Effectiveness of ICF features for collection-specific CBIR

Nabeel Mohammed and David McG. Squire

Faculty Of Information Technology, Monash University. Clayton Campus.
Wellington Road. Victoria 3800, Australia.
{nabeel.mohammed, david.squire}@monash.edu

Abstract. This study aims to find more effective methods for collection-specific CBIR. A lot of work has been done in trying to adapt a system by user feedback, in this study we aim to adapt CBIR systems for specific image collections in an automated manner. Independent Component Analysis (ICA), a high order statistical technique, is used to extract Independent Component Filters (ICF) from image sets. As these filters are adapted to the data, the hypothesis is that they may provide features which are more effective for collection-specific CBIR. To test this question, this study develops a methodology to extract ICF from image sets and use them to extract filter responses. In developing this method, the study uses image cross-correlation and clustering to solve issues to do with shifted/duplicate filters and selecting a smaller set of filters to make CBIR practical. The method is used to generate filter responses for the VisTex database . The filter response energies are used as features in the GNU Image Finding Tool (GIFT). The experiments show that features extracted using ICF have the potential to improve the effectiveness of collection-specific CBIR, although some more work in this area is required.

1 Introduction

The primary aim of this study is to establish more effective techniques for Content Based Image Retrieval (CBIR) in collection-specific cases. With the wide availability of computing resources, a very large number of images are being produced, used and stored. However image searching has still not become nearly as effective or useful as text search [18]. This has to do with the various difficulties in searching for images, which can based on either textual queries or image queries [16]. The latter is of interest to us in this study, as searching for images by using images as queries is the essence of CBIR.

While there have been successful results in using generic approaches (at times customised) for CBIR, the effectiveness of these systems can still be improved [15]. The current research laboratory prototypes of CBIR systems are still far behind from being available as effective commercial products [14]. As described in [16] it might be impossible to create a generic CBIR system (CBIRS) which performs well in all cases. A lot of research has taken place to improve CBIR performance based on user feedback. In this study we aim to explore a method to perform collection-specific CBIR, adapted for the collections, as it may be an area where great improvements are achievable.

Independent Component Analysis (ICA), a high-order statistical technique, has been used with success to extract independent component filters (ICF) from images. These filters are are extracted in an unsupervised manner and are adapted to the images [12]. The research question is whether texture features extracted using these filters are more effective in CBIR compared to generic texture features. While there has been studies where ICA have been used to directly extract image features, to the knowledge of the authors no study has been carried out to establish whether the ICF extracted using ICA would lead to better features for CBIR. This study aims to answer this question. If the texture features extracted using ICF are shown to be more effective compared to generic approaches then this study would contribute new techniques to improve the effectiveness of collection-specific CBIR.

2 Background

2.1 Independent Component Analysis (ICA)

ICA is defined as "a method for finding underlying factors or components from multivariate (multidimensional) statistical data" [9]. 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 [9]. In fact BSS is a good example to describe 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 $i = 1, 2, 3$

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

In this situation, 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, there is a matrix W, with w_{ij} as coefficients, which would allow the separation of each s_i, as

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

That is, $W = A^{-1}$. This is the basic mathematical problem. ICA provides a solution to this seemingly hard problem by the assuming that the signals are statistically independent. That is, if 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 [9]. So, in essence 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 sum of non-Gaussian random variables will be closer to a Gaussian compared to the original ones. Therefore finding maxima in non-Gaussianity in a linear combination of the mixture variables ($y = \sum_i b_i x_i$) gives us the independent components [9].

2.2 ICA in CBIR

As ICA finds underlying components from a dataset and has been applied to images, various studies have attempted to use it for CBIR. [11] compares ICs extracted from the query image and images in the database to determine the results. 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. [19] also uses ICA in conjunction with Generalized Gaussian Density for the purposes of CBIR. The results shown in the paper are very encouraging, however their method also suffers from the use of ICA in the image feature extraction process. [1] uses Probabilistic ICA to extract image features and uses the z-values of ICs to find a component-wise similarity bipartite. [21] uses ICA features and other low-level features, along with a learning algorithm for image retrieval and they show very promising results. Although not directly related to CBIR, there has been some work done in applying ICA for image features for a variety of tasks, including segmentation, classification, dimensionality reduction, etc. [8] [17] [13] [19] [22].

