MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

6—3 Flexible features in texture similarity

Joseph K.P. Kuan *   Dan W. Joyce   Paul H. Lewis

Multimedia Research Group,
Dept. of Electronics and Comp. Science,
University of Southampton,
Highfield, Southampton, SO17 1BJ
United Kingdom

Abstract

In this paper, we demonstrate a texture technique in the context of content-based retrieval (CBR). The method is based on our previous work with some modifications. Although CBR has been widely studied, the novelty of our technique enables the users to adjust the edge features used for matching and the results show visual correspondence to the modified configuration. Hence this gives more freedom to the users over the control of a target texture. The development of the method is in a preliminary stage. However the potential of this method can lead to a number of significant developments, such as front end text based description with back end content based retrieval and a hierarchical model of a *texture dictionary* which can facilitate different abstraction levels of CBR.

1 Introduction

In [6], we demonstrated a new texture representation that outperformed two other methods, Multi-resolution Simultaneous Auto-regressive model (MRSAR) [7] and Statistical Geometrical Features (SGF) [2], with the entire Brodatz texture database [1] and complex texture patterns from a commercial catalogue. Our method uses edge and plain region information as a basis for the texture features.

In this paper, we investigate another potential benefit of our technique. The approach mainly makes use of edges which can be seen as primitive visual features in human vision. Therefore, we can extend this model into a content based retrieval (CBR) application with more control over structural features. In QBIC [3], the well known texture method by Tamura et al. [9] is adopted and has shown that each feature has close correspondence to human vision. However, these features (*coarseness, regularity, line-likeness, contrast, roughness* and *directionality*) do not provide control over the texture structure nor do they match based on edge skeletons since they used global visual features. An advantage in our application is that it lets the user adjust the features for further content based matches according to their conception of the texture structure. The other advantage is that if users want to start browsing for simple texture patterns, they only need to specify the desired features for similar matches without knowing what textures are stored in the database.

Several modifications have been made from our previous method and are described in section 2. In section 3, we demonstrate how the features interface can interact to retrieve texture corresponding to the adjustments made by the users and also present some results from CBR with visual similarity.

2 Edge Features

The method in [6] uses three different groups of features, namely label ratios, contrasts across edges, and conditional probabilities. Each group of features are based on 4 directional edge labels known as horizontal (H), vertical (V), left-diagonal (LD), and right-diagonal (RD) and homogeneous region label (B).

Firstly, the features are rearranged into three types: global, local, and coarse local. Global features are label ratios (i.e. the proportion of the total label counts for each label type) calculated over the whole image. Local features are the conditional probabilities of two labels appearing within a 3×3 neighbourhood. The coarse local features are intermediate between both local and global and indicate the continuity of each edge in a larger mask.

Before detecting any directional labels with a 3×3 mask, the significant changes of brightness values inside the mask have to be decided first. A different approach is used to check the change of pixel values

* E-mail: {jk94r, dwj95r, phl}@ecs.soton.ac.uk

Neighbouring Labels

Center Label	H	V	LD	RD	B
H	p(H|H)	p(V|H)	p(LD|H)	p(RD|H)	p(B|H)
V	p(H|V)	p(V|V)	p(LD|V)	p(RD|V)	p(B|V)
LD	p(H|LD)	p(V|LD)	p(LD|LD)	p(RD|LD)	p(B|LD)
RD	p(H|RD)	p(V|RD)	p(LD|RD)	p(RD|RD)	p(B|RD)

Figure 1: Conditional Probability Matrix

which is similar to an erosion operation in morphological techniques [4]. The procedure is performed as follows:

```
For each pixel (except centre) in the mask
   if (centre brightness -
       pixel brightness >= T)
     increment the count
If (count > 0 && count < mask size - 1)
  detect edge direction
else
  label the centre pixel as B
```

T is set to 12 in this paper. For finding the initial edge directions, we apply the method by Patel et al. [8] which is also used in our previous work [6]. Hence a labelled image is created with H, V, LD, RD and B. The ratio features, R, are simply the count of each label occuring in an image divided by the total number of all labels in the image.

