Recognition of Logo Images Using Invariants Defined from Higher-Order Spectra

S. Lowther, V. Chandran and S. Sridharan
Research Concentration in Speech, Audio and Video Technology
Queensland University of Technology, GPO Box 2434, Brisbane QLD 4001, Australia
s.lowther@qut.edu.au, v.chandran@qut.edu.au, s.sridharan@qut.edu.au

Abstract

This paper presents a system for the automatic recognition of logo images extracted from electronic document images. It uses feature extraction by use of higher-order spectra and classification by means of a Nearest Neighbour classifier. Results demonstrate the robustness of the technique for logo recognition where the images suffer from imaging defects, varying degrees of Gaussian, speckle and salt and pepper noise as well as slight variations in logo structure and inclusion of background artefacts. The technique has been applied to two independent databases and achieved recognition accuracies as high as 99.6%.

1. Introduction

Logos are used by organisations to identify themselves on documents. When scanned paper documents are analysed to produce structural electronic representations or for the purpose of sorting by corporate identity, logo recognition becomes an important component. Successful recognition of logos facilitates automatic source classification of document images and may also be used to determine how best to process the information contained within these particular documents.

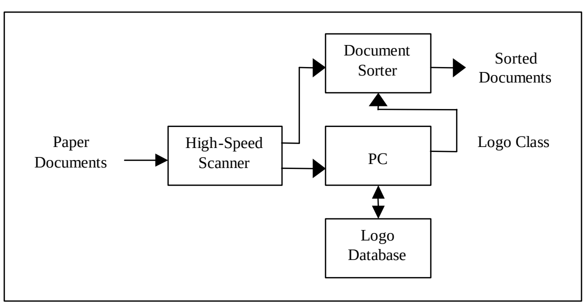

Figure 1. Example application where logo recognition facilitates the automatic sorting of paper documents according to corporate identity

An example of an application of logo recognition is presented in Figure 1. In this application, paper documents are fed into a high-speed scanner. The digitised document images are then passed on to a computer where logos are extracted and matched against a logo database. The results of the logo recognition are used to sort the documents from the scanner according to the company logo on each. Such an application is useful wherever a large collection of business documents requires sorting into organisational classes.

There are a number of challenges associated with logo recognition. One such challenge includes robustness to noise since many digitised documents containing logos have some degree of noise. A primary source of noise is due to the printing, scanning and faxing of paper documents. Another source of noise is introduced from environmental factors such as smudging of ink, damaged documents and dirty documents. Another challenge with logo recognition is accurate logo segmentation. A logo image may not be segmented accurately due to the logo detection and segmentation process and as such, a perfectly skew and scale normalised logo image may not be available. These extracted logo images may also contain surrounding text and background artefacts from the original document. The type of logos available also varies quite considerably with types such as line style logos, dense graphical logos, textual logos and a combination of these. With these challenges in mind, a logo recognition algorithm must be robust to noise, slight skew angles, scale and logo structure type. Figure 2 presents a few example images depicting some of these challenges.

While there has been considerable work in the general area of image recognition, few researches have developed algorithms designed specifically for recognition of logos. Some of the techniques developed for the specific task of logo recognition include neural networks [1, 2] and a multi-level approach using text, contour features and similarity invariants [3].

The neural-network approaches [1, 2] have proven to be very effective in discriminating between binary logo images. While most logo images will contain a single

