An attention based similarity measure for fingerprint retrieval

Fred W M Stentiford*
UCL@Adastral Park, Martlesham, Ipswich, Suffolk, UK IP5 3RE

1. INTRODUCTION

The volume of digital images has increased dramatically in recent years and as a result a crisis is now taking place within a broad range of disciplines that need and use visual material. Whilst storage and capture technologies are able to cope with the huge numbers of images, poor image retrieval is in danger of rendering many repositories valueless because of the difficulty of access. These problems have motivated much research into content based image retrieval [1,2,3,4,5] and many commercial and laboratory systems are described in the literature [6-13].

These conventional approaches suffer from some disadvantages. Firstly there is a real danger that the use of any form of pre-defined feature measurements will preclude solutions in the search space and be unable to handle unseen material. Secondly the choice of features in anything other than a trivial problem is unable to anticipate a user's perception of image content. This information cannot be obtained by training on typical users because every user possesses a different subjective perception of the world and it is not possible to capture this in a single fixed set of features and associated representations.

Relevance feedback is often proposed as a technique for overcoming many of the problems faced by fully automatic systems by allowing the user to interact with the computer to improve retrieval performance. However, retrieval should not require the user to have explicit knowledge of the features employed by the system and users should not have to reformulate their visual interests in ways that they do not understand.

This paper employs a robust similarity measure [14] that imposes only very weak assumptions on the nature of the features used in the recognition process. This approach does not make use of a pre-defined distance metric plus feature space in which feature values are extracted from a query image and used to match those from database images, but instead generates features on a trial and error basis during the calculation of the similarity measure. This has the significant advantage that features that determine similarity can match whatever image property is important in a particular region whether it be a shape, a texture, a colour or a combination of all three. It means that effort is expended searching for the best feature for the region rather than expecting that a fixed feature set will perform optimally over the whole area of an image and over every image in the database. By generating thousands of random features and applying them on a trial and error basis as an integral part of the calculation of the similarity value, it is shown that a consistent measure is obtained that is not dependent upon any one or group of specific pattern measurements or representative training sets. The similarity measure is applied to the problem of distinguishing between fingerprint images.

2. COGNITIVE VISUAL ATTENTION

The model of Cognitive Visual Attention (CVA) described in this paper is a generalization of an earlier model of Visual Attention [15,16] and relies upon the similarity of pairs of neighbourhoods taken from patterns A and B.

Let a set of measurements **a** on pattern A correspond to a location **x** in A in bounded n-space $(x_1, x_2, x_3, \ldots, x_n)$ where

$$\mathbf{x} = (x_1, x_2, x_3, \ldots, x_n) \quad \text{and} \quad \mathbf{a} = (a_1, a_2, a_3, \ldots, a_p)$$

Define a function **F** such that $\mathbf{a} = \mathbf{F}(\mathbf{x})$ wherever **a** exists. It is important to note that no assumptions are made about the nature of **F** eg continuity. It is assumed that **x** exists if **a** exists.

Consider a neighbourhood N_x of **x** where

$$\{\mathbf{x}' \in N_x \text{ iff } |x_i - x'_i| < \varepsilon_i \ \forall\ i\}$$

Select a set of m random points $S_\mathbf{x}$ in N_x where

$$S_x = \{\mathbf{x}'_1, \mathbf{x}'_2, \mathbf{x}'_3, \ldots, \mathbf{x}'_m\} \text{ and } \mathbf{F}(\mathbf{x}'_i) \text{ is defined.}$$

Select a location **y** corresponding to the set of measurements **b** on pattern B for which **F** is defined where

---

* f.stentiford@ee.ucl.ac.uk; http://picard.ee.ucl.ac.uk/~fstentif

$$|x_i - y_i| \leq s$$

Define the set $S_y = \{\mathbf{y}'_1, \mathbf{y}'_2, \mathbf{y}'_3, ..., \mathbf{y}'_m\}$ where

$$\mathbf{x} - \mathbf{x}'_i = \mathbf{y} - \mathbf{y}'_i \text{ and } \mathbf{F}(\mathbf{y}'_i) \text{ exists.}$$

The neighbourhood of **x** is said to match that of **y** if

$$|F_j(\mathbf{x}) - F_j(\mathbf{y})| < \delta_j \text{ and } |F_j(\mathbf{x}'_i) - F_j(\mathbf{y}'_i)| < \delta_j \ \ \forall \text{ i,j.}$$

In general δ_j is not a constant and will be dependent upon the measurements under comparison ie.

