4

Artefact matching and retrieval using the Generalised Hough Transform

Peter Durham, Paul Lewis & Stephen Shennan

(University of Southampton, Southampton, UK)

4.1 Introduction

One of the main tasks of the archaeologist is concerned with the classification or identification of artefacts and much previous work has been reported on the application of computer technology to the storage and retrieval of artefact shape data. The bulk of this work has been based on artefact profiles following the manual approach to artefact classification using line drawings of outlines.

For example, Hall and Laflin (1985) used the B-splines curve-fitting technique to generate object outlines, and suggested that this method could be used for storing and matching profiles. Gero and Mazzullo (1984) used closed-form Fourier series to describe an object's shape, and suggested that the Fourier series could be used for matching. Peter Main has carried out much work on artefact shape: on databases for archaeological artefacts, representation of their shapes by tangential profiles schemes, and on artefact retrieval by shape from such databases (Main 1986, 1987).

In this paper we describe a prototype System for Matching ARTefacts, (SMART). It uses a versatile approach to artefact classification based initially on the generalised Hough Transform (GHT) although a range of classification techniques are anticipated, including texture which is discussed briefly in section 4.6.4. The possible use of the GHT for artefact shape matching was proposed in a paper by Lewis and Goodson in 1990 (Lewis & Goodson 1991) as it is a robust matching algorithm capable of producing a response when only a partial shape is available.

Here we demonstrate the effectiveness of the technique not just as a tool for profile matching using line drawings, but for a range of artefact retrieval and matching tasks. These include retrieval of similar shapes when only a fragment is available and retrieval of line drawings when the input image is from a photograph.

One of the other benefits of the approach described here is that the shape data for the database does not have to be extracted and represented explicitly either by manual digitisation of the outlines or by some other means. A raster image of the drawing or the photograph is simply captured into the image database either by scanning the paper line drawings or the photograph using a flatbed scanner or by using a video camera and digitiser. No further pre-processing is needed before the application of the algorithms.

Section two of the paper describes the way in which the SMART system appears to the archaeologist, in other words the current user interface for the search system and the options available. The third section describes the methods used for handling the images. Section four describes the details of the generalised Hough Transform and the techniques for calculating the similarity index used for establishing the matches. The fifth section provides data from a test set of images and shows the results achieved and in the final section proposed enhancements to the SMART system are discussed.

4.2 The user interface

SMART allows the user to find a match or a set of matches for an image of an "unknown" object with a set of images of known objects held in the image database. The database of images is essentially a collection of raster image files the names of which indicate the object whose image is held therein. The database is established by capturing the images of known objects using one of several available techniques as described in section 4.3.

The system runs under X-windows and was initially developed on a Sun workstation. When it is invoked, the user may choose the source for the input image (the unknown). This may be either an existing image file or an image capture device such as a video camera. If an existing image file is to be used the user is asked to provide its name. If a new image is to be captured, step by step instructions are given on how to position the object or drawing, and the image is then captured. In either case, the image is displayed on the control panel. The user can then select the way in which to perform the match based retrieval from a menu of options. The default options are for matching the whole shape of an image input from a pre-existing file. A sample screen display for this step is illustrated in Fig. 4.1.

The two matching options currently available are:

- The whole shape. If this option is selected, the whole of the input image is used for the search and compared with all the shapes in the database.
- Part of shape. If this option is selected, the user specifies a subsection of the input image to be used for matching by drawing around the appropriate section with the mouse. In this way, the user may instigate a search for a "form" sub-part of the shape, such as the lip of a bowl, or a surface sub-part of the object, such as a decorative design. The section to be used is highlighted on the input image.