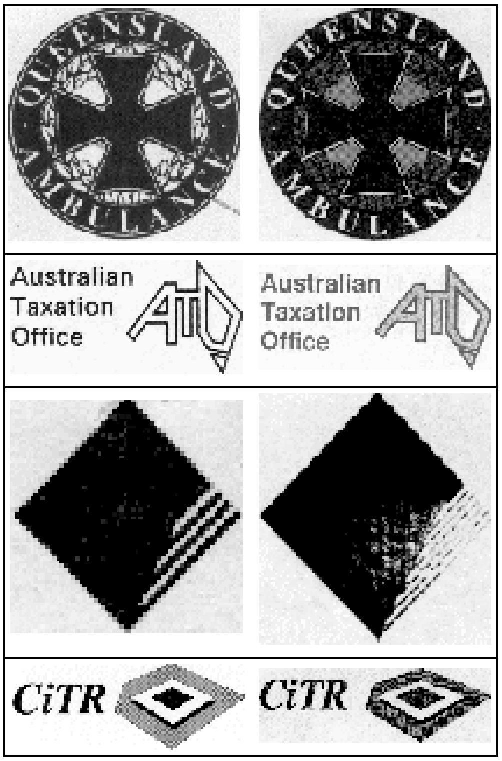

Figure 2. Some examples illustrating variations between instances of logos from the same logo class

intensity, some logo images may consist of varying shades. If a binary threshold is applied to these logos, the resulting logo image may appear very different from another logo image from the same class. An example of this effect is depicted in Figure 3. Neural network techniques also generally require considerable time for training (for [1, 2] between 47 and 137.56 hours). It is

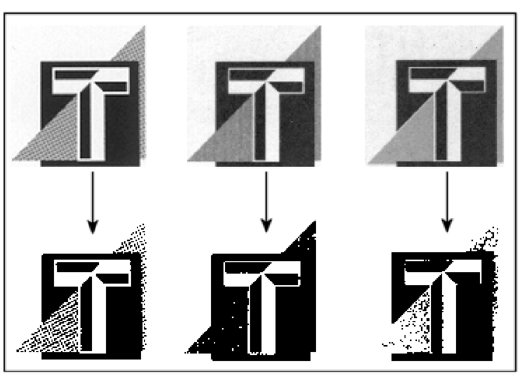

Figure 3. Example of applying a binary threshold to greyscale logo images from the same logo class

also unclear how robust the above techniques are to variations in logo style and inclusion of surrounding text or background artefacts extracted from the original document.

2. Higher-Order Spectral Invariant Features

The technique we propose is based on previous work in the area of higher-order spectra [4, 5]. As shown in [5], higher-order spectra can be used effectively for the task of image recognition by extracting a series of 1-D projections of an image in angle increments and extracting features from the deterministic bispectrum, a triple product of Fourier coefficients, of these projections. The technique is robust to the presence of Gaussian noise by virtue of the integration in the bispectral plane used to calculate the features and also by the use of 1-D projections of the input image. It is also robust to translation and scale and can also be made robust to rotation at the expense of reducing discrimination power. The robustness offered by the technique is highly suited to the task of logo recognition, allowing robust features to be extracted for most logo classes. An explanation of how to calculate bispectral features for 2-D images is given below.

Let g(x,y) be an NxN image containing a logo. Let $P_\theta(m)$ be the Radon transform projection at angle θ. The projections are zero padded to length 2N. Let R(k) be the DFT of $P_\theta(m)$. Then the triple product,

$$B(k_1,k_2) = R(k_1)R(k_2)R^*(k_1 + k_2) \tag{1}$$

is referred to as the deterministic bispectrum of R(k). If the bispectrum is radially integrated,

$$I(a) = \int_{k_1=0+}^{1/(1+a)} B(k_1, ak_1)dk_1 \tag{2}$$

the features $\Phi(a)$ = arg[I(a)]. The integrated bispectrum can be obtained by interpolating in the discrete bifrequency plane or by an indirect method using DTFTs [6]. These features, $\Phi(a)$, have been mathematically proven to be translation- and scale-invariant [4]. Using a number of projection angles, we can extract rotation invariant features, also at the expense of discrimination power.

Because expected values of the triple product in equation (2) are zero for a variety of random processes including Gaussian, these features are robust to noise. Background artefacts such as a line of text, typically alter the DC level in a projection along the y-axis and this does not affect the HOS features from the projection appreciably either because DC values are ignored.

Features are extracted from the phase of the bispectrum, since the phase information contains important shape information of the original image. The basic algorithm for calculating features for recognition based on this technique is given in the next section. For a more detailed understanding of how bispectral features are calculated, see [4, 5].

3. Algorithm

This section presents the basic algorithm for the extraction of bispectral features for logo recognition. Classification is performed using a simple Nearest Neighbour classifier. Results for the algorithm are presented in the next section.