$$\delta_j = f_j(\mathbf{F}(\mathbf{x}), \mathbf{F}(\mathbf{y}))$$

In the case of two-dimensional visual patterns the CVA score of a pixel **x** in the pattern A should be high if a high proportion of randomly selected $S_\mathbf{x}$ match S_y for a fixed **y** in pattern B**.**

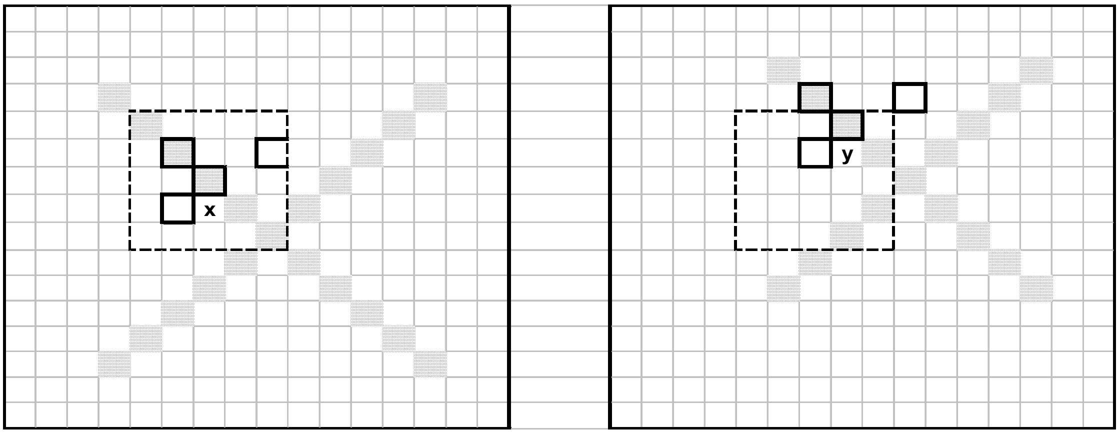

Figure 1 Neighbourhood at location x matching at location y

The CVA score of a location **x** is incremented each time one of the set of M neighbourhoods $S_\mathbf{x}$ matches a neighbourhood S_y surrounding some **y** in pattern B. This means that pixels **x** in A that achieve large numbers of matches between a range of M neighbouring pixel sets $S_\mathbf{x}$ and pixel neighbourhoods around **y** in B are assigned high CVA scores. In Figure 1, m = 3 pixels **x'** are selected in the neighbourhood of a pixel **x** in pattern A and matched with 3 pixels in the neighbourhood of pixel **y** in pattern B. Each of the pixels might possess three colour parameters, so **F**(**x'**) = **a** = (r, g, b) and the neighbourhood of the second pixel **y** matches the first if the colour parameters of all m + 1 corresponding pixels have values within δ_j of each other.

A parameter s is introduced to limit the area in pattern B within which the location **y** is randomly selected. S=2 defines the dotted region in figure 1. This improves the efficiency of the algorithm in those cases where it is known that corresponding objects in the two images are shifted by no more than s pixels. In effect s represents the maximum expected mis-registration or local distortion between all parts of the two images.

Some image analysis techniques carry out comparison calculations between images using patches that are contiguous neighbourhoods in which values from all the pixels are employed in the calculation. Patches match when a measure of correlation exceeds a certain threshold. This approach is unable to make best use of detail that is smaller than the size of the patch except in the case in which the correlation measure is designed to identify specific features. The random pixel neighbourhoods $S_\mathbf{x}$ described here do not suffer from this disadvantage.

3. APPLICATION TO BLACK AND WHITE IMAGES

In the case of black and white images F(**x**) = {0,1} and we take ε=0 so that a match occurs only if all m+1 pixels in a neighbourhood in A match exactly those in B.

Define the CVA score of pattern A with respect to B as

$$C_{AB} = \sum_{x \in A} \Big(\sum_{M,\, y \in B} (1 \,|\, S_x matches\, S_y, 0 \,|\, otherwise) \Big/ \sum_{A} black\ pixels$$

C_{AB} is normalised to be independent of image size. Furthermore the score is not incremented if any neighbourhood N_x contains solely black pixels because such $S_\mathbf{x}$ would tend to attach value to matching large expanses on black in A and B.