2.3 Our approach

In image processing, ICA can be used to extract components from sets of images. These components can be transformed back into patches to form filters. Each of the filters is known as an Independent Component Filter (ICF) [6]. Hateren et al. [6] state that when ICA is applied to images of natural scenes, it produces components similar to the receptive fields in simple cells in the visual cortex. [2] 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.

In the studies mentioned in 2.2, ICA has been used to extract image features which were then used to perform CBIR. This study attempts to use techniques from work such as [6] and [2] and use it in the context of CBIR. This differentiates our work from that of [11], [19] etc. 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.

3 Experiments

This study used 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 1. The original database had 512×512 images. The version used for this study,

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

uses slightly modified version of the original, which was first used in the development of the Viper/GIFT system at the University of Geneva [18]. 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 [18] 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, as can be seen from Figure 1.

Fig. 1. 3 images from the VisTex database

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 all the images in the database, three sets of features were extracted. They are explained here.

3.1 Feature sets

Existing GIFT features using a bank of Gabor filters A complete explanation of the feature types of GIFT can be found in [18]. Here we will present information related only to the texture features and only the information deemed to be most relevant.

GIFT employs a bank of real, circularly symmetric Gabors:

$$f_{mn}(x, y) = \frac{1}{2\pi\sigma_m^2} e^{-\frac{x^2+y^2}{2\sigma_m^2}} \cos(2\pi(u_{0_m} x \cos\theta_n + u_{0_m} y \sin\theta_n)), \tag{3}$$

where m indexes filter scales, n their orientations, and u_{0_m} is the centre frequency. The half peak radial bandwidth is chosen to be one octave, which determines σ_m. The highest centre frequency is chosen as $u_{0_1} = 0.5$, and $u_{0_{m+1}} = u_{0_m}/2$. Three scales are used. The four orientations are: $\theta_0 = 0$, $\theta_{n+1} = \theta_n + \pi/4$. The resultant bank of 12 filters gives good coverage of the frequency domain, with little filter overlap. The mean energy of each filter is computed for each of the smallest blocks in the image, and quantized into 10 bands. A feature is stored for each filter with energy greater than the lowest band. These are treated as local features. Each image has at most 3072 of the 27648 such possible features. Histograms of these features are used to represent global texture characteristics.

ICF features For filter extraction, one image was chosen to represent each texture class. We implemented the FastICA algorithm [9], 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. For some of the experiments, a 17×17 Hamming window was applied on each patch. However a normal $2D$ Hamming window has a sharp rise and a narrow peak, leading to massive reduction in dimensionality when using PCA. In an attempt to get a slightly wider peak, the window used in this study is an elementwise square root of the original Hamming window. The use of a Hamming window has been recommended in literature [3], but there is no evidence showing whether ICFs extracted from image patches which had a Hamming window applied to it, were more effective than ICFs which were extracted from unprocessed image patches. Although it is not the main goal of this study to perform such a comparison, this study will provide some insight about the usefulness of Hamming windows for patch processing. For the purpose of this paper, the features extracted using ICFs with the Hamming window shall be called HammingICF features and the features extracted using ICFs without the Hamming window shall be called the ICF features.

As the patch sizes used in this study is 17×17 the original dimension was 289. PCA was used to drop the least significant 1% of the dimensions before applying ICA. Table 1 shows the dimensionality after PCA and the number of ICFs extracted.

Table 1. New dimensions after running PCA and number of ICs

Hamming Window	New dimensions	No. of IC
Yes	180	110
No	231	146

[20] mentions the problem where ICA extracts ICFs which are apparently shifted or near duplicate versions of each other. As these filters would extract very similar features, it is preferable to eradicate the shifted/duplicate versions. Also, ICA extracts very large number of ICFs. For practical CBIR, we need to select a smaller subset of filters. [4] 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.