In Fig. 4.2 the whole of the image has been selected as the basis for the search. Whether a whole or part shape is selected, once initiated the search tries to find a match for the selected area in each image in the library and the names of the images in which the best N matches are found are listed, together with a similarity index based on the quality of the match. This is shown in Fig. 4.3. In our examples we work with N set to five.
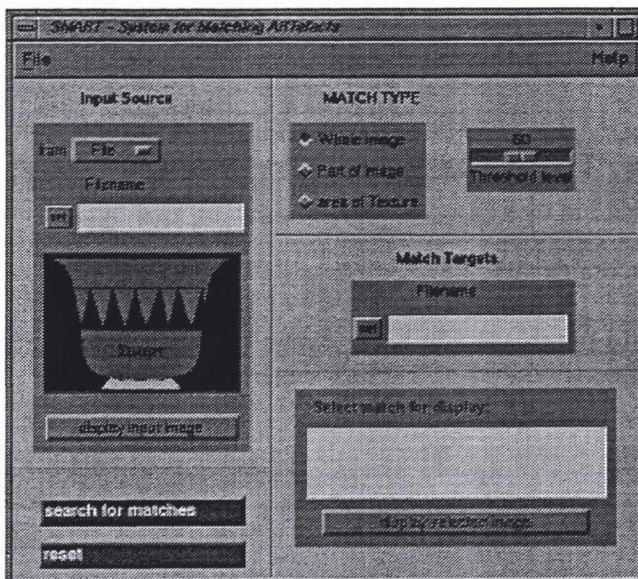
Figure 4.1: Image input from file, and whole-matching selected.

Figure 4.2: Input image and target list selected.

The user may select one or more of the images from the best match list for display alongside the input image. If the part-matching option was used, the location of the area which provided the match is again highlighted on each image as shown in Fig. 4.8.

The system has recently been transferred to a Silicon Graphics workstation running Motif, a graphical user interface running on top of X-windows, enabling a more user-friendly interface to be developed (see Figs 4.1 – 4.3). This consists of a control panel on which the user sets the desired options for the source of the input image, its name, the type of matching required, and the library of images to be searched. The control panel also contains buttons to start the matching, capture and display of images, etc.

4.3 Image capture and display

Images may be captured with a video camera or document scanner attached to an image digitiser. At least one of these special facilities must be available on the computer system to be used unless the images are to be imported from elsewhere. The separate "CAPTURE" utility guides the user through the process of setting up the source object or drawing, adjusting the conditions to obtain an appropriate contrast and capturing the image into a file. The filename in which the image is to be stored must be provided and the captured images are held in pgm format files. This format is widely used for image file handling and a large number of public domain utilities are available for manipulating them. The SMART system at present assumes that grey scale images are being used although extensions to colour images are planned.

Facilities exist within the SMART system for displaying any image from the database by naming the image file and the image is displayed in a relocatable window on the display.
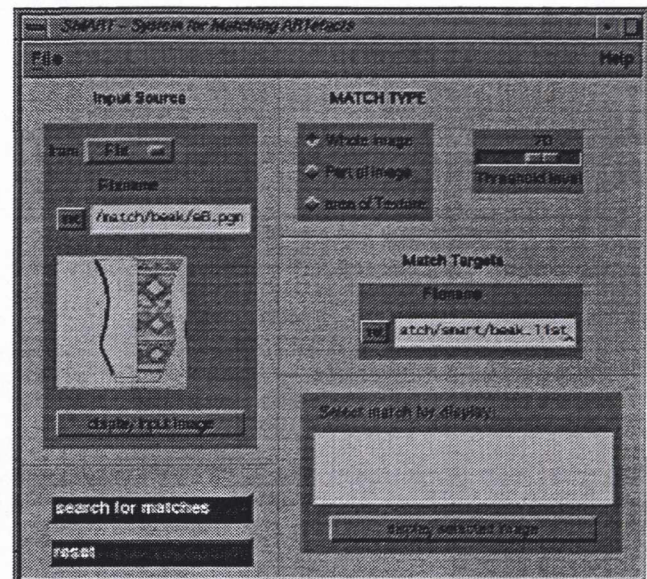
4.4 The Generalised Hough Transform

The Hough transform was first proposed as a method of finding lines and circles in images. It has been shown by Ballard (1981) that the method may be generalised to detect arbitrary, complex shapes, and has the potential to be invariant to translational, scale or rotational differences. A definition of the translation-invariant version of the algorithm, as used in the current work, is given here:

4.4.1 The algorithm

First, a reference point for the input object image is chosen. Commonly, the reference point is the centroid of the object, but in this case the origin of the co-ordinate system (i.e. the top-left corner) was used. The edge points for the input object are then detected (see below). For each point on the edge of the input object, a vector r from that edge point to the reference point is calculated. The vector r is stored in a table (called the r-table) where it is indexed by the angle of orientation of the edge from which it was derived.

