2014 International Symposium on Nonlinear Theory and its Applications
NOLTA2014, Luzern, Switzerland, September 14-18, 2014

Shape description based on projected edges and global statistical features

Attila Stubendek, Kristóf Karacs and Tamás Roska†

†Pázmány Péter Catholic University, Faculty of Information Technology and Bionics
Práter u. 50/A, 1083 Budapest, Hungary
Email: stubendek.attila@itk.ppke.hu, karacs@itk.ppke.hu, roska@ppke.hu

Abstract– A projected edge based shape descriptor extended by global features is presented along with a related learning method. We also propose a two level classification method, corresponding the two distinct feature sets. Our experimental results show that the combination of independent features leads to increased recognition robustness and speed. The core algorithms are appropriate for cellular architectures and dedicated VLSI hardware.

1. Introduction

The key to efficient shape recognition is to use an appropriate representation that compresses all important characteristics of a shape into a compact descriptor. A shape description is considered to be efficient from a recognition point of view, if

- the representation is compact
- a metric for the comparison of the feature vectors can be efficiently computed
- the representation is insensitive to minor changes and noise, and
- the description is scale and rotation invariant.

However, in case of certain shapes and tasks (e.g. recognition of arrows), orientation may also encode information, thus rotation invariance is required only up to a small degree, or a relative orientation to a predefined axes has to be detected as well.

Description of shapes can be classified to contour-based and region-based techniques. Each method extracts specific features that encompass some meaningful aspects of the information in the shape. Using only one feature thus limits the description power of the descriptor in terms of discriminative power and classification performance.[1] Combining different features include information about different essence of the shape and may increase robustness. [2][3][4]

However, compound feature vectors may provide increased complexity, and require decision method that is able to handle differences between the parts of the description.

In Section II we present our proposed compound description called Extended Projected Principal Shape Edge Distribution. In Section III a gradual classification method is presented including a limited nearest neighborhood decision. Finally in Section IV we conclude with future directions.

2. The Extended Projected Principal Shape Edge Distribution

We suggest a shape description that combines shape features in order to represent different aspects. To cover the most of the potential aspects optimally and avoid excessive redundancy, independent features are utilized.

The descriptor consists of three parts:

a) A general header including eccentricity and area fill ratio
b) A region-based feature set with histogram moments
c) A contour-based edge description employing modified Projected Principal Edge Distribution description for shapes

In the following two paragraphs we briefly summarize the terms mentioned in points a) and b), and then we introduce the PPED and expound the edge descriptor in point c).

2.1. General region-based features

The first two features, eccentricity and area ratio represent the basic outline of the shape. The smaller the eccentricity is, the closer is the shape to a circle, while shape with eccentricity value of one is a line. The area ratio is the ratio of the area occupied by the shape and the area of the minimal rectangle covering the shape.

Regional feature set consist of the first four moments of horizontal and vertical histograms of the shape (Fig. 1). Statistical moments are frequently used as shape descriptors.[1] Using more moments enables us to describe the shape more in details, but loose the general recognition ability.

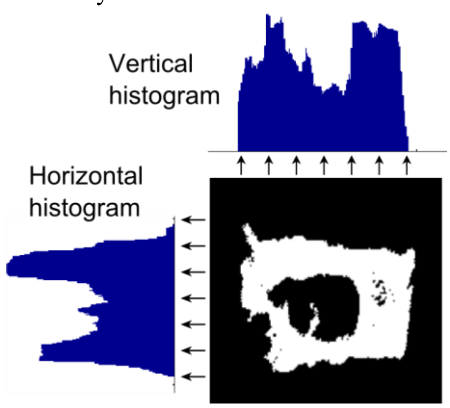

Figure 1. Vertical and horizontal histograms of a shape

2.2. Contour-based features

Projected Principal Edge Distribution (PPED) is a grayscale image descriptor that characterizes principal edges of the 64x64 pixel moving image window. To highlight important edges, for every pixel a local threshold is defined as the median of differences of neighboring pixel values. Edges are detected in four directions with a convolution. For every pixel location of the edge map only the largest edge value is kept, and values below the threshold are set to zero. [5]

The base of the contour-based edge description is the PPED that is uses as a shape descriptor with some modifications. The method is shown in Fig. 2. To achieve scale-invariant shape analysis the shape is resized to a uniform 64x64 pixel, preserving the original aspect ratio. Out measurements in previous works showed that resizing to bigger size is unnecessary. The differences between neighboring pixel gray-values in a binary image are 0 or 1 (pixel value 1 for in-shape pixels and 0 for others), consequently the median value is also 0 or 1. We experimented with different threshold values, and we concluded that the best results can be achieved by using a threshold value of 2.

