FEATURE CLASSIFICATION BASED ON LOCAL INFORMATION

Ling Shao, Ihor Kirenko, Ping Li

Philips Research Laboratories
Prof Holstlaan 4, 5656 AA, Eindhoven, The Netherlands
phone: + (31) 402746127, fax: + (31) 402742630, email: l.shao@philips.com

ABSTRACT

Feature extraction is the first and most critical step in various vision applications. The detected features must be classified into different feature types before they can be efficiently and effectively applied on further vision tasks. In this paper, we propose a feature classification algorithm that classifies the detected regions into four types including blobs, edges and lines, textures, and texture boundaries, by using the correlations with the neighbouring regions. The effectiveness of the feature classification is evaluated on image retrieval.

1. INTRODUCTION

Good performance in many computer vision tasks often depends upon the reliable selection of a sufficient number of image regions or salient features. Many region or feature extraction algorithms have been presented in the literature. For example, Lowe [3] proposes an efficient recognition method based on local extrema of difference-of-Gaussian filters in scale-space. Mikolajczyk and Schmid [4] detect interest points using a multi-scale framework and then apply scale selection to extract characteristic points. In [5], they adapt the technique to detect affine invariant interest points. Baumberg [1] selects interesting points at multiple scales and then adapts the shape of the region to the local image structure using an iterative procedure based on the matrix of second moments. Kadir and Brady [2] proposed an algorithm to select salient regions based on the entropy of local descriptors. Shao et al. [6] further develop the salient region detection algorithm that improves the repeatability significantly.

All the above techniques extract low level features or regions, but they don't classify the extracted features into different types. Different features will be useful for different vision applications. Therefore, feature classification plays an important role after features and regions are extracted by various methods. We demonstrate the feature classification algorithm on regions detected using the technique in [6]. It can be straightforwardly applied on features or regions detected by other algorithms with little modification.

The remainder of this paper is organized as follows. Section 2 describes the feature classification algorithm, which classifies the detected features into four types. In Section 3, the classified features are applied on some image retrieval application to evaluate the usefulness of the proposed technique. We conclude our paper in Section 4.

2. FEATURE CLASSIFICATION

In [6], Shao et al. proposed a novel salient region selection algorithm that improves the repeatability substantially under different transformations, which means that a region selected in the first view of a scene is very likely also to be selected in a second view of the same scene. The regions selected by the algorithm are also intra-class invariant, i.e. the algorithm selects similar regions semantically in different objects of the same category. For example, in the category of human face images, regions corresponding to the eyes, nose tips, mouth corners and ears, are always selected. However, there are also some other regions selected which are not particularly salient for this category of images. Those regions mostly correspond to edges, lines and textures. Evidently, it would be better to classify selected regions into different feature types and remove those less salient regions and retain only those that are more useful for further applications. For example, in human face images, regions with blobs are obviously more useful than regions with edges or textures for applications such as recognition or matching, because the distinctive features in human faces are mostly contained in blob regions. Of course, for some other images edges or texture may be more useful than blobs. The salient region selection algorithm is a bottom-up method, that is, we don't have any information about the image content. Any regions considered salient by the algorithm are selected but not classified. It would be very useful to distinguish those regions selected using the low-level technique according to some high-level knowledge of the image.

The regions selected usually include blobs, edges, lines and textures, as illustrated in Figure 1. In many applications, we are more interested in regions with blobs. In the following sections, we attempt to classify regions into those different feature types.

2.1 Edges and Lines

Regions with edges and lines have a common characteristic: they both have regions along the edges or lines which resemble them in appearance. Therefore, we can model regions with edges and lines as regions which have similar regions in their neighbourhoods. We use correlation to calculate the similarity between regions.