To detect objects similar to the input object in an image from the database using the Hough transform, the edges in the database image must first be detected, and a 2-D accumulator array is set up, whose axes represent the co-ordinates of possible locations of the reference point in the image. Then, for each edge point in this image, the r-table is searched for vectors which refer to edges with the same orientation as that edge point. The cell in an accumulator array corresponding to the position of the other end of the vector is incremented, for each of the vectors with this orientation. Once every edge point in this image has been processed, the accumulator cell containing the highest value corresponds to the location of the reference point for the shape in the current image. Each edge point may cause many cells to be incremented, but providing the shape is actually present in the image only one cell will be incremented by *every* edge point.
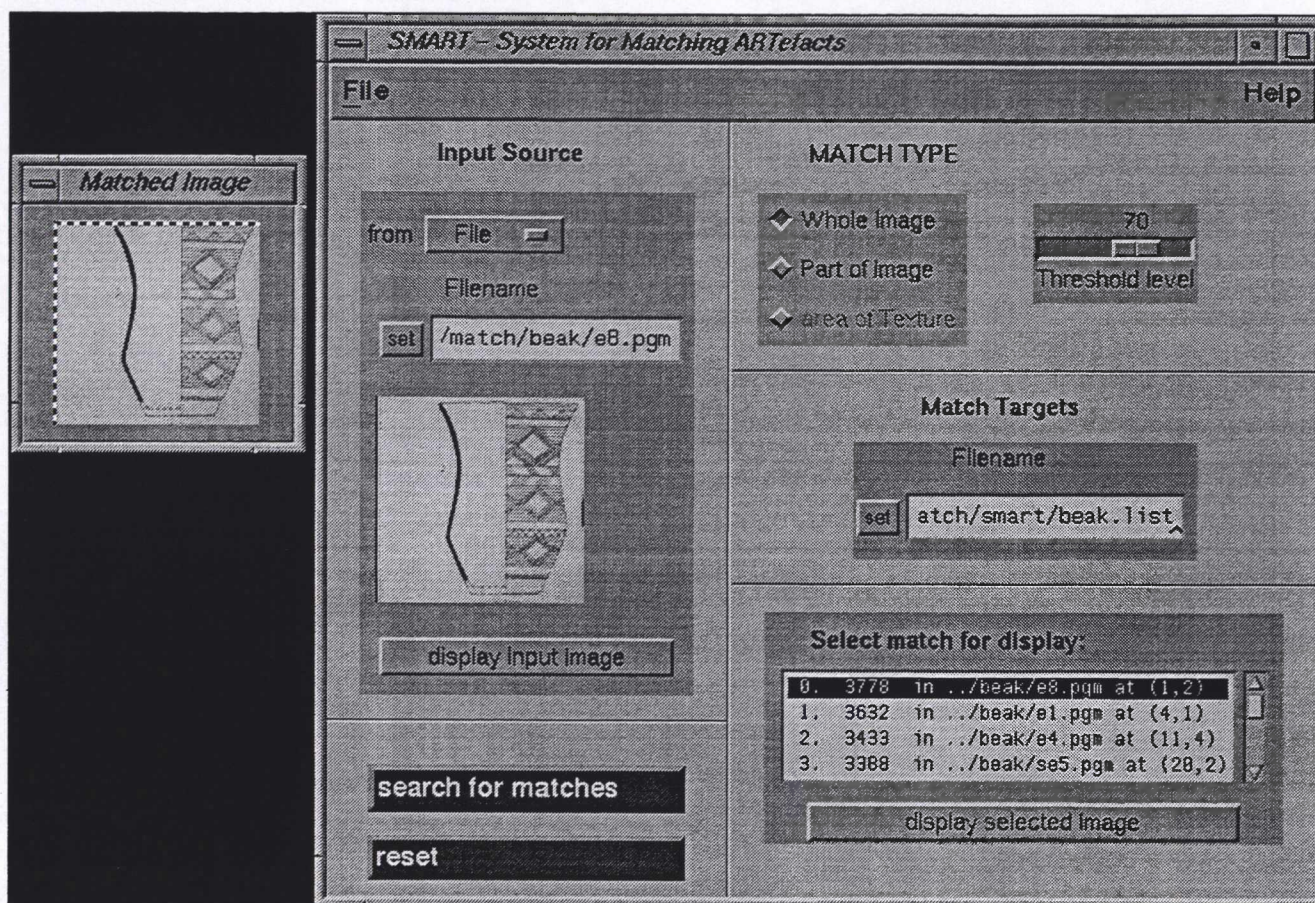
Figure 4.3: Displaying image selected from best match list.

