International Journal of Engineering and Applied Sciences (IJEAS)
ISSN: 2394-3661, Volume-4, Issue-11, November 2017

Fourier Descriptor Based Static Sign Language Recognition

Yong HU, Wei-yan SHEN, Wei ZHENG

Abstract—**Sign language recognition is an important and challenging research area in the field of human-computer interaction. The work presented in this paper aims at developing a robust system for automatic translation of static gestures of alphabets. The proposed methods used the Support Vector Machine classifier and features extracted from a novel signature, Inner-Centroid Distance. In the experiments, a comparison between the proposed methods and Zernike moments is performed by using a public sign database. Experiments show that the proposed approach is effective and simple to implement.**

Index Terms—**Sign Language Recognition, Fourier Descriptor, Inner-Centroid Distance, Zernike moments, Support Vector Machine**

I. INTRODUCTION

Sign language (SL) recognition began to appear in the 1990s and has become one of the most important research areas in the field of human-computer interaction. It aims at providing an efficient and accurate mechanism to translate sign language into text or speech. There are many existing surveys within the area of Sign language recognition [1]. In general, gestures can be divided into two groups, static gestures (hand postures) and dynamic ones. A static sign is determined by a certain configuration of the hand, while a dynamic gesture is a moving gesture determined by a sequence of hand movements and configurations. The aim of sign language alphabets recognition is to provide an easy, efficient and accurate mechanism to transform sign language into text or speech [2].

There are two main categories in sign language recognition, one is using electromechanical devices (data gloves) and the other is visual approach. Instrumented glove approaches simplified the recognition by measuring the different gesture parameters such as hand's position, angle, and the location of the fingertips. But it is expensive, less user friendly and complicates the hardware. The vision based approaches only use machine vision and image processing techniques to create visual based hand gesture recognition systems. The system deals with images of bare hands, which allows the user to interact in a natural way. It is widely used because of its characteristic of user friendly and affordable.

M. A. Amin et al. [3] developed a recognition system for all 26 alphabets of American SL alphabet gestures. Gabor filters and Principle Component Analysis (PCA) were used for recognition after some required image pre-processing. Their system had achieved an average accuracy of 93.23%.

Q. Munib et al. [4] proposed a developing system for automatic translation of static gestures of American alphabets by using Hough transform and neural networks. The system was implemented and tested using a data set of 300 samples of hand sign images. Experiments revealed that their system was able to recognize selected ASL signs with an accuracy of 92.3%. J. Ravikiran et al. [5] introduced an efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet of the American SL. Finger detection is accomplished based on the concept of boundary tracing and fingertip detection. D. Kelly et al. [6] present a user independent framework based on Support Vector Machine (SVM). An eigenspace size function and Hu moments features were used to classify different hand postures. Experiments based on two different hand posture data sets show the robustness of their approach. F. H. CHOU et al. [7] presented a GMM-based (Gaussian Mixture Model) processing algorithm for the gesture images detection and recognition. Based on their presented algorithm, the correct recognition rate is about 94% in average. K. C. Otiniano-Rodrıguez et al. [8] proposed two methods for Sign Language Recognition using the SVM classifier and features extracted from Hu and Zernike Moments. In the experiments, the proposed methods are performed using a database composed of 2040 images for recognition of 24 symbol classes. Promising results were obtained by comparing the proposed method with the Hu moments. P. Rajathi and S. Jothilakshmi [9] proposed a three stage system to recognize static gestures representing Tamil words. The proposed sign language recognition system is able to recognize images with 91% accuracy.

In this work, a vision-based approach is proposed to recognize static signs of American SL alphabets. The recognition system is based on the Inner-Centroid Distance signature and Support Vector Machine classifier. Firstly, each sign image was processed as preprocessing, which include resize, gray conversion and filtering. And then, a boundary tracing method and Fourier descriptor were implemented to extract features from the sign images. In the last step, a SVM classifier was utilized as pattern classifier in this work, and the implemented results were compared with the other feature extraction method for promising the robustness and effectiveness.