Correlation is a bi-variant measure of association (strength) of the relationship between two variables. It varies from 0 (random relationship) to 1 (perfect linear relationship) or -1
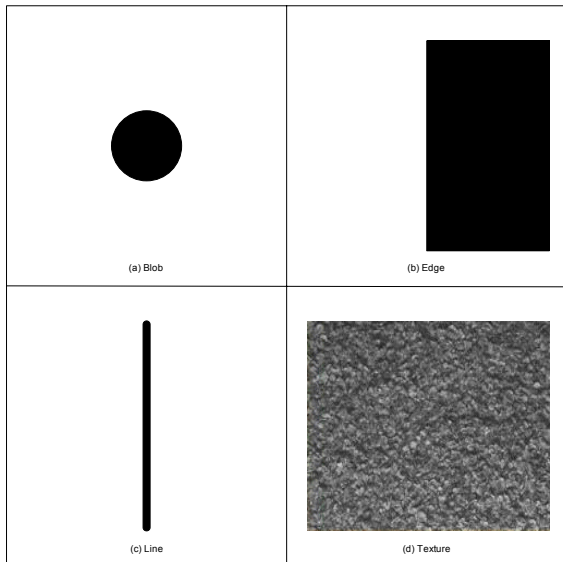
Figure 1: The idealised examples of four image feature types: Blob, Edge, Line and Texture.

(perfect negative linear relationship). The Cross Correlation between two image regions is defined as follows:

$$\mu_x = \frac{1}{N}\sum_{i \in R} I_x(i) \qquad (1)$$

$$\mu_y = \frac{1}{N}\sum_{i \in R} I_y(i) \qquad (2)$$

$$Corr(I_x, I_y) = \frac{\sum_{i \in R}(I_x(i) - \mu_x) \cdot (I_y(i) - \mu_y)}{\sqrt{\sum_{i \in R}(I_x(i) - \mu_x)^2} \cdot \sqrt{\sum_{i \in R}(I_y(i) - \mu_y)^2}} \qquad (3)$$

where I_x, I_y are the intensities of the two image regions, and R depicts the region area.

For a particular region, we compute the correlations between this region and the regions in the neighbourhood. A region is considered to be an edge or line if there is at least one correlation between that region and one region of the neighbourhood larger than the threshold. Figure 2 shows the results of removing regions with edges and lines from some human face images with different thresholds. We can see that most regions with edges and lines are removed, but other regions are retained.

2.2   Texture

The characteristic of a textured region is that its neighbourhood is homogeneous. We can model a region with texture as one whose neighbouring regions have similar intensity distributions. Two regions with similar intensity distributions should have similar histograms, hence similar entropies. We compute the standard deviation of the entropies of regions in the neighbourhood of the textured region selected as follows:

$$\mu_H = \frac{1}{N}\sum_{i \in R} H(i) \qquad (4)$$

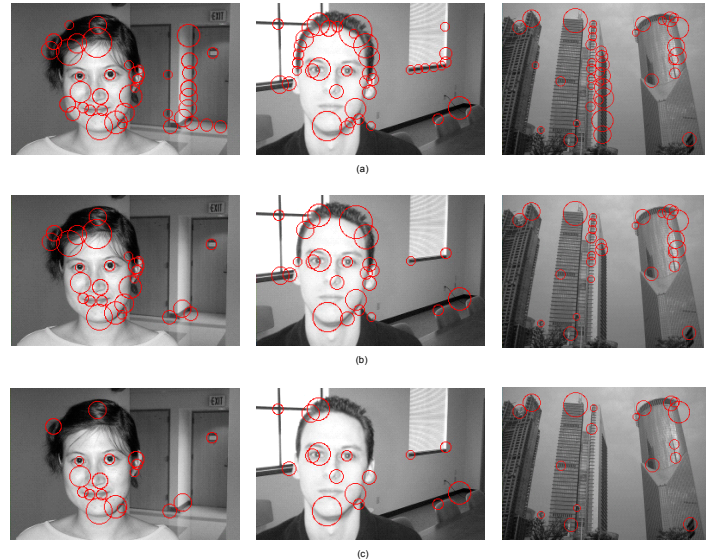$$\sigma_H = \sqrt{\frac{1}{N}\sum_{i \in R}(H(i) - \mu_H)^2} \qquad (5)$$

Figure 2: Removing regions with edges and lines: (a), Original method without removing edges and lines; (b), Threshold = 0.95; (c), Threshold = 0.85.

where $H(i)$ is the entropy of a region in the neighbourhood of the textured region, R is the neighbourhood area and N is the number of pixels in the neighbourhood. Since textured regions have homogeneous neighbourhoods, the standard deviation of entropies of such regions in the neighbourhood, σ_H, tends to be small.