The basic algorithm for logo recognition is as follows:

1. Transform the logo image to greyscale in the range [0,255]
2. Resize the logo image to fit within a N x N image block (while maintaining the logo's aspect ratio)
3. Apply a smoothing filter to the logo image
4. Take a series of angle increments between 0 and 180 degrees
5. For each angle increment
    a. Take the 1-D projection at the specified angle
    b. Find M bispectral features
    c. Concatenate with the previous bispectral features obtained for the other angle increments
6. Perform classification using a Nearest Neighbour classifier.

Our implementation of the above algorithm uses N=64 with 6 angles in increments of 30 degrees and 2 bispectral phase features extracted for each angle.

4. Results

Testing of the proposed logo recognition technique has been performed on both the University of Maryland logo database [7] and our own logo database. Multiple degraded copies of each logo were generated according to Baird's Document Image Defect Model [8] using deformations for Blur (mean, m = 0.7 and standard error, e = 0.3), Sensitivity (m = 0.06, e = 0.02), Skew (m = 0, e = 0.5), Width (m = 1, e = 0.05), Height (m = 1, e = 0.05) and Kerning in the X and Y axes (uniform over [0, 1]). Resolution deformation was modelled as re-sampling uniformly over the range of resolutions {300dpi, 250dpi, 200dpi, 150dpi, 100dpi, 75dpi} using the assumption that original images were scanned at 300dpi.

The University of Maryland logo database provides only a single instance of 105 individual logo classes. 60 artificially degraded logo images were generated from each image in the database, giving a total of 6300 images

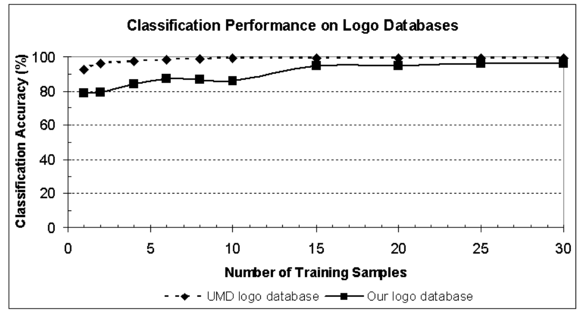

Figure 4. Classification accuracies obtained from testing our algorithm on the University of Maryland logo database (105 classes) and our own logo database (26 classes)

(3150 training and 3150 testing images). Testing was conducted on training sets of different size as seen in Figure 4. Solid recognition performance (above 90%) is achieved even when only a single training sample per class is used. The highest accuracy obtained was 99.6% using 30 training samples.

Our own logo database was created to address the lack of multiple original instances of individual logo classes in the University of Maryland logo database. Our database provides between 3 and 14 instances of 26 individual logo classes, allowing evaluation of logo recognition across independent training and testing sets. For our tests 3 instances of each logo class were used to generate 30 degraded training images per logo class. The remaining 71 images were used to generate a total of 1420 degraded testing images. Testing was again conducted on training sets of different size as seen in Figure 5. On this database, high recognition performance (95%) is attained when 15 training samples per class are used. The highest accuracy obtained was 96.3% using 30 training samples.

Testing of the algorithm for robustness to noise has also

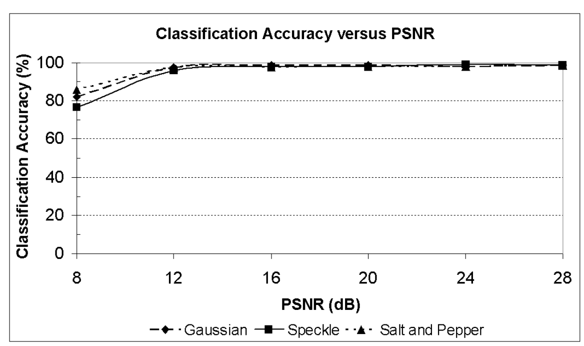

Figure 5. Results of testing the preposed algorithm over a range of PSNR levels of the testing set

been evaluated on our logo database using the above training set with 30 training samples per logo class. Test images were generated by corrupting the original 71 test images with Gaussian, speckle and salt and pepper noise over the range {8, 12, 16, 20, 24 and 28dB}, creating a test set of 1278 images. As can be seen in Figure 5, high recognition performance (above 96%) is achieved for all PSNR values of 12dB or greater.