For the local features, we adopted the conditional probability matrix in [6]. This is computed by sampling using a 3×3 mask on each pixel in the labelled image, I, and calculating the conditional probability, $p(j|e)$, of label, j, appearing in the mask given that the centre of the mask has a label, e. The probability is then averaged, summing and dividing by the number of e labels appearing in I. In this paper, we only used the conditional probabilities for the four directional edge labels, as shown in figure 1.

The edge continuity features, $cont_e$, are processed with a larger mask, 33×33, in order to capture the semi-local properties. The distance between each mask application is 12 pixels. For a given mask application, we calculate the length of the continuous edge in the mask passing through the central pixel. These edge lengths are then used to calculate the average edge length, $cont_e$, for each edge label, e. The $cont_e$ values are normalised to be in the range 0 to 1.

Once all the global, local and coarse local features are evaluated, a weighted Euclidean distance measurement is computed for this technique. The label ratios (global features) are used as the weights, such that the weight of a label e, w_e, is calculated as the average of R'_e and R''_e which are the ratio features of query and stored images respectively:

$$w_e = (R'_e + R''_e)/2 \tag{1}$$

Then the square of the Euclidean metric is calculated as:

$$\textstyle\sum_e w_e \times s_e \tag{2}$$

$$s_e = \textstyle\sum_j (p'(j|e) - p''(j|e))^2 + (cont'_e - cont''_e)^2$$

The terms $p'(j|e)$ and $p''(j|e)$ are the conditional probabilities features of query and stored images respectively, whereas $cont'_e$ and $cont''_e$ are the continuity features of edge label e of query and stored images respectively. Both summations are over appropriate edge labels.

3 Experiment

In this section, we demonstrate some of the similarity matching results selecting five different textures with random (D32), quasi-structural (D62, D67, D70) and structural (D56) appearances as query images.

Figure 2 shows the visual similarity on the nearest matches from all the 112 digitised (256×256) Brodatz textures for the new edge method. Each row represents from the query image (far left) and the match results which are the 5 far right images. Most of the matches are visually similar.

A user interface was built for the technique with Java[1] Swing that enables users to control the values of feature ranges (shown in figure 3). The user selects a net-like texture as the base query image and the texture's statistics are calculated and displayed. From this, the user is able to modify the relationships between features, enabling them to base the query on one image, but modify certain properties to reflect their specific interest. For example, in Figure 3 the user wishes to query on an image with the general net-like properties of the base query image displayed. Then the user can modify the features values and retrieve other similar images where visual similarity corresponds to the adjustment. Figure 4 shows the results of burlap texture, D104, after the following changes:

- Increase the horizontal ratio and decrease the vertical ratio
- Increase the bindings of both horizontal and vertical to blank, $p(B|H)$ and $p(B|V)$.

