmentation, classification, dimensionality reduction, etc. [9], [10], [11], [6], [12].

C. Our approach

In our work documented in [1] we presented a different use of the ICA in CBIR, compared to the work presented above. The initial inspiration of the work came from the work of Hateren et al. and Bell and Sejnowski . Hateren et al. state that when ICA is applied to images of natural scenes, it produces components similar to the receptive fields in simple cells of the visual cortex [13]. Bell and Sejnowski describe the use of ICA to extract filters from images of natural scenes, and say those filters are edge filters, noting their resemblance to Gabor filters [14].

In the studies mentioned in II-B, ICA has been used to extract image features which were then used to perform CBIR. Our work attempted to use techniques from work such as [13] and [14] and use it in the context of CBIR, essentially using ICA to extract feature extractors (filters), through which we can extract image features. The advantage of this approach is that ICA only needs to be executed once on each image set, at the time of learning the filters. Once the filters have been found, we can extract features from the images using filter responses. This should be faster and more scalable for use in production quality systems.

However, one of the issues with this approach is dealing with the large number of filters that is extracted by ICA. In our work ICA extracted over 200 filters [1]. Obviously using such large number of filters is not practical. So, there needs to be some mechanism to choose a smaller subset of filters. One method is described by Borgne and Guerin-Dugue, where they state that independent component filters have the properties of sparseness and dispersal [2]. However only dispersal is used to select filters in that work. The idea is the variance of a filter's response indicate how useful the filter is for encoding the images. Dispersal is calculated as the normalised variance, where the largest variance is set to 1. So, if V is a list of variances, where V_i is the variance of the responses of filter i . Then the list of dispersal values D is constructed as follows:

$$D_i = \frac{V_i}{max(V)} \tag{3}$$

Where $max()$ is a function which finds the highest value in a list.

For our purpose the actual calculation of dispersal is not required, so from now onwards we will refer to the this work as the variance based method. This method does not solve the problem mentioned in [3], where it is shown that ICA may extract filters which are seemingly shifted/duplicate versions of each other. An example is shown in Figure 1.

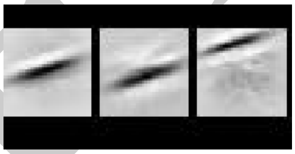


Fig. 1: Filters extracted by ICA which are seemingly shifted/duplicate versions of each other.

By using variance, the method has the weakness of choosing filters with similar characteristics, giving very similar features without providing any extra useful information.

[15] uses cross-correlation with Self Organising Maps (SOMs) to select filters, but provides no motivation for using SOMs for this problem—there is no apparent reason to expect the ICFs to lie on a two-dimensional manifold. More importantly they report no experiments at all.

We used a similar but simpler approach, where the cross-correlation values are used as a distance metric in a complete-link clustering algorithm, followed by filter selection from each cluster [1].

Of the three methods described, only the last one has been applied to a CBIR scenario and the performance of the other methods are unknown. The aim of this study is to compare the performance of the different filter selection approaches, specifically the approach which uses variance measures and the approach which uses clustering. The next section describes the filter selection methods in more detail and outlines the experiments we conducted.

III. Experiments

This section is divided into three parts. The first part describes the method used for filter extraction. Next the different filter selection methods are presented, which is the central part

Fig. 2: 3 images from the VisTex database

of this study. After that we present a short description of the experimental setup.

A. Filter Extraction

The image collection used was a modified version of the VisTex database[1]. It is a database holding a fixed set of images of various kinds of texture. Some example images are shown in Figure 2. The original database had 512×512 images. The version used for this study, uses slightly modified version of the original, which was first used in the development of the Viper/GIFT system at the University of Geneva [16]. Ten 256×256 patches were taken randomly from the images of the database and downsized to 128×128. Using this version allowed the study in [16] to proceed from an established ground-truth, as all the patches taken from a single image can be taken to be similar, specially considering that most images in the VisTex database have a uniform texture throughout.

However, as some of the original images in the VisTex database were very similar, the relevance judgement based on sub-images did not seem to be accurate. To address the problem, this study carried out relevance judgements for 10 sample query images to be used in the experiments.

For filter extraction, one image was chosen to represent each texture class. We implemented the FastICA algorithm [4], and applied it to these selected images. Experimentation was done with various different patch sizes, however this paper only presents the results gathered from 17×17 patches. $10,000$ random patches were extracted from the training set.

As the patch sizes used in this study is 17×17 the original dimension was 289. ICA extracted 231 components, which were re-formed to give 231 filters.