The gain of the scoring mechanism is increased significantly by retaining the pixel location **y** if a match is detected, and reusing **y** for comparison with the next of the M neighbourhoods. It is likely that if a matching pixel configuration is generated, other configurations will match again at the same point, and this location **y** once found and reused, will accelerate the rise of the CVA score provided that the sequence is not subsequently interrupted by a mismatch. If however, S_x subsequently mismatches at that location, the score is not incremented, and an entirely new location **y** in pattern B is randomly selected ready for the next comparison. In this way competing locations in pattern B are selected against if they contain little commonality with the neighbourhood of **x** in pattern A.

The size of neighbourhoods is specified by the maximum distance (ε_i) of configuration components to the pixel in A being scored. The neighbourhood is compared with the neighbourhoods of M other randomly selected pixels in pattern B that are more than a distance epsilon from the boundary of the pattern. Typically $\varepsilon_i = 2$ and M = 50, with m = 3 neighbouring pixels selected for comparison. Larger values of s and the ε_i are selected according to the scale and distortion in the patterns being analysed.

4. CLUSTERING

The CVA scores were computed for all pairs in a small set of 1000x1000 binary fingerprint images. The examples are shown in Figure 2 with scores tabulated in Table 1.

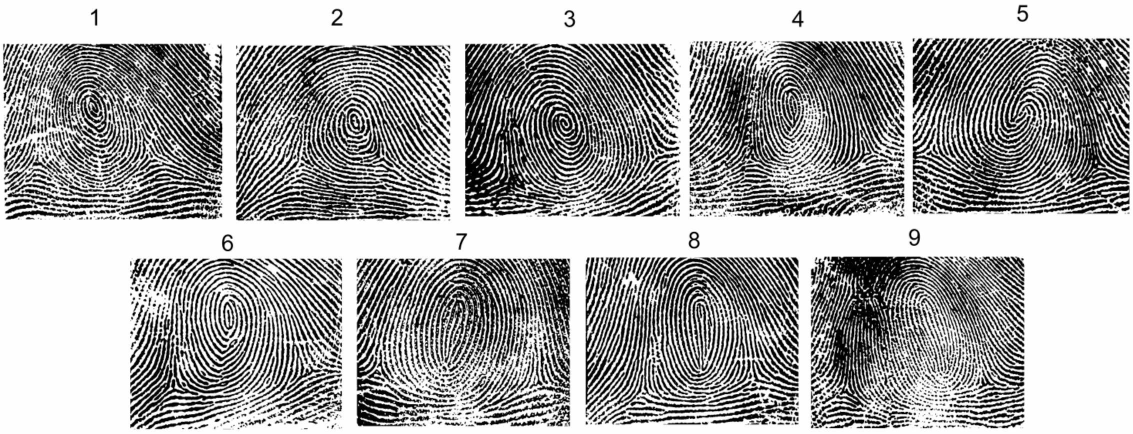

Figure 2 Fingerprint images

In all cases the self similarity score C_{AA} is greater than C_{AB} for all B≠A. The greatest sum $\max_i(\sum_j C_{ij})$ of the scores in each row indicates the pattern that is the most similar to the others (print 8), and the highest column sum $\max_j(\sum_i C_{ij})$ indicates the pattern to which most other patterns are the most similar (print 8). The asymmetry of the CVA score is apparent from the matrix of scores. In this case print 8 can be considered the ‘centre’ of the cluster when judging whether new patterns are members of the same group. The similar concept of ‘vantage’ patterns has been used for efficient image retrieval in work by Vleugels et al [17]. Candidate patterns would be members of the cluster if the CVA scores with print 8 were greater than the minimum score already in the cluster (print 9 score = 77.2).

..	1	2	3	4	5	6	7	8	9	Sum
1	**28.5**	9.9	9.7	10.4	10.2	9.9	9.5	10.7	9.6	79.9
2	9.9	**29.7**	10.1	9.6	10.5	10.4	9.6	11.0	9.3	80.3
3	9.2	10.1	**30.4**	9.8	10.9	10.0	10.2	10.0	9.5	79.8
4	9.8	9.4	9.9	**28.3**	10.3	10.3	9.5	11.6	9.4	80.1
5	9.7	10.3	10.7	10.1	**29.2**	9.9	10.5	10.8	9.5	81.5
6	10.4	11.4	11.0	11.6	11.0	**31.2**	11.0	13.1	10.0	89.6
7	9.4	9.6	10.2	9.8	10.9	10.3	**28.0**	11.2	8.9	80.3
8	10.8	11.0	10.3	12.1	11.5	12.5	11.5	**30.5**	10.9	90.5
9	9.3	9.3	9.8	9.4	9.7	9.0	8.8	10.6	**29.0**	75.9
Sum	78.5	81.1	81.6	82.7	85.1	82.2	80.5	89.0	77.2	