[1] Trademark of Sun Microsystems

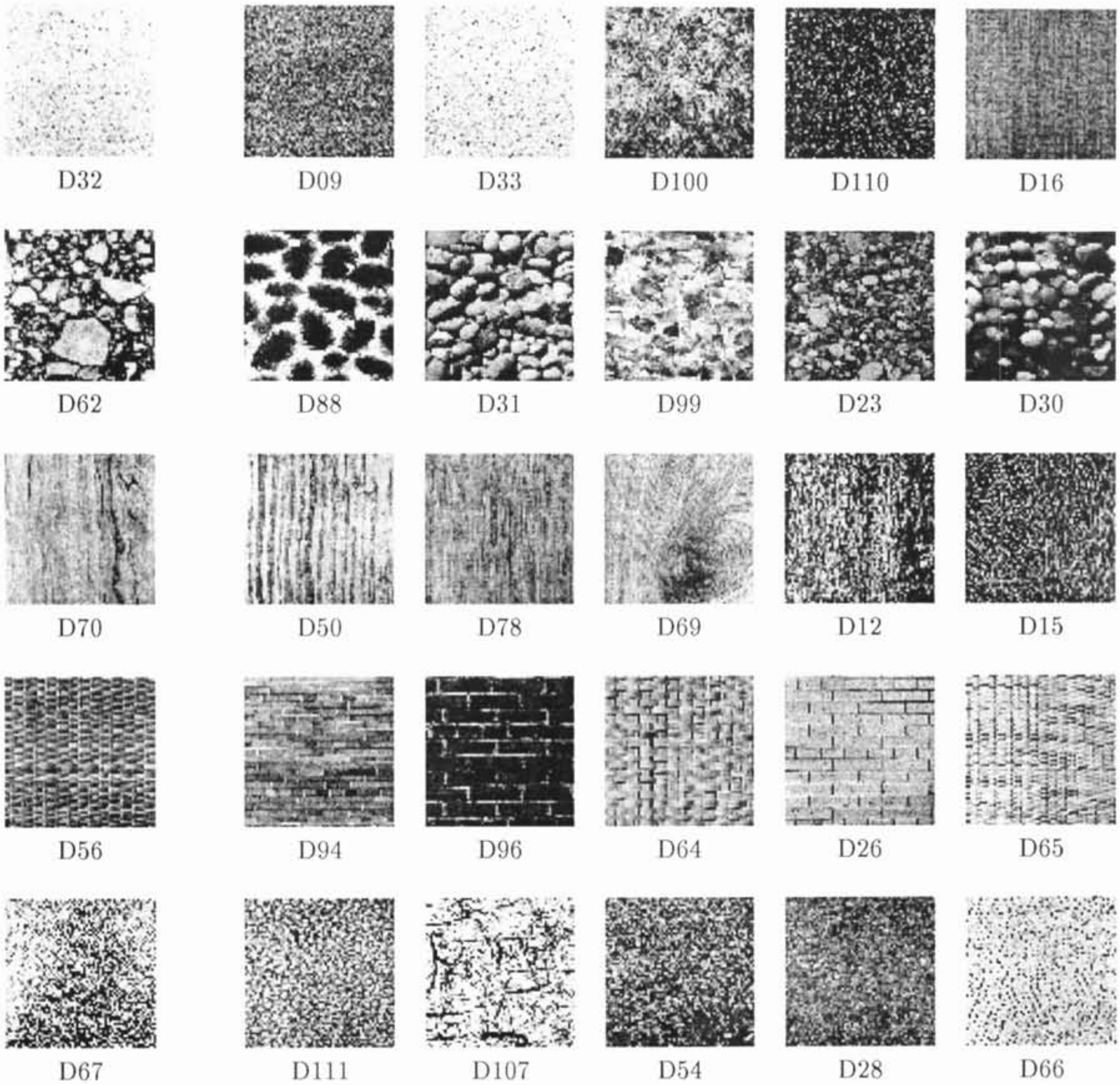

Figure 2: Similarity results on the entire Brodatz album

The results shown on the right of figure 4 demonstrate the effect of modifying the query image. Note that homogeneous regions are of similar granularity to the query image and horizontal edges dominate the retrieved textures. Also the global scale properties are similar in the top two retrieved textures and gradually increase (that is, the size of rectangular homogeneous regions grows-the size of the bricks) as the matching distance increases [2]. Qualitatively we might say that the modifications to the query image amount to straightening the burlap pattern but preserving the granularity of homogeneous regions.

These modifications are comparatively intuitive because they are specified in terms of global properties of the texture which are visible via the query image. This seems to have the effect of making the relationships specified by the sliding controls qualitative because of their reference to the displayed base query image. The user is not forced to conceptualise a texture *purely* in terms of the relationships between primitives because they can browse a set of images, and then use one as a base query image.