Therefore, a region can be considered to be textured if the standard deviation of entropies of the neighbouring regions is smaller than a threshold. In order to observe the effectiveness of the above region classification method, we apply the method on a synthesized image that has different types of image features. We first use the model of edges and lines, then the model of textures, and the remaining regions are considered to be blobs. A result of region classification is showed in Figure 3, green circles indicate regions with edges or lines, blue circles indicate textured regions, and red circles indicate blob regions. We can see that most regions are correctly classified, with only few false positives in the boundaries of multiple features.

2.3   Texture Boundary

From Figure 3, we can see that most false positives correspond to texture boundaries, which comprise both texture and an edge. The multiple features make regions in the texture boundary hard to identify by the models of edge and texture discussed above. The characteristic of a region in the texture boundary is that part of the region is texture and there is an edge in the region. Therefore, we model a region in the texture boundary as follows: (1) it is not a pure edge or line; (2) it is not just texture; (3) it has at least one region of the same size in the neighbourhood which has a similar histogram as the region in question. The similarity measure of histograms in two regions is estimated by calculating the distance between the histograms as follows:
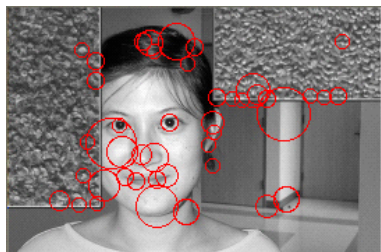
$$D_{mn} = \frac{1}{N}\sum_{i=1}^{N}\left|B_m(i) - B_n(i)\right| \qquad (6)$$
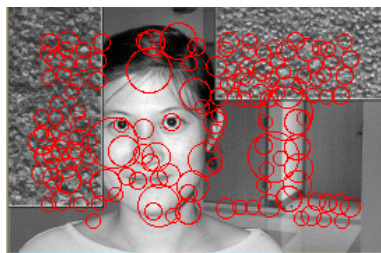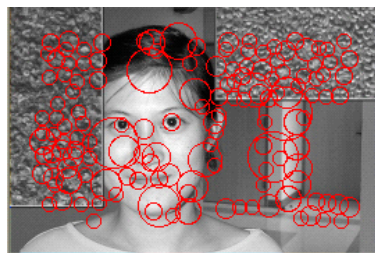
(a)
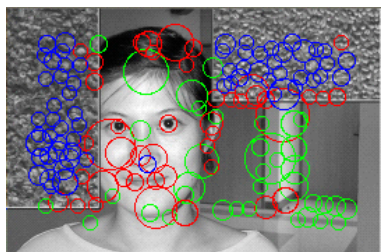
(b)

(c)

Figure 3: Feature classification: (a), 200 regions selected using the salient region detector; (b), Regions classified into 3 feature types: Red circles indicate blobs, Green circles indicate edges or lines, Blue circles indicate texture; (c), Removing region with edges, lines and texture.

where D_{mn} stands for the distance between two histograms, $B_m(i)$ and $B_n(i)$ represent the values of the ith bins in the histograms of regions m and n, N is the number of bins used in the histograms. If the histogram distance between one region and a region in its neighbourhood is smaller than some threshold, this region is considered to be texture boundary. In order to avoid classifying blob regions as texture boundaries and to make the algorithm more efficient, only regions which are tangent to the region in consideration are considered to be neighbours.

To test the effectiveness of the modelling of texture boundary, again we apply all the feature classification models on the synthesized image which contains different feature types. Figure 4 shows the result of feature classification and the result that only retains the blob regions. We observe that most regions in texture boundary are identified successfully.
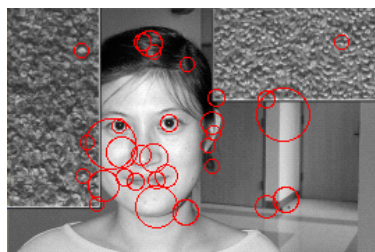
3. IMAGE RETRIEVAL EXPERIMENT