B. Filter selection methods

For this study, we have employed three different techniques to select a smaller subset of filters. They are

- Filter selection through complete-link clustering. CMCLUST
- Filter selection through calculating the variance of the filter energy, where
  - Variance is calculated across all the pixels. (ALLVAR)
  - Variance is calculated for block energies. (BLCKVAR)
- Filter selection through variance after clustering, which is a merge of the above two techniques.
  - CMCLUST followed by ALLVAR. (CMCLUSTALLVAR)
  - CMCLUST followed by BLCKVAR. (CMCLUST-BLCKVAR)

1) Filter selection through clustering (CMCLUST): If ICA extracts N filters, and for each filter f_i where $1 \leq i \leq N$, we calculate the cross-correlation with every other filter. From the cross-correlation matrix of f_i and f_j we choose the highest value as the cross-correlation value of the two filters. Using these values we construct a $N \times N$ matrix D. We then use the matrix D as a distance matrix where D_{ij} is the distance between f_i and f_j. These distances are used in an implementation of complete-link clustering, where filters with a distance less than a given threshold t are grouped together in a single cluster. From each cluster, the filter which has the highest average correlation with other filters in the same cluster is chosen for feature extraction. By varying t we can get different number of clusters, and hence have different number of filters for feature extraction. Figure 3 shows a cluster of filters, grouped by complete-link clustering. Notice how the filters have some similarity in terms of orientation and magnitude.

Our work presented in [1] used one of these methods (CMCLUST) to choose a smaller set of filters.

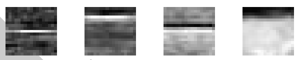


Fig. 3: Complete link clustering grouped these filters in a cluster with t=0.14

Setting $t = 0.2$ gave 19 filters when using CMCLUST.

2) Filter selection through calculating variance of filter energies: The above method does not take into account how filters are responding to the actual image set. The method proposed by [2] has been described in §II-C. As mentioned earlier, this approach does not eliminate shifted/duplicate versions in the filter set. However the filters selected may be more attuned to the image set compared to the previous method.

To extract ICF using ICA we already had a set of training images. Once the filters were extracted, for each candidate filter f_i and each training image I_j we calculate the filter energy by squaring the 2 dimensional convolution matrix $C_{i,j}$. $C_{i,j}$ is calculated as follows:

$$C_{i,j}(x,y)) = \sum_{k_1=1}^{height_{f_i}} \sum_{k_2=1}^{width_{f_i}} I_j(x+k_2, y+k_1) f_i(x,y) \quad (4)$$

The result is a matrix R_{ij} which gives us the filter energy at every pixel. We use these filter energies to select a smaller subset of filters through the following two schemes.

- ALLVAR: The variance is calculated for all the pixel values of the filter energy.
- BLCKVAR: For a filter f_i, its responses over the training images R_{ij}, is divided in 16×16 blocks. The average

[1]Vistex Database is available from http://vismod.media.mit.edu/vismod/imagery/VisionTexture/

filter energy at every block is calculated and the variance of these block energies are calculated for the filter. This closely resembles the operation of the GNU Image Finding Tool (GIFT), which is the CBIR system (CBIRS) we use for the experiments.

For each of the schemes, we choose the top 20 filters with the highest variance.

C. Filter selection through measure of standard deviation after clustering

The third method is a mixture of the two mentioned above. We chose the following thresholds for CMCLUST to give us smaller sets of candidate filters.

TABLE I: Thresholds and number of candidate filters.

Threshold (t)	No. of candidate filters
0.25	37
0.3	68
0.35	106

These candidate filters are then processed through both ALLVAR and BLCKVAR to give us the CMCLUSTALLVAR, CMCLUSTBLCKVAR.

D. Experiments

As our main aim is to use the filters for CBIR, we use CBIR performance as the measure of how well each filter set is performing. In total we have 5 sets of features from the 5 different filter sets. These features were integrated into GIFT. Automated test scripts were used along with the established relevance judgements(also used in [1]). The query results were examined and Precision-Recall graphs were generated to evaluate the efficacy of the filter selection methods.

IV. Results and Discussion

Figure 4 shows the Precision-Recall graphs generated when the same 10 images were used to query the GIFT system for the different feature sets generated by the different filter selection strategies. It also shows the performance of GIFT's own set of Gabor filter based features. The precision values are averaged for sampled values of recall. For this set of images, the features extracted by the CMCLUST filters provide the best results overall.