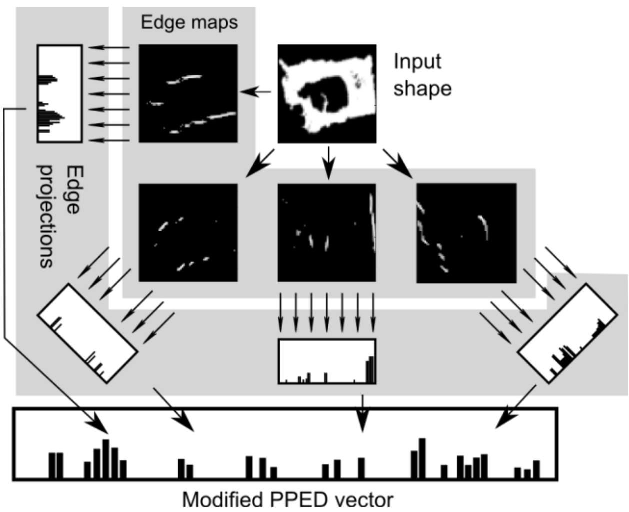

Figure 2. Construction of the contour-based part of the description. Edges are detected in four directions, then thresholding and maxima selection is applied, finally projections are concatenated and normalized.

A major deficiency of the PPED is that it is not invariant with respect to rotation. To achieve rotation invariance we chose to detect a characteristic angle and normalize the shape angularly. The orientation of the shape (defined as the orientation of the ellipse having the same second moment) serves well as a characteristic angle, since it is consistent in the sense that orientation values of similar shapes are close to each other. (Mathematical orientation may significantly deviate from the orientation value estimated by a human observer.)

Orientation provides invariance up to rotation by $k*\pi$, resulting in two distinct possibilities. To make the rotation unambiguous, shape is rotated with π if the mass center of the shape is located on the right side.

3. Classification

To demonstrate the effectiveness of the proposed description, we tested it in a shape classification task. Nearest neighborhood classifiers are typical when using PPED type descriptors.

The drawback of the nearest neighborhood method is that it might be slow (due to many comparisons), representation set scales poorly, and there is no option to classify an input as not belonging to any class. Furthermore, the EPPSED as a compound descriptor enables us a special comparison method, since the parts of the vector represents different features. Compound classifiers are frequently used techniques to handle separate parts, but generally they do not exploit the meaning of each part of the vector.

We suggest a gradual classification scheme. Shape classification is performed by comparison of the descriptor to labeled points in the feature space denoted as templates. The set of templates used for comparison is denoted as representative set. First global and statistical features are compared, then, if a satisfactory match is achieved, the final decision is computed from the differences between the contour features.

3.1. Filtering

The first phase of the decision selects candidates for the second phase of the decision and rejects obviously dissimilar template vectors. For every feature vector element a threshold is determined. An input descriptor matches the labeled template vector if the number of elements with difference higher than the threshold is within certain limit. Threshold and limit is based on preliminary measures.

3.2. Limited nearest neighborhood

The second phase of the classification employs limited nearest neighborhood decision to define the final class. The disadvantage of the nearest neighborhood decision is its inability to reject distant inputs that do not belong to any class (denoted as zero-class elements). Without specifying additional constraints there is always a nearest element to every input vector, even if the distance is high. We propose a method that selects the nearest template in the Euclidean space, but it is accepted only if the distance is smaller than the minimal acceptance threshold of the template.

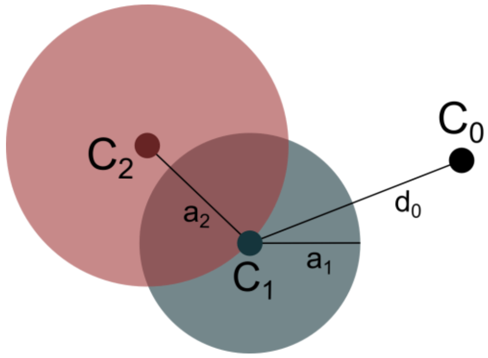

Figure 3. Automatic definition of minimal acceptance range. C_1 and C_2 represent two in-class elements, C_0 a zero-class element. Acceptance threshold for C_1 will be d_0 as the half of the distance to the closest zero-class element. Acceptance threshold for C_2 will be a_2 as the distance to the closest element of another class.

To determine the minimal acceptance threshold we used an automatic generation algorithm. We employed a shape database as a train set consisting of various shapes consisting also elements that do not belong to any class. The acceptance threshold for a template is determined based on the other elements that match the template by the filtering described above. We have set the threshold to be half of the distance to the nearest zero-class element. In case of the absence of any zero-class elements, the distance to the nearest element labeled with different class is used as the acceptance threshold. If no element with a different class label is in the filter range, the acceptance threshold is set as the distance to the farthest element within the same class.