This study has employed a similar but simpler approach. For each pair of filters a cross-correlation matrix was calculated. From each of these matrices the highest value was chosen as the cross-correlation value of the two filters. Using these values a matrix of correlation values of each pair of filters was constructed. This matrix was used as a similarity measure in an implementation of complete-link clustering [10]. To find filters to use in CBIR, a decision was made on how many filters we want to use for feature extraction. The threshold value given to the clustering algorithm was adjusted appropriately to return the desired number of filters. As an example, when about 20 filters were required, a threshold value of 0.14 was used giving 22 filters. This is one crucial area of this study and needs further research to establish the validity of this technique. However a visual inspection was conducted and the clustering technique seemed to be grouping filters with similar visual layout. An example of a cluster is shown in Figure 2. The filters have been scaled for visual presentation.

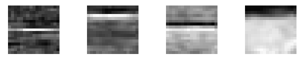


Fig. 2. Complete link clustering grouped these filters in a cluster when a threshold of 0.14 was used.

From each cluster, the filter which has the highest average correlation with other filters in the same cluster is chosen. The chosen filters were used in $2D$ convolution and GIFT used the filter energies to extract image features using the approach mentioned for the bank of Gabor filters.

GLCM based features As both the Gabor features and the ICF features are extracted using filters, we wanted to compare their performance with other types of texture features. One of the most common and intuitive method for texture feature extraction is using Grey Level Co-occurrence matrices. [5] suggested the use of GLCMs and provided a good set of features which can be extracted from the GLCMs of images. A GLCM conveys information about the frequency of two grey level values, separated by a certain vector, appearing in an image. Changing the angle and size of the vector will give different GLCMs. There have been many different methods of applying GLCMs, but we chose to implement the method used by [7]. The work in [7] was carried out for medical images, which could be classified as a specialised collection.

The study in [7] divided each images in 7×7 tiles and for each tile calculated 16 GLCMs. The GLCMS were generated for vectors of size 1,2,3 and 4 pixels and orientations 0, $\frac{\pi}{4}$, $\frac{\pi}{2}$ and $\frac{3\pi}{4}$. For each GLCM $P(i,j)$, they calculated a homogeneity feature H_p,

$$H_p = \sum_i \sum_j \frac{P(i,j)}{1 + |i-j|} \tag{4}$$

Using these features they calculated the Manhattan distance between the query images and the images in the database.

For our work, we implemented their scheme and used it as local GLCM features. For global GLCM features, we calculated GLCMs for whole images, rather than breaking it up into tiles. When combining global and local features, we followed the recommendation in [7] and divided distances of each feature by their median and then summing up the values.

3.2 Setup and feature sets

GIFT as a CBIRS has a unique property that it uses MRML, an XML-based communication protocol, to communicate between the CBIR client and the CBIRS and also between its own components[2]. This allowed the development of a suite of scripts which can be used to automatically run tests and gather the results.

For the tests, 10 query images were selected from the image database. The existing features, extracted using a bank of Gabor filters, performed adequately on some of these images but not very well on most of them. Example of such images are shown in figure 3.

(a) (b)

Fig. 3. Two example query images.The Gabor features perform well on (a) and poorly on (b)

Precision-Recall graphs were generated based on the relevance judgements and used to compare CBIR performance for features extracted using three different filter collections and the GLCMs, giving the following feature sets.

- Features extracted using a bank of 12 Gabor filters (Gabor features).
- Features extracted using ICF without Hamming window (ICF features).
- Features extracted using ICF with Hamming window (HammingICF features).
- Features extracted using GLCM features (GLCM Features).

Experiments were conducted to find the CBIR performance for global features, local features and using both global and local features and the results are described next.