Figure 5, shows the performance of the features sets extracted using the CMCLUST, CMCLUSTALLVAR and CMCLUSTBLCKVAR filter sets. As pointed out before the CMCLUST filters provide the most useful features. However the filters selected using the variance at a block level seem to provide more useful features compared to filters selected using variance at every pixel value.

This result is also confirmed when compare ALLVAR and BLCKVAR (Figure 6). One explanation for the better results is that BLCKVAR more closely resembles how GIFT uses the filter energies. This makes a case of exploring methods which

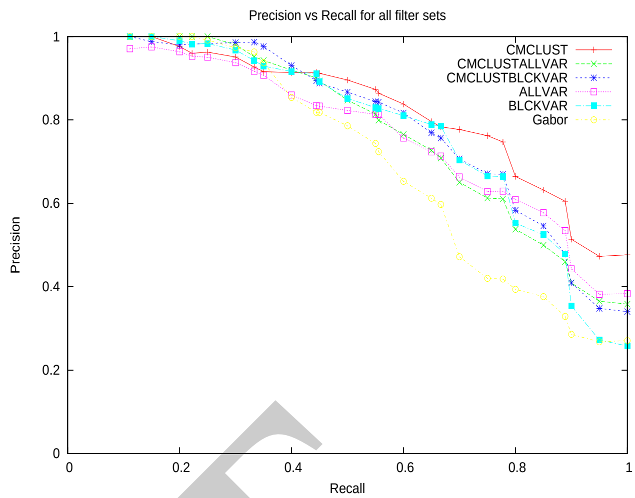


Fig. 4: PR graph depicting the performance of all the filter sets

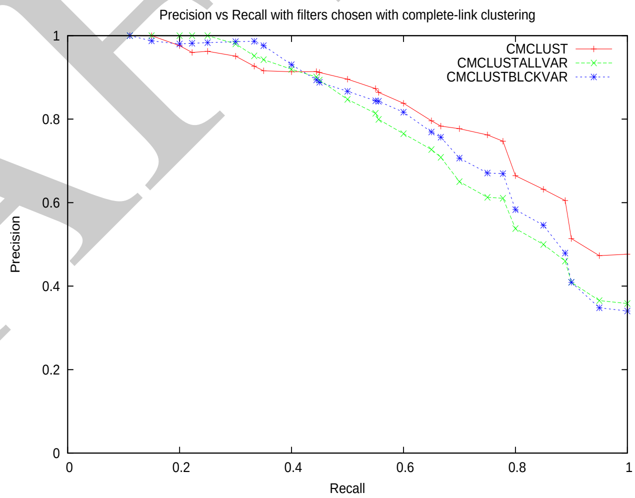


Fig. 5: PR graph depicting the performance the CMCLUST, CMCLUSTALLVAR and CMCLUSTBLCKVAR filter sets.

not only adapts to the image set, but also to the underlying CBIRS.

However the issue with the variance based techniques can be demonstrated with Figure 7. It shows the 20 filters selected by ALLVAR. It is clear that some of the filters are very similar which would provide very similar energy responses and features extracted from them would be very similar. This is further demonstrated by figure 8 and figure 9. Figure 8 shows 5 of the 20 filters chosen using ALLVAR which have a very strong vertical orientation. Figure 9, shows a sample image (9a) being filtered by two of these filters (8a, 8b). As can be seen from Figures 9b and 9c, the energy responses are

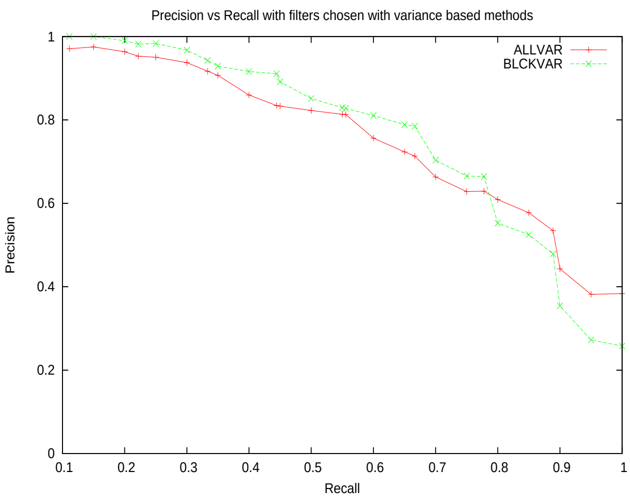


Fig. 6: PR graph depicting the performance the ALLVAR and BLCKVAR filter sets.

very similar for each filter, leading to features which do no provide any extra useful information about the images.