The height and spread of the peak in the accumulator indicate the goodness of the match. A perfect match should produce a peak cell with a count similar to the number of edge points in the shape. Accordingly, the similarity index used here is based on the ratio of the peak's value to the number of edge points.

4.4.2 Edge detection

In the original Hough Transform, the edge points for the original shape are generated using the equation which describes that shape.

The generalised Hough Transform uses an image of the object being sought rather than an equation or set of equations. This enables it to search for arbitrary and/or complex objects which it is not possible to describe by equations. The edge information must be generated using one of the several edge detection algorithms that exist.

In this study, a Sobel edge detection template has been used. This is a template which returns a value for edge strength in a set direction at a given point (*i.e.* the gradient of the grey-scale values at the current point in the set direction is calculated). The edge strength is calculated by weighting the adjacent pixels' grey-scale values using the Sobel mask (see Fig. 4.4) and adding them together. To find edges in all orientations, the artetion is done using several different Sobel masks for several different orientations. The values calculated are then combined trigonometrically. At present four directions are used, at 0°, 45°, 90° and 135° to the vertical. The results are held

in an edge map, a 3-dimensional array the same size as the original image, containing values of edge strength and edge orientation. The Sobel edge detector is used to calculate an edge strength for each pixel in the image.

4.4.3 Thresholding

In many cases the strength of the edge will be zero or very small. The Hough Transform does not take edge strength into account, and so very weak edges will have the same influence as very strong ones. Not only does this slow down the algorithm but in general the stronger edges are more important in defining the object's shape. In this implementation, the edge map has been thresholded, with edges below a certain strength being set to zero. The default level of thresholding is 70% of the maximum edge magnitude in the image. When using the part-matching facility to match surface features, these may not be pronounced enough to survive the thresholding. The user is therefore able to lower the threshold when the part-matching option is selected.

The image which is being searched for the object also needs to be searched for edges, to enable the Hough Transform to be applied to it. The edges are detected using the thresholded Sobel edge detector in the same way as described above.

It is perhaps worth noting that, unlike the original Generalised Hough Transform, this implementation makes use of internal edges as well as boundary edges so that patterns

27

-1	0	1
-2	0	2
-1	0	1

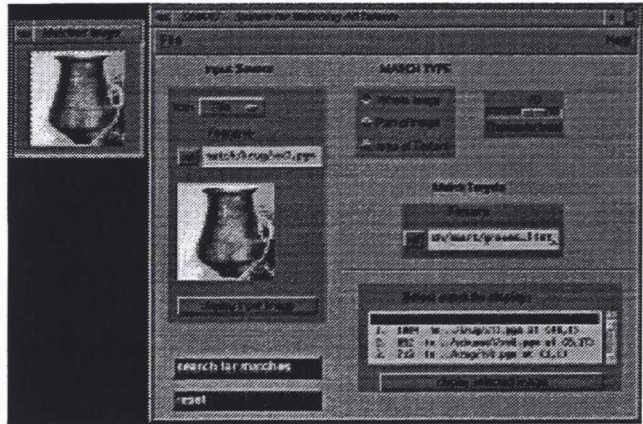*Figure 4.4: Sobel edge detection mask for vertical edges.*