4 Results and Discussion

Figures 4, 5 and 6 shows the Precision vs Recall graphs generated from the experiment results for the VisTex database, where precision is average over all query images. As can be seen from the graphs, the ICF features have overall better performance compared to the Gabor features. The HammingICF features perform better than the Gabor features, however not as well as the ICF features. This is a bit surprising, as it was theorised that eliminating hard edges in the filters would produce more useful results, but the use of the Hamming window seems hamper the extraction of useful image features. When using local features only, the performance of the GLCM features is clearly superior to the filter based features, although not close to the global features.

For a query image, the GIFT interface shows 20 images that GIFT deems to be most relevant depending on the features used. For each query image, the precision in the first 20 results were calculated. These results are shown in Figures 7, 8 and 9. It is clear from

[2] `http://www.mrml.net/`

these graphs that for certain image classes the Gabor features give good results, but for a majority of the query images the ICF feature and the HammingICF features outperform the Gabor and GLCM features . This measurement of performance is important as it can be directly translated to a better user experience.

As is evident from the results, the global features outperform the local features. Using the combination of global and local features actually seems to worsen performance. This can be explained by the nature of the images in the VisTex database. As mentioned previously, most of the images have a uniform texture; it is reasonable that features which express image characteristics as whole would work better. Local feature which would identify differences in image regions might actually be counter-productive for a database of images such as VisTex. The only exception to this seems to be the performance of the GLCM features when querying with images of buildings. The combination of local and global GLCM features gives perfect result. This needs further study to ascertain whether for certain types of images the GLCM features are able to capture more useful information.

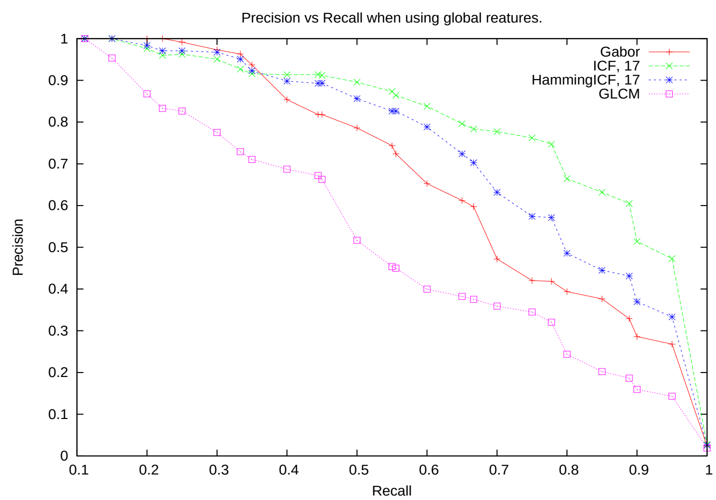


Fig. 4. Precision vs Recall graph for queries using global features.

It was theorised that ICF extracted using ICA would be adapted to the data and hence would help in finding texture features to which the Gabor filters and other predetermined methods would be blind. An example of such a case is shown in the two query images in Figures 10(a) and 10(b). Both of them are images of very similar fabrics. Figure 10(a) shows a fabric which has larger texture compared to the fabric shown in Figure 10(b). When using Gabor features, the first query image (Figure 10(a)) has

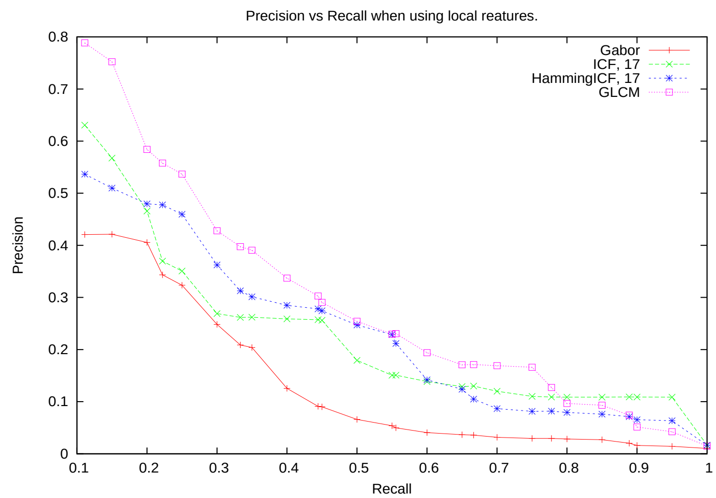


Fig. 5. Precision vs Recall graph for queries using local features.

