Image Retrieval Methods for a Database of Funeral Monuments

Jonathan Howell and David Young

School of Cognitive and Computing Sciences
University of Sussex, UK

Collaborators

Nigel Llewellyn
David Hickman
Hilary Tunley

With assistance from the Photographic and Design Unit, University of Sussex

Funding

Arts and Humanities Research Board

2

Objective

To investigate form-based recognition in a database of images of funeral monuments, obtained as part of a new census of post-Renaissance monuments in English churches.

The monuments form an important social and art-historical record. Image-based search may provide a valuable adjunct to conventional text-based search for users of the database.

The Database

Monuments may be flat, two-dimensional with relief, or three-dimensional. They incorporate a variety of sculptural, architectural and decorative elements, and often substantial amounts of text.

Access may be difficult, and camera angles are constrained by walls and church furniture. Lighting is highly variable. It is thus impossible to standardise the scale or orientation of the images.

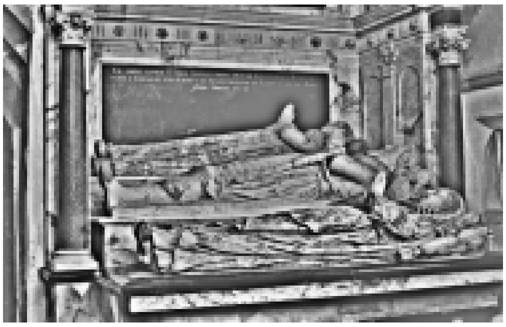
Image preprocessing

The effects of grey-level and colour changes due to illumination variation can be reduced by removing low spatial frequencies from intensity images.

This is achieved by subtracting a smoothed copy of the image from the original image — in effect, a Difference of Gaussians convolution where the inner Gaussian has a width of less than a pixel. There is one free parameter — the scale for smoothing.

It would be possible to go further and use an edge map, but it is likely that this would discard too much information.

Although this process is intended to facilitate matching, the results are also useful for display, in that they show detail in dark or bright areas in the way a good drawing does.



5

We search the database by identifying a target patch and matching it against all the images at a variety of positions and scales. There are many possible approaches to this: we adopted Gabor coefficient matching as a method with proven success in other domains.

Matching by Gabor coefficients

Preprocessing is applied to each image offline, and is slow. Vectors of Gabor coefficients are used to represent each image.

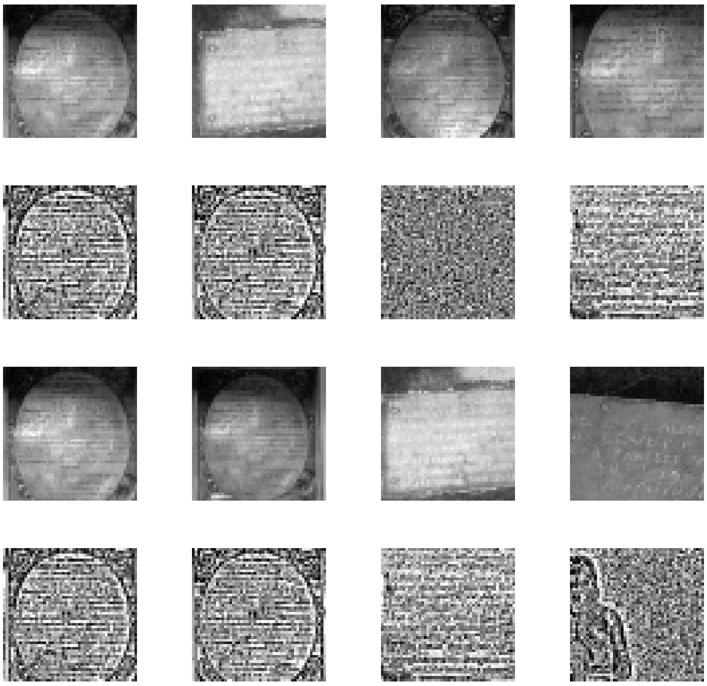
Images are divided into overlapping patches, each 200 × 200 pixels or 300 × 300 pixels. The centres of the patches lie on a grid of spacing equal to 5% of the patch size.

For each patch a vector of 126 Gabor coefficients is found. This breaks down as follows:

- The vector is made up of 21 subvectors.
- Each subvector has 6 Gabor coefficients: cosine and sine at 3 orientations.
- One subvector uses masks at the largest scale, centred on the centre of the patch.
- 4 subvectors use intermediate scale masks, centred on 4 quadrants of the patch.
- 16 subvectors use small scale masks, centred on a 4 × 4 grid across the patch.

Matching is done by comparing the coefficient vector for a target patch with the vectors for every patch in every image in the search set.

At present Euclidean distance is used. This is reasonably fast. We looked at both the full coefficient vector and matching using the Gabor amplitudes rather than the sine and cosine terms, to give a measure of position independence within the patch.



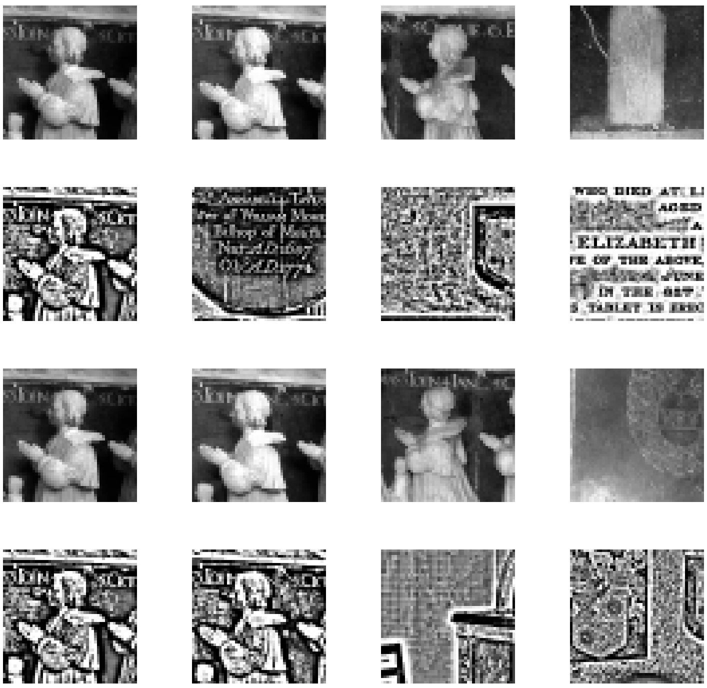
Columns: Leftmost column is target patch (which is also the best match in the database). The remaining columns are the 3 next-best matches in order.

Rows: From top to bottom: original images, all coefficients; processed images, all coefficients; original images, amplitudes only; processed images, amplitudes only.

The approach picks out inscriptions which have similar texture to the targets, but shape matching is difficult.

(Figure on next two pages have the same layout.)

6



Further work

1. A central problem is identifying regions of interest. Centring our search patches on a rectangular grid is computationally expensive and unsatisfactory. We need efficient ways to find candidate positions and scales in each image, against which to try matching the target.

2. Gabor coefficients are an effective choice of representation, but we will evaluate our choice of number of orientations and scales more systematically.

3. The matching metric is too simple: we will replace it with a learnt metric or one based on an analysis of the coefficient distribution.

4. The matching methods should be adapted to the specific needs of users — e.g. for texture matching or for script matching.