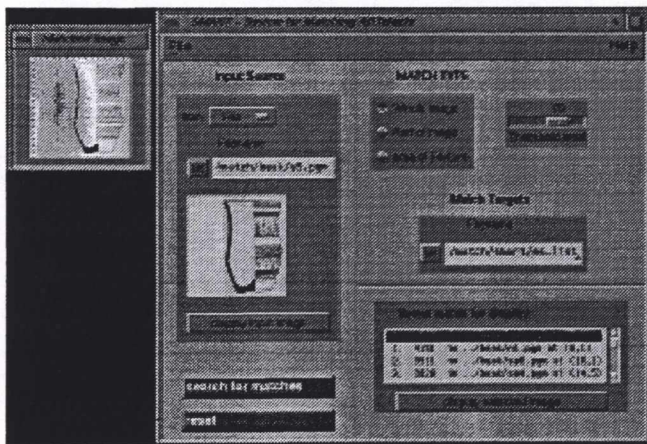*Figure 4.5: Photo vs. photo.*

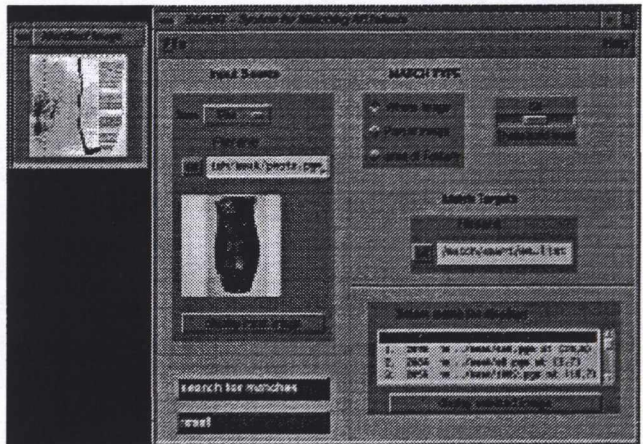

Figure 4.6: Drawing vs. drawing.

Figure 4.7: Matching a photo to drawings.

in the artefact can contribute to the matching as well as the profile.

4.5 Testing and examples

The matching program was tested using two sets of images. The photographs are of objects from the Únetice culture from the region of Lovosice in the Czech Republics. They have been input from photographs published by the National Museum, Prague (Moucha 1961). The individual codes have no particular meaning. Those images prefixed "beak" are Neolithic beakers from the British Isles. The images have been input from line drawings of 34 beakers, taken from more than a thousand such drawings published by D. L. Clarke (Clarke 1970). He also made a classification of the beakers, based on his personal visual comparison and a certain amount of quantitative analysis. The letters in the beakers' individual codes refer to Clarke's types. All these images were input via a video camera using the CAPTURE utility described above in section 4.3:

1 **Photograph with Photograph — Fig. 4.5.** This example is a test of the whole-matching option. The list of possible matches includes the input image, which can be seen to match best against itself as one would expect.

2 **Drawing with Drawing — Fig. 4.6.** Here beaker e5 is matched best against e4 (confirming Clarke's classification in this instance). It is interesting to note that although the *best* match was with another member of the same type, the next best ones were of different types. This happens because Clarke's classification was based on more than shape alone.

3 **Photograph with Drawing — Fig. 4.7.** This is a test of matching a photograph (which gives the same image as the real object) against line drawings. The pot in the photograph has matched well with beaker n5. This is an extremely important capability, as most pot finds are published in Britain as line drawings and hence most database images will be line drawings. It is usually easier for the human eye to see the shape of a pot's form and decoration in a line drawing. Photographs cannot show a cross-section (except by destroying the object), and the contrast in a photograph will never be as pronounced as with a line drawing, making surface patterning more difficult to discern. This is likely to be the most common mode of using the matching system.