5. Discussion

From the results presented in the previous section, it can be seen that the proposed method is well suited to the task of logo recognition. The features extracted show significant robustness to noise and even a simple classifier can be used to achieve a high degree of accuracy. The features used for classification have also shown robustness to variations in logo structure between instances from the same class. Recognition accuracies have been as high as 99.6% on the University of Maryland logo database and 96.3% on our own logo database. Learning time for new logo classes is in the order of seconds and recognition of a single logo image can be performed in less than 1 second.

While the technique has not been shown to be robust for significant skew angles, due to the nature of the intended application, techniques such as those proposed in [9, 10] can be used to correct the skew of documents before logo segmentation and recognition is performed.

Further research is now planned to improve recognition accuracy such as incorporation of complementary features and improved classification models. Some of the complementary features that will be investigated include statistical details of the logo images such as pixel density, aspect ratio and connected component information. Another avenue of future research is to develop accurate methods of detection and extraction of candidate logo regions from documents. Logo detection is currently being investigated and it is envisaged that a complete system will incorporate a highly coupled detection and recognition algorithm.

6. Conclusion

We have presented a new technique for automatic recognition of logo images. The technique is based on classification of features extracted from the phase of the bispectrum of a logo image. Results have shown that even a simple Nearest Neighbour classifier can be used to achieve a high degree of recognition accuracy. The use of bispectral features has ensured that the technique is robust against noise and slight logo structure variations.

7. Acknowledgements

The authors would like to thank the Defence Science and Technology Organisation (DSTO) for their support in our document analysis research.

8. References

[1] E. Francesconi, P. Frasconi, M. Gori, S. Marinai, J. Q. Sheng, G. Soda and A. Sperduti, “Logo Recognition by Recursive Neural Networks”, Proceedings of the 2nd IAPR Workshop on Graphics Recognition, pp. 144—151, Springer Verlag, 1997.

[2] F. Cesarini, E. Francesconi, M. Gori, S. Marinai, J. Q. Sheng and G. Soda, “A Neural-based Architecture for Spot-noisy Logo Recognition”, Proceedings of the 4th International Conference Document Analysis and Recognition, pp. 175—179, August 1997.

[3] D. Doermann, E. Rivlin and I. Weiss, “Logo Recognition”, UMD Technical Report CS-TR-3145, University of Maryland, 1993.

[4] V. Chandran and S.L. Elgar, “Pattern Recognition Using Invariants Defined from Higher Order Spectra—One-Dimensional Inputs”, IEEE Transactions on Signal Processing, Vol. 41, No. 1, pp. 205—212, January 1993.

[5] V. Chandran, et. al., “Pattern Recognition Using Invariants Defined from Higher Order Spectra: 2-D Image Inputs”, IEEE Transactions on Signal Processing, Vol. 6, No. 5, pp. 703—712, May 1997.

[6] C. L. Nikias and A. P. Petropulu, “Higher-Order Spectra Analysis : A Nonlinear Signal Processing Framework”, Englewood Cliffs, N.J., Prentice Hall, 1993.

[7] D. Doermann, University of Maryland, Document Processing Group Logo Database, http://documents.cfar.umd.edu/resources/database/UMDlogo.html

[8] H. S. Baird, “Document Image Defect Models”, Structured Document Image Analysis, pp. 546—556, Springer Verlag, 1992.

[9] H. S. Baird, "The Skew Angle of Printed Documents," in *Document Image Analysis*, L. O'Gorman and R. Kasturi, Eds.: IEEE Computer Society Press, pp. 204—208, 1995.

[10] G. S. Peake and T. N. Tan, "A General Algorithm for Document Skew Angle Extimation," presented at Proceedings of the 1997 International Conference on Image Processing, pp. 230—233, 1997.