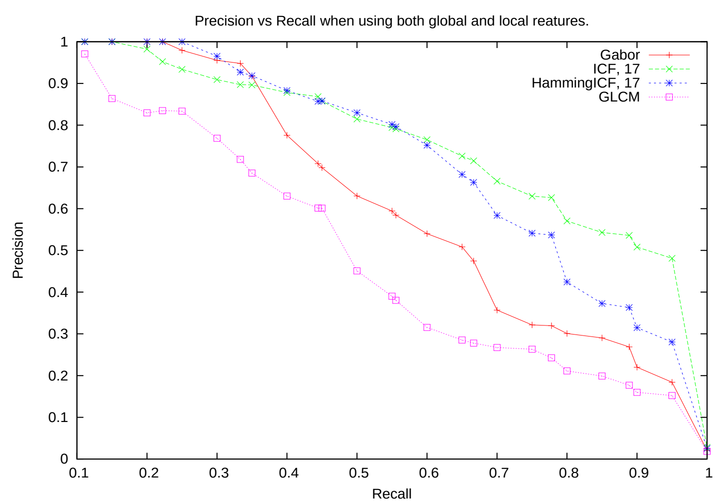


Fig. 6. Precision vs Recall graph for queries using both global and local features.

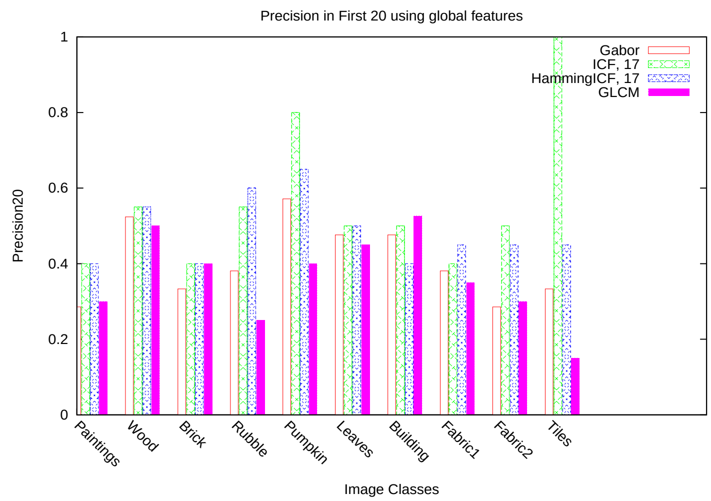


Fig. 7. Precision for different image classes in the top 20 results, using Global features only.

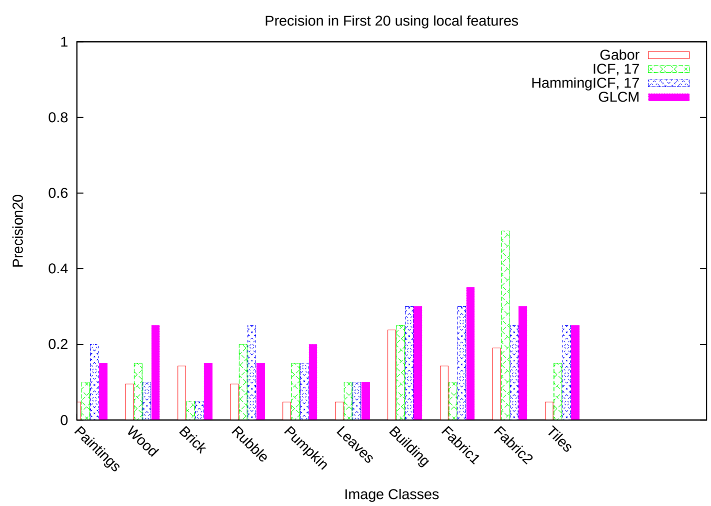


Fig. 8. Precision for different image classes in the top 20 results, using Local features only.