Table 1 Pattern similarity scores for fingerprint set

6. EVALUATION

Real scene of crime prints will be distorted in a variety of ways and identification cannot be achieved by simple template matching. The effects of translation, rotation, scaling and noise on the scores have been investigated by transforming one of the prints and observing changes in the score relative to the centre of the cluster (print 8). It is then possible to identify the level of distortion which causes the pattern to move outside the bounds of the cluster as an indication of the robustness of the approach. Results indicated that a 2-3% change in scaling could be tolerated, the score was reliable provided any translation was less than the value of the shift parameter, s, and a rotation of up to 10% was also acceptable. The approach lends itself to the analysis of more general images to establish similarity relationships between image regions.

REFERENCES

1 Y. Rui, T. S. Huang, and S. F. Chang, "Image retrieval: current techniques, promising directions and open issues," Journal of Visual Communication and Image Representation, Vol 10, No 1, pp39-62, March 1999.

2 A.W.M. Smeulders, M. Worring, S. Santini, A. Gupta and R. Jain, "Content-Based Image Retrieval at the End of the Early Years," IEEE Trans PAMI, Vol 22, No 12, pp 1349-1379, December 2000.

3 C. C. Ventners and M. Cooper, "A review of content-based image retrieval systems," JISC Technology Applications Programme, http://www.jtap.ac.uk/reports/htm/jtap-054.html.

4 R. C. Veltkamp and M. Tanase, "Content-based image retrieval systems: a survey," March 2001, http://www.aa-lab.cs.uu.nl/cbirsurvey/cbir-survey/.

5 E. Izquierdo, V. Mezaris, E Triantafyllou, and L-Q. Xu, "State of the art in content-based analysis, indexing and retrieval", IST Project IST-2001-32795, Network of Excellence in Content-Based Semantic Scene Analysis and Information Retrieval, Sept. 2002, http://www.iti.gr/SCHEMA/library/index.html

6 M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, M. Gorkani, J. Hafner, D. Lee, D. Petkovic, D. Steele, and P. Yanker, "Query by Image and Video Content: The QBIC System," IEEE Computer, 1995.

7 A. Pentland, R. W. Pickard, and S. Sclaroff, "Photobook: Content-Based Manipulation of Image Databases," SPIE Storage and Retrieval of Images and Video Databases II, No 2185, Feb 6-10, San Jose, 1994.

8 C. Carson, S. Belongie, H. Greenspan, and J. Malik, "Blobworld: Image Segmentation using Expectation-Maximisation and its Application to Image Querying," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol 24, No 8, pp 1026-1038, August 2002.

9 J. R. Smith and S-F. Chang, "VisualSEEk: a fully automated Content-Based Image Query System," Proc. ACM Int. Conf. Multimedia, pp 87-98, Boston MA, November 1996.

10 W-Y. Ma and B. S. Manjunath, "NeTra: a Toolbox for Navigating Large Image Databases," Multimedia Systems, Vol 7, pp 184-198, 1999.

11 A. Gupta and R. Jain, "Visual Information Retrieval," Communications of the ACM, Vol 40, No 5, pp 70-79, May 1997.

12 J. Dowe, "Content Based Retrieval in Multimedia Imaging," Proc SPIE Conf. Storage and Retrieval for Image and Video Databases, 1993.

13 J. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Gorowitz, R. Humphrey, R. Jain, and C. Shu, "Virage image search engine: an open framework for image management," Proceedings of the SPIE, Storage and Retrieval for Image and Video Databases IV, San Jose, CA, pages 76-87, February 1996.

14 F. W. M. Stentiford, "An attention based similarity measure with application to content based information retrieval," Storage and Retrieval for Media Databases conference at SPIE Electronic Imaging 2003, 20-24 Jan, Santa Clara.

15 F. W. M. Stentiford, "An estimator for visual attention through competitive novelty with application to image compression," Picture Coding Symposium, Seoul, 24-27 April, 2001.

16 F. W. M. Stentiford, N. Morley and A. Curnow, "Automatic identification of regions of interest with application to the quantification of DNA damage in cells," Proc SPIE Vol 4662, San Jose, 20-26 Jan 2002.

17 J. Vleugels and R. C. Veltkamp, "Efficient image retrieval through vantage objects," Pattern Recognition, vol 35, pp 69-80, 2002.