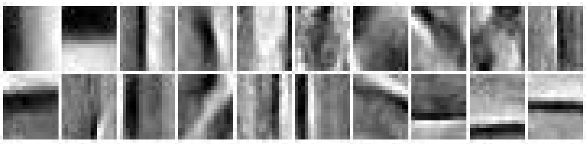
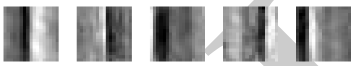


Fig. 7: Filters chosen by ALLVAR.

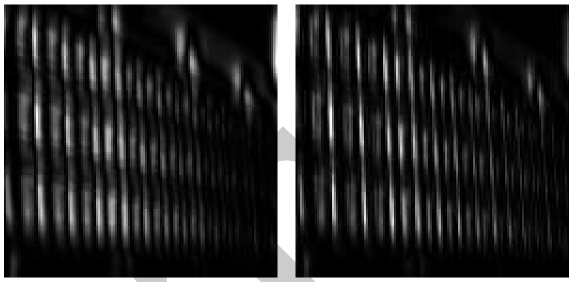


(a) (b) (c) (d) (e)

Fig. 8: Filters from Figure 7 which have vertical orientation

It can be theorised that using ALLVAR or BLCKVAR to choose filters from a smaller set of candidate filters (chosen through clustering) would yield better results. Figure 10 shows the filters which were chosen by CMCLUSTALLVAR. $t = 0.3$ gave 68 candidate filters, from which the top 20 filters where chosen by ALLVAR. It is interesting to note that there are still some shifted/duplicate filters in the selected set, although not as many. The correlation value for the similar looking filters are greater then the threshold, however they were not grouped in the same cluster. This has to do with the implementation of complete-link clustering where filters are clustered together on a first-found basis. This lead to some very similar filters not being grouped together. There is also the case where a filter is the negative version of another. An example of such a case if Figures 8b and 8d. Both the filters are vertical edge detectors, however the in the case of 8b the transition is from black to white. In the case of 8d the transition is from white to black.

(a) Sample image



(b) Energy response from 8a (c) Energy response from 8b

Fig. 9: Filters chosen by ALLVAR giving very similar image features, leading to redundancy

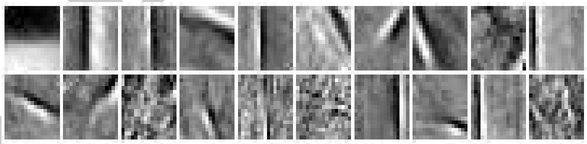


Fig. 10: Filters chosen by CMCLUSTALLVAR with t=0.3

Figure 11 shows that a marginal improvement in CBIR performance can be achieved by clustering for the variance based methods. However the improvement is not universal and can only be seen when the clustering process has discarded more than half of the original filters.

The problem with the variance based approaches can be overcome by using CMCLUST. Figure 12 shows the 19 filters chosen with $t = 0.20$.

The filters chosen through CMCLUST are different from the filters selected through the variance based methods. The filters are more reflective of the underlying textures of the image set unlike the filters chosen by the variance based methods, which are more edge detectors. This is probably why CMCLUST perform best for this data set. One of the causes of this might be inferred from what ICA is meant to do. ICA should extract a set of basis vectors (which we form into filters). A weighted sum of these basis vectors should describe the image patches we used in filter extraction. After clustering, from each cluster we choose the filter with the highest average correlation with other filters in the same cluster. This step is probably what makes this technique choose more effective filters compared to using the responses from the images themselves.