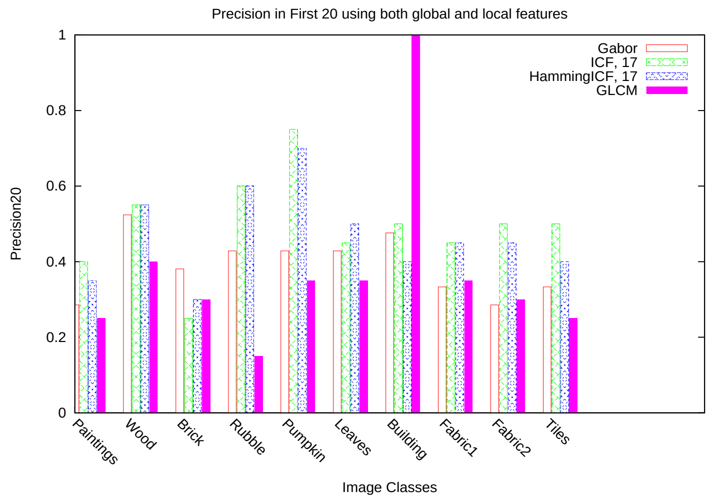
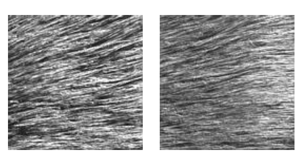


Fig. 9. Precision for different image classes in the top 20 results, using both local and global features.

a precision value of 0.4 in the first 10 images retrieved compared to a precision value of 0.5 for the GLCM features and 1 for the ICF-based features. On the other hand for Figure 10(b) the Gabor, GLCM and ICF-based features, all have a precision value of 0.7 for the first 10 images retrieved. It is also interesting to note that for Figure 10(a) when using Gabor features, images of water are judged to be similar and when using GLCM features, images of powdered food is judged to be similar. This is an example where the pre-determined size and orientation of the bank of Gabor filters and GLCMs leads to inaccurate results.



(a) (b)

Fig. 10. Two very similar query images. For (a) the Gabor features perform poorly, for (b) they perform well.

5 Conclusion

The results indicate that for images with globally consistent texture features, the ICF features work well. Analysing the precision at the top 20 results, the global ICF-based features seem to perform better than both the Gabor and GLCM features. However for a single image (buildings), the global GLCM features perform better. The ICF-based features seem to describe image characteristics which the Gabor and GLCM features have been unable to capture. This is a strong indication that learning filters from image sets and using those filters to extract features can be a viable way to perform CBIR feature extraction, without incurring the cost of having the overhead of running ICA repeatedly.

Extracting the ICF is more computationally expensive than using Gabor filters. Further studies need to explore how this process can be made less expensive. There are other areas which may give ICF features better performance, crucial among those are choosing the filters to use in CBIR. This study uses cross-correlation followed by clustering, however other methods need to be explored to ensure that the optimal set of filters is chosen.

It is also acknowledged that this study has the same limitations as other CBIR studies. The results are based on the image database chosen and also the relevance judgements made. Work is currently under progress to implement similar processes across different image collections, including a set of images with varying texture within the image and also one of skin lesions. Also, an effort is being made to establish a comprehensive relevance judgement set for the VisTex database and the other databases. The results of this study gives encouragement to pursue this further work.