3.2. Experimental results

We tested the descriptor and the classification method on a shape set extracted from Hungarian Forint banknotes in the Bionic Eyeglass Project. [6] Table 1. and Table 2. summarize the results.

The two shape datasets contain 6172 and 7024 shapes images respectively, including negative elements that do not belong to any class. The representative set had 271 elements.

Shapes in the test sets represent characteristic graphical patches (portraits, drawings) extracted from banknotes, but also shadows, joined patterns and other patches from the background. Nonzero-class shapes represented 9 classes with high variation due to morphologic extraction. Input images are shown on Figure 4.

Global accuracy is the ratio of correctly recognized shapes including zero-class inputs. Cover corresponds for the ratio of correctly identified and all nonzero-class inputs. Precision is the ratio of correctly classified nonzero-class inputs and all classified nonzero inputs. Average lookup time was measured on a standard computer (Core2 Quad CPU @ 2.66 GHz, 3 GB memory).

	Global accuracy	Cover	Prec.	Av.lu. time (ms)
W/out Filters	84,4%	45%	98,3%	57,2
With Filters	88,5%	58,3%	99,4%	6,4

Table 1. Experimental results of test set 1.

	Global accuracy	Cover	Prec.	Av.lu. time (ms)
W/out Filters	88,5%	46%	99,6%	58,0
With Filters	91,5%	60,1%	99,7%	6,4

Table 2. Experimental results of test set 2.

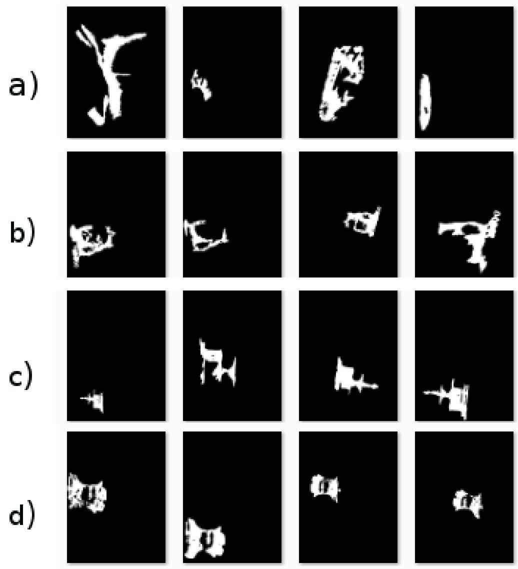

Figure 4. Fragment on the test sets. In the row a) are zero-class shapes, rows b)-d) show nonzero-class shapes.

4. Conclusion

We presented a new shape description and classification method. Key characteristics of our approach are the compound descriptor and classifier that join the region and contour-based features.

Results show high precision and lower cover. The reason to have lower cover is that the test sets consist of also highly deformed shapes, which were classified as non-zero elements.

The computation time allows real-time recognition on standard CPUs, and the architecture core of the algorithm is appropriate on cellular architectures. In the future our plan is to implement the descriptor on CNN architecture.[7]

Acknowledgement

The support of the Swiss Contribution is gratefully acknowledged.

References

[1] D. Zhang, G. Lu, “Review of shape representation and description techniques”, *Pattern Recognition* 37, pp.1 – 19, 2004

[2] J. Iivarinen, M. Peura, J. Srel, A. Visa, “Comparison of Combined Shape Descriptors for Irregular Objects, *8th British Machine Vision Conference,* 1997 http://www.cis.hut.fi/research/IA/paper/publications/bmvc97/bmvc97.html

[3] M. Hasegawa, S. Tabbone: “A Shape Descriptor Combining Logarithmic-scale Histogram of Radon Transform and Phase-only Correlation Function”, *International Conference on Document Analysis and Recognition*, Beijing, pp.182 – 186, 2011

[4] S. Khanam, S. Jang, W. Paik, “Shape Retrieval Combining Interior and Contour Descriptors”, *International Conference FGCN*, Jeju Island, pp.120-128, 2011

[5] M. Yagi, T. Shibata, “An Image Representation Algorithm Compatible with Neural-Associative-Processor-Based Hardware Recognition Systems,” *IEEE Trans. Neural Networks, Vol. 14*, No. 5, pp. 1144-1161, September, 2003

[6] Z. Solymar, A. Stubendek, M. Radvanyi, K. Karacs, “Banknote Recognition for Visually Impaired”, *Proc. of the European Conference on Circuit Theory and Design (ECCTD’11)*, Linköping, Sweden, Aug 2011.

[7] T. Roska and L. O. Chua, “The CNN universal machine: an analogic array computer,” *IEEE Trans. Circuits Syst. II, vol. 40,* pp. 163–173, Mar. 1993.