In this section, we use the above proposed feature classification algorithm on the image retrieval application. The dataset we use for experiment is the Caltech Human Face Dataset (http://www.vision.caltech.edu/html-files/archive.html),

(a)
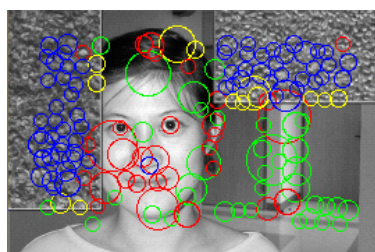
(b)

(c)

Figure 4: Feature classification: (a), 200 regions selected using the salient region detector; (b), Regions classified into 4 feature types: Red circles indicate blobs, Green circles indicate edges or lines, Blue circles indicate texture, and Yellow circles indicate texture boundaries; (c), Removing regions with edges, lines and texture and texture boundary.

which contains 435 images of different people taken against different backgrounds and with different expressions. We remove 5 single instance images from the dataset, because we need at least two images of a particular human face for retrieval experiment of a specific face. Thus, 26 sets of 430 images are used and each set contains several images of a particular human face.

Regions selected from different instances of the same object category would have distinctive appearances. For example, the eyes of different people are often distinctive. This characteristic can be used for image retrieval of a specific object from image databases. We can compute the correlations between one or more regions selected from the query image and regions selected from images in the database. If the correlations of a number of regions of the query image and an image from the database are larger than an appropriate threshold, we can conclude that the particular image contains the same object as the query image. Objects in different images would have different overlaps or occlusions, which are obstacles for recognition or retrieval. However, our method does not suffer from those problems, because only part of the regions between two images to have high correla-

tions would lead to the conclusion that they contain the same object.

The image retrieval algorithm is described as follows:

1. Regions are first selected from each image in the dataset using the region detection technique above;
2. Each region is normalised with mean and standard deviation of the pixel intensities within that region;
3. Compute the correlations between the regions of each query image and the regions of each image in the dataset;
4. The Similarity Score between each query image and each image in the dataset is computed, and true positives and false positives are obtained according to a certain threshold.

The similarity score between two images is calculated on the sum of correlations of a number of most correlated regions. In human face images, blob regions are the most distinctive. Therefore, we use the above discussed feature classification technique to remove regions of lines, edges, texture and texture boundaries, and only retain blob regions before correlations are calculated. For comparison, retrieval result without feature classification, i.e. using all the detected regions, is also obtained. The true positive and false positive results are presented in Table 1. It can be seen that the proposed feature classification significantly improves the true positives and reduces the false positives.

Table 1: Retrieval results with and without feature classification

	TP	FP
Feature Classification	80.5%	3.14%
No Feature Classification	57.2%	12.3%

4. CONCLUSION

In this paper, we presented a feature classification technique that relies on the correlations of neighbouring regions. Detected regions are classified into four feature types including blobs, edges and lines, texture and texture boundaries. In order to evaluate the effectiveness of our proposed feature classification algorithm, an image retrieval experiment is done. Results show that feature classification significantly improves the image retrieval performance.

In future work, we will test our feature classification technique on other region or feature extraction algorithms. A more general classification algorithm may be proposed afterwards.

REFERENCES

[1] A. Baumberg. Reliable feature matching across widely separated views. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Hilton Head Island, South Carolina, USA*, pages 774-781, 2000.

[2] T. Kadir and M. Brady. Scale, saliency and image description. *International Journal of Computer Vision*, 45(2):83-105, 2001.

[3] D. Lowe, Distinctive image features from scale invariant keypoints, *International Journal of Computer Vision*, 2(60):91-110, 2004.

[4] K. Mikolajczyk and C. Schmid. Indexing based on scale invariant interest points. In *Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada*, pages 525-531, 2001.

[5] K. Mikolajczyk and C. Schmid, Scale and Affine invariant interest point detectors, *International Journal of Computer Vision*, 1(60):63-86, 2004.

[6] L. Shao, T. Kadir and M. Brady. An invariant salient region detector. In *Proceedings of the Irish Machine Vision and Image Processing Conference, Belfast, UK,* 2005, pp. 105-112.