References

1. Bai, B., Kantor, P., Shokoufandeh, A., Silver, D.: fMRI brain image retrieval based on ICA components. In: Current Trends in Computer Science, 2007. ENC 2007. Eighth Mexican International Conference on. pp. 10 –17 (24-28 2007)
2. Bell, A.J., Sejnowski, T.J.: The “independent components” of natural scenes are edge filters. Vision Research 37(23), 3327–3338 (December 1997), http://dx.doi.org/10.1016/S0042-6989(97)00121-1
3. Borgne, H.L., Guerin-Dugue, A.: 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)
4. Guo, C., Wilson, C.: Use of self-organizing maps for texture feature selection in content-based image retrieval. In: IEEE International Joint Conference on Neural Networks. pp. 765 –770 (jun 2008)
5. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. Systems, Man and Cybernetics, IEEE Transactions on 3(6), 610–621 (1973), http://dx.doi.org/10.1109/TSMC.1973.4309314
6. Hateren, J.H.V., Schaaf, A.V.D.: 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)
7. Howarth, P., Yavlinsky, A., Heesch, D., Rger, S.: Medical image retrieval using texture, locality and colour. In: Peters, C., Clough, P., Gonzalo, J., Jones, G., Kluck, M.,

Magnini, B. (eds.) Multilingual Information Access for Text, Speech and Images, Lecture Notes in Computer Science, vol. 3491, pp. 918–918. Springer Berlin Heidelberg (2005), http://dx.doi.org/10.1007/1151964572, 10.1007/1151964572

8. Huang, B., Li, J., Hu, S.: Texture feature extraction using ICA filters. In: Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on. pp. 7631 –7634 (2008)
9. Hyvarinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley-Interscience (May 2001), http://www.amazon.ca/exec/obidos/redirect?tag=citeulike09-20&path=ASIN/047140540X
10. Jain, A.K., Murty, M.N., Flynn, P.J.: Data clustering: a review. ACM Comput. Surv. 31(3), 264–323 (September 1999), http://dx.doi.org/10.1145/331499.331504
11. Khaparde, A., B.L.Deekshatulu, M.Madhavilatha, Farheen, Z., Sandhya Kumari, V.: Content based image retrieval using independent component analysis. IJCSNS International Journal of Computer Science and Network Security 8(4), 327–332 (2008)
12. Le Borgne, H., Guérin-Dugué, A., Antoniadis, A.: Representation of images for classification with independent features. Pattern Recognition Letters 25(2), 141–154 (January 2004), http://dx.doi.org/10.1016/j.patrec.2003.09.011
13. Li, Z., Wu, S., Wang, X., Ye, H., Wang, M., Ye, J.: Dimensional reduction based on independent component analysis for content based image retrieval. In: Artificial Intelligence, 2009. JCAI '09. International Joint Conference on. pp. 741 –745 (2009)
14. Marques, O., Furht, B.: Content-Based Image and Video Retrieval. Kluwer Academic Publishers, Norwell, MA, USA (2002)
15. Müller, H., Rosset, A., Vallee, J., Geissbuhler, A.: Comparing feature sets for content based image retrieval in a medical case database. In: In Proceedings of the Medical Informatics Europe Conference (MIE2003). St. Malo, France (May 2003)
16. Müller, W.: Design and implementation of a flexible Content–Based Image Retrieval Framework - The GNU Image Finding Tool. Ph.D. thesis, Computer Vision and Multimedia Laboratory, University of Geneva, Geneva, Switzerland (September 2001)
17. Smoaca, A., Coltuc, D., Lazarescu, V.: Pattern segmentation in textile images. In: Signals, Circuits and Systems, 2009. ISSCS 2009. International Symposium on. pp. 1 –4 (2009)
18. Squire, D.M., Müller, W., Müller, H., Raki, J.: Content-based query of image databases, inspirations from text retrieval: inverted files, frequency-based weights and relevance feedback. In: Pattern Recognition Letters. pp. 143–149 (1999)
19. Sun, G., Liu, J., Sun, J., Ba, S.: Locally salient feature extraction using ICA for content-based face image retrieval. Innovative Computing ,Information and Control, International Conference on 1, 644–647 (2006)
20. Trojan, N.: 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
21. Wang, F., Dai, Q.: A new multi-view learning algorithm based on ica feature for image retrieval. In: MMM (1). pp. 450–461 (2007)
22. Zhang, R., Zhang, X.P., Guan, L.: Wavelet-based texture retrieval using independent component analysis. vol. 6, pp. VI –341 –VI –344 (sep 2007)