4 **Part matching — Fig. 4.8.** This is an attempt to match the base-form of the pot. The best match was not a base match, but other good matches in the list were. This is probably because the differences in the edge shapes are too small for the resolution being used (see section 4.6.2).
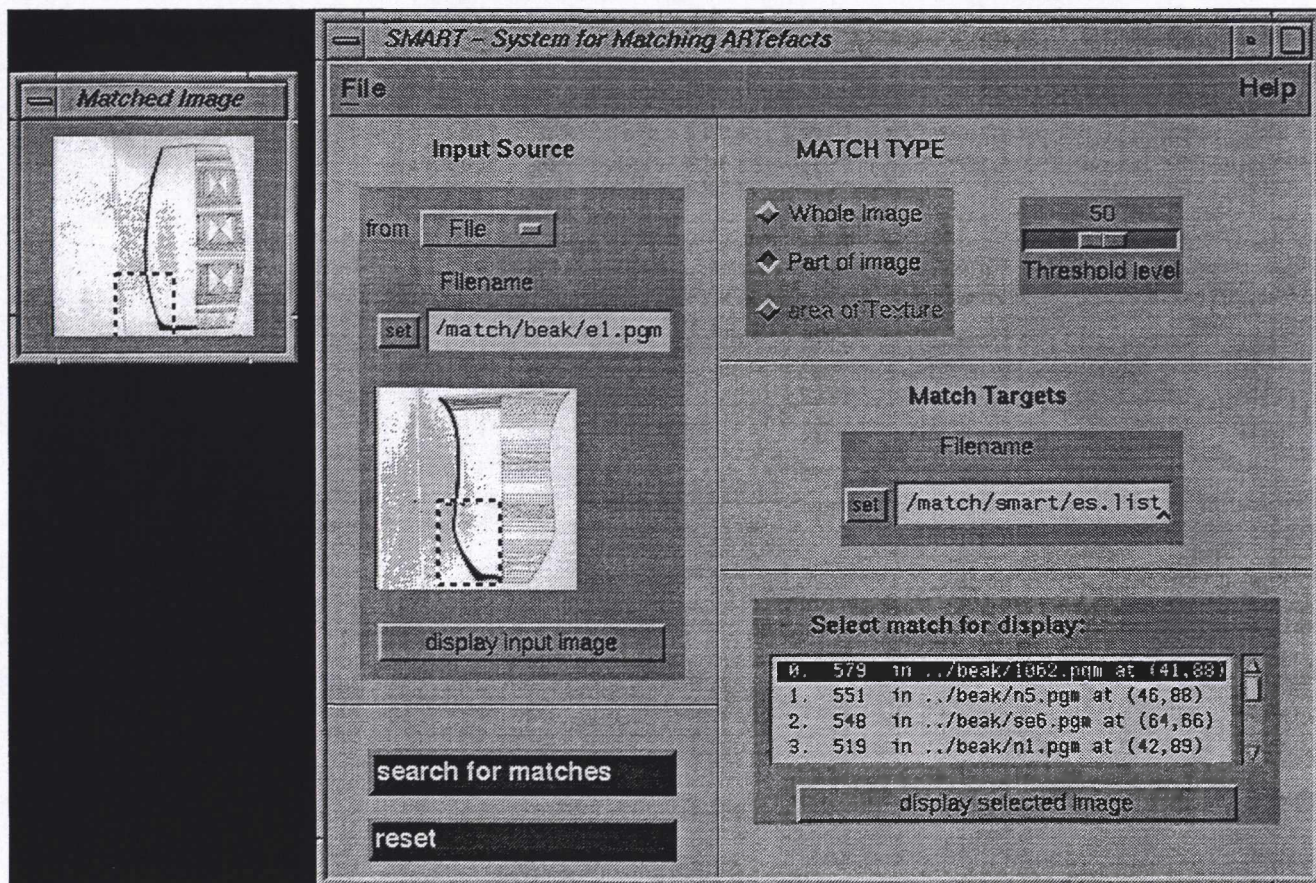
28

Figure 4.8: Part-match example.

4.6 Proposed Enhancements

4.6.1 Speed

The first version of the system, which ran on a Sun 3/270 workstation, was relatively slow. There are several ways in which the basic algorithm can be modified to speed it up, but we have taken advantage of advancing technology to use a faster computer. The matching system now runs on a Silicon Graphics Indigo (R4000) workstation, and is at least ten times faster. Example 2 took over 40 minutes on the Sun, but takes 65 seconds on the new machine.

4.6.2 Part-matching

The part-matching option did not work perfectly as originally implemented. This was probably due to the low resolution of the images enforced by the speed of the system. (The speed of the algorithm is inversely proportional to the size, and therefore resolution, of the image.) The features selected were only represented by a few pixels, which decreased the number and accuracy of the vectors in the r-table. It should now be possible to increase the image resolution in view of the system's improved speed.