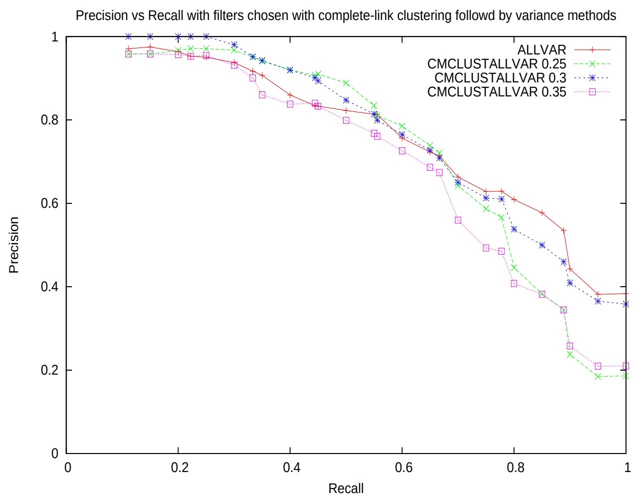


Fig. 11: Performance of CMCLUSTALLVAR filter sets for different t

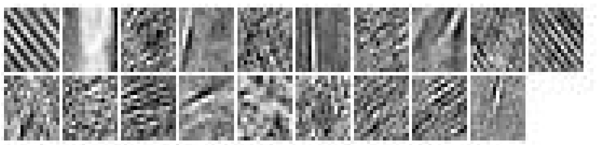


Fig. 12: Filters chosen by CMCLUSTFILTERS with t=0.2

V. Conclusion

We have proposed a new method to select a smaller subset of independent component filters. On our test data, the filters chosen by the CMCLUST extract more useful features than any of the variance based methods. Further work is required to compare across other image collections and across larger data sets. A weakness with the clustering technique used has also been identified. We are pursuing work in trying to use more suitable techniques. It is expected that an improved clustering technique would only enhance the results of CMCLUST, which are already very encouraging.

References

[1] N. Mohammed and D. M. Squire, "Effectiveness of icf features for collection-specific cbir," in *Proceedings of the 9th International Workshop on Adaptive Multimedia Retrieval, co-located with the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, Barcelona, Spain, July 18–19 2011. [Online]. Available: http://www.csse.monash.edu.au/~davids/publications/postscript/2011/AMR2011.pdf

[2] H. L. Borgne and A. Guerin-Dugue, "Sparse-dispersed coding and images discrimination with independent component analysis," in *Proceedings. Third International Conference on Independent Component Analysis and Signal Separation (ICA'2001)*, December 2001.

[3] N. Trojan, "CBIR-based dermatology diagnostic assistant," Honours Thesis, School of Computer Science and Software Engineering, Monash University, Clayton, Victoria, Australia, November 2004, supervised by Dr. David Squire.

[4] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. Wiley-Interscience, May 2001. [Online]. Available: http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/047140540X

[5] A. Khaparde, B.L.Deekshatulu, M.Madhavilatha, Z. Farheen, and V. Sandhya Kumari, "Content based image retrieval using independent component analysis," *IJCSNS International Journal of Computer Science and Network Security*, vol. 8, no. 4, pp. 327–332, 2008.

[6] G. Sun, J. Liu, J. Sun, and S. Ba, "Locally salient feature extraction using ICA for content-based face image retrieval," *Innovative Computing ,Information and Control, International Conference on*, vol. 1, pp. 644–647, 2006.

[7] B. Bai, P. Kantor, A. Shokoufandeh, and D. Silver, "fMRI brain image retrieval based on ICA components," in *Current Trends in Computer Science, 2007. ENC 2007. Eighth Mexican International Conference on*, 24-28 2007, pp. 10 –17.

[8] F. Wang and Q. Dai, "A new multi-view learning algorithm based on ica feature for image retrieval," in *MMM (1)*, 2007, pp. 450–461.

[9] B. Huang, J. Li, and S. Hu, "Texture feature extraction using ICA filters," in *Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on*, 2008, pp. 7631 –7634.

[10] A. Smoaca, D. Coltuc, and V. Lazarescu, "Pattern segmentation in textile images," in *Signals, Circuits and Systems, 2009. ISSCS 2009. International Symposium on*, 2009, pp. 1 –4.

[11] Z. Li, S. Wu, X. Wang, H. Ye, M. Wang, and J. Ye, "Dimensional reduction based on independent component analysis for content based image retrieval," in *Artificial Intelligence, 2009. JCAI '09. International Joint Conference on*, 2009, pp. 741 –745.

[12] R. Zhang, X.-P. Zhang, and L. Guan, "Wavelet-based texture retrieval using independent component analysis," vol. 6, sep. 2007, pp. VI –341 –VI –344.

[13] J. H. V. Hateren and A. V. D. Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex," in *Proceedings. Biological sciences/The Royal Society*, vol. 265(1394), 1998.

[14] A. J. Bell and T. J. Sejnowski, "The "independent components" of natural scenes are edge filters," *Vision Research*, vol. 37, no. 23, pp. 3327–3338, December 1997. [Online]. Available: http://dx.doi.org/10.1016/S0042-6989(97)00121-1

[15] C. Guo and C. Wilson, "Use of self-organizing maps for texture feature selection in content-based image retrieval," in *IEEE International Joint Conference on Neural Networks*, jun. 2008, pp. 765 –770.

[16] D. M. Squire, W. Müller, H. Müller, and J. Raki, "Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback," in *Pattern Recognition Letters*, 1999, pp. 143–149.