In order to solve this, we are investigating a mapping from the primitives used in matching (e.g. the edge labels and conditional probabilities that express the relationship between them) to a quasi-natural language interface. This enables the user to specify simple texture queries such as "strong vertical edges and weak horizontal edges". These are then parsed and an appropriate matrix of conditional probabilities is generated. The corresponding stored values which describe the quantification of the natural language terms were empirically derived and seem to produce reasonable signatures for reliable matching. Empirically, it transpires that quan-

[2] Recall that similarity is inversely proportional to matching distance

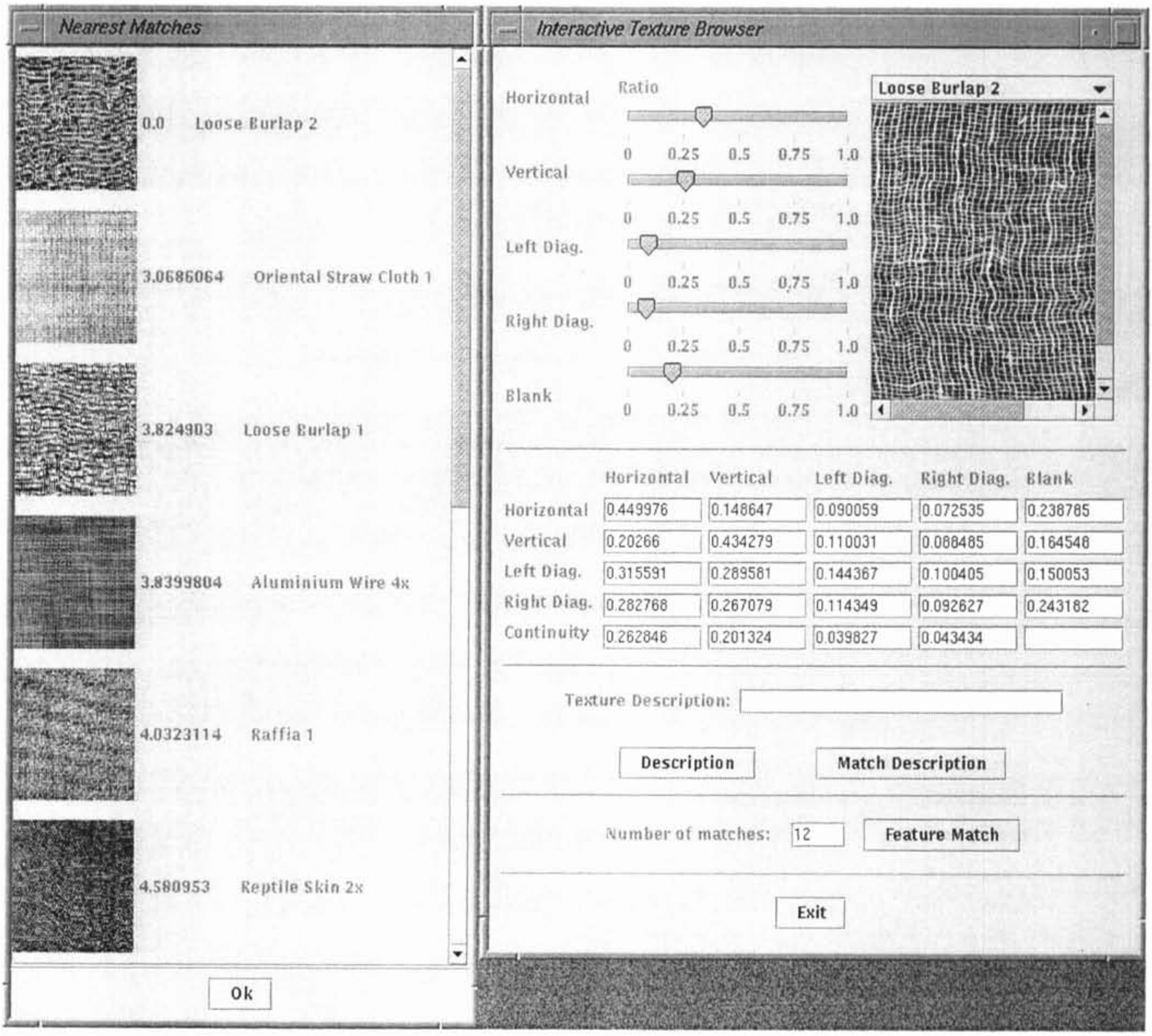

Figure 3: Similarity match on burlap pattern

tifying terms such as "many", "few", "strong" and "weak" that describe qualitative characteristics of the texture can be achieved and provides relatively robust retrieval rates from the Brodatz database. It is unlikely that such an approach is anywhere near generalisable to other content based retrieval problems, but it attempts to move texture based retrieval towards coping with what Jain called "emergent image semantics" [5], where we attempt to capture certain properties of the implicit content in *types* of image. In our case, we have proceeded empirically by finding fortuitous mappings between the primitives used in the matching algorithm and visually apparent features in the texture images.