4.6.3 Scale- and rotation-invariance

It would be a relatively simple operation to introduce scale- and rotation-invariance to the Hough Transform functions used in this work. For each of these parameters, another dimension has to be added to the Hough accumulator. For example, for rotation, the new dimension would be small angular intervals. Each plane of this dimension would be in effect a complete 2-dimensional accumulator using the translation parameters as before, except the object would be rotated through the appropriate angle. The new planes do not need to be created from the edge information, but can be calculated by a simple mapping of the vectors from the first plane (which has zero rotation). Similarly, the scale dimension would be quantised by small changes in scale, and the information is calculated by a mapping from the vector information in the original 2-dimensional r-table.

It was decided not to implement these extensions to the Hough Transform in the first instance as the extra processing involved would slow down the program. However the system has scope for its shape-matching facility to be enhanced in this way in due course.

4.6.4 Texture matching

It is intended that the system will be extended to enable objects to be matched on the basis of similarities in visual texture (surface patterning). Visual texture is concerned with the arrangement of pattern units, rather than the properties of the units themselves. If the object's surface decoration consists of a few large units, the part-matching option described should be used (see section 4.2). The procedure for using texture will be the same as with the part-shape matching option, except that the user selects the area of pattern they wish to compare.

Visual texture is quantifiable from the pixel values in the image. A practicable method of texture comparison

has been demonstrated by Haralick et al. (1973), which use a grey-scale co-occurrence matrix to generate a set of texture statistics from a given image. Preliminary work indicates that these statistics enable various pattern textures to be differentiated (e.g. plain, straight lines, *etc.*).

4.7 Conclusions

We have shown that the Generalised Hough Transform is a practicable and robust tool for matching whole and partial artefact shapes. We have described a prototype system for artefact matching and retrieval based on the technique and have indicated further avenues which will enhance the system to assist with archaeological artefact classification.

Bibliography

BALLARD, D. H. 1981. "Generalising the Hough Transform to detect arbitrary shapes", *Pattern Recognition* 13(2): 111-122.

CLARKE, D. L. 1970. *Beaker Pottery Great Britain and Ireland*, Vol.2. Cambridge University Press, Cambridge.

GERO, J. & J. MAZZULLO 1984. "Analysis of artifact shape using Fourier series in closed form", *Journal of Field Archaeology* 11: 315-322.

HALL, N. S. & S. LAFLIN 1985. "A computer aided design technique for pottery profiles", *in* Laflin, S. (ed.), *Computer Applications in Archaeology 1984*, pp. 178-188. University of Birmingham, Birmingham.

HARALICK, R. M., K. SHANMUGAM & I. DINSTEIN 1973. "Textural features for image classification", *IEEE Transactions on Systems, Man and Cybernetics* SMC-3(6): 610-621.

LEWIS, P. H. & K. GOODSON 1991. "Images, databases and edge detection for archaeological object drawings", *in* Lockyear, K. & S. P. Q. Rahtz (eds), *Computer Applications and Quantitative Methods in Archaeology 1990*, pp. 149-153. BAR International Series 565, Oxford.

MAIN, P. L. 1987. "Accessing outline shape information efficiently within a large database", *in* Laflin, S. (ed.), *Computer Applications in Archaeology 1986*, pp. 243-251. University of Birmingham, Birmingham.

MAIN, P. L. 1988. "Accessing outline shape information efficiently within a large database II : database compaction techniques", *in* Ruggles, C. L. N. & S. P. Q. Rahtz (Eds), *Computer and Quantitative Methods in Archaeology 1987*, pp. 243-251. BAR International Series 393, Oxford.

MOUCHA, V. 1961. *Nálezy Únetické Kultury na Lovosicku, Fontes Archaeologici Pragenses*, Vol.4, Museum Nationale Pragae, Prague.

Authors' addresses

P. Durham and S. J. Shennan,
Department of Archaeology,
University of Southampton
GB- Southampton.

P. H. Lewis
Department of Electronics and Computer Science
University of Southampton
GB- Southampton.