4 Conclusion and Future Work

We described modifications from our previous texture matching technique in order to enhance visual similarity matching. The features are grouped into three domains: global, local and coarse local. The technique uses a very simple mathematical algorithm to extract the features and it is extended in a content based retrieval application, so that we can demonstrate the flexibilty for the user over the control of the query texture structure.

Although the user interface does require some efforts from the user to understand the meaning of the features and changes to values in the user interface, we have developed a simple lingusitic approach which maps a set of lexemes onto the feature set of

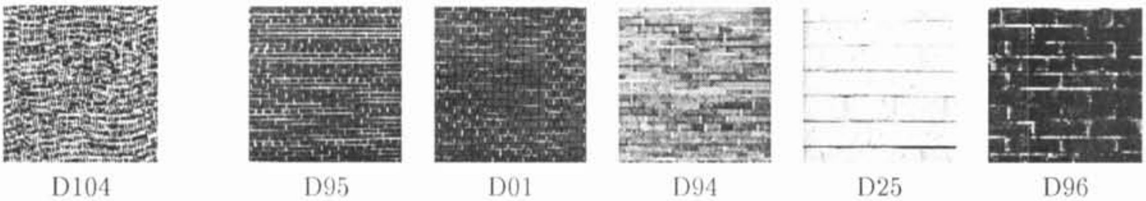

Figure 4: Similarity retrievals after the parameters changed

our technique. As a result, users will benefit from more freedom (or a more natural fashion) that uses less effort to express the query and is also easier to understand than the features. The description query basically lets the user define part of features that they are interested in and then a similarity match is performed on partial feature space. Other recent developments include building a hierarchical model of terms defined by a set of lexemes with human prior knowledge. This means that the query can be formulated at a higher level of abstraction. These developments will appear in future publications with detailed evaluations of the technique.

References

[1] P. Brodatz. *Textures: A Photographic Album for Artists & Designers*. New York: Dover, 1966.

[2] Y.Q. Chen, M.S. Nixon, and D.W. Thomas. Statistical geometrical features for texture classification. *Pattern Recognition*, 28(4):537–552, April 1995.

[3] Myron Flickner, Harpreet Sawhney, Wayne Niblack, Jonathan Ashley, Qian Huang, Byron Dom, Monika Gorkani, Jim Hafner, Denis Lee, Dragutin Petkovic, David Steele, and Peter Yanker. Query by image and video content: The QBIC system. *IEEE Computer*, 28(9):23 – 32, September 1995.

[4] Robert M. Haralick, Stanley R. Sternberg, and Xinhua Zhuang. Image analysis using mathematical morphology. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-9(4):532 – 550, July 1987.

[5] Ramesh Jain. Immersive, interactive, navigational – needed for information access. *IEEE Multimedia*, 4(4):1, October/December 1997.

[6] Joseph Kuan and Paul Lewis. Complex texture classification with edge information. In *Proc. of 2nd International Conference on Visual Information Systems*, pages 157 – 162, San Diego, CA, December 1997.

[7] Jian Chang Mao and Anil K. Jain. Texture classification and segmentation using multiresolution simultaneous autoregressive models. *Pattern Recognition*, 25(2):173 – 188, 1992.

[8] D. Patel and T. J. Stonham. Texture image classification and segmentation using RANK-order clustering. In *11th IAPR International Conference on Pattern Recognition*, pages 92 – 95, The Hague, The Netherlands, 30, August – 3, September 1992.

[9] Hideyuki Tamura, Shunji Mori, and Takashi Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. on Systems, Man, and Cybernetics*, SMC-8(6):460 – 473, June 1978.