This paper is organized as follows. In Section 2, the materials and methods used in this paper were briefly introduced. In Section 3, the proposed framework for sign language recognition includes preprocessing, feature extraction and classification is presented in detail. Experimental results are reported in section 4. Finally, the conclusion is provided in Section 5.

Yong HU, School of Software Engineering, Jinling Institute of Technology, Nanjing, China, +86-25-86188709
Wei-yan SHEN, School of Software Engineering, Jinling Institute of Technology, Nanjing, China, +86-25-86188709
Wei ZHENG, School of Computer Engineering, Jinling Institute of Technology, Nanjing, China, +86-25-86188709

II. THE PROPOSED RECOGNITION ALGORITHM

A. *System Architecture*

Figure 1 illustrates the architecture of the proposed sign language recognition system, which is composed of three steps. The first stage is the pre-processing stage, that including hand segmentation, gray conversion and filtering. The second stage is the feature extraction stage; the segmented binary image is used for extracting the Zernike moments and the Fourier descriptor (Inner-Centroid Distance signature). In the training/classification stage, the combined feature vectors are used as the input of the SVM classifier.

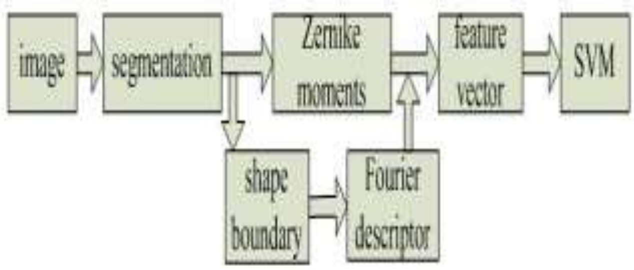


Figure 1: Architecture of the proposed system

B. *Inner-Centroid Distance (ICDs)*

In this section, we introduce a novel signature, the Inner-Centroid Distance (ICDs) signature [10] which is based on the Centroid Distance (CD) signature and Inner-Distance [11]. The ICDs is developed to overcome some of the shortcomings of existing signatures, such as ignoring articulation and part structures of complex shapes. In the Centroid Distance signature, only the distances of the boundary points from the centroid of the shape are concerned and thus the features extracted from CDs cannot characterize the fundamental properties of the complex shape boundary. The inner-distance is articulation insensitive and more effective at capturing complex shapes with part structures and is proved as a natural replacement for the Euclidean distance in shape descriptors. So, inner-distance is used here to extend the Centroid Distance (CD) signature for image retrieval.

The inner-distance is defined as the length of the shortest path between landmark points within the shape silhouette in [11]. Consider two points $x, y \in O$, where O is a shape defined as a connected and closed subset of $\boldsymbol{R}^2$. The inner-distance between $\boldsymbol{x}$ and $\boldsymbol{y}$, denoted as $d(x, y; O)$, is defined as the length of the shortest path connecting $\boldsymbol{x}$ and $\boldsymbol{y}$ within O. If shape O is convex or the line segment connecting $\boldsymbol{x}$ and $\boldsymbol{y}$ falls entirely within the silhouette, the inner-distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ reduces to the Euclidean distance.

Figure 2 depicts how the Inner-Centroid distance is calculated. For point A, part of the line segment AC is out of the shape silhouette. Thus the inner-distance between A and C (Centroid point) is calculated as the length of the shortest path within O, which is denoted by dashed line. For point B, the Inner-Centroid distance is the Euclidean distance between C and B. It indicates that the inner-distance is influenced by part structure and characterizes some fundamental properties of the complex shape boundary.

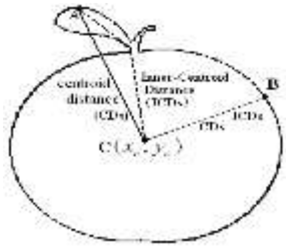


Figure 2: Basic concept of the ICDs signature

C. *Support Vector Machine*

Support vector machine (SVM) is an effective classification technique on the foundation of statistical learning theory. Training SVM is equivalent to solving a linearly constrained quadratic programming problem so that the solution of the SVM is always unique and globally optimal.

Consider the two-class linear separable training sets

$$(x_i, y_i),\ i = 1,2,\cdots,n,\ x \in R^d,\ y \in \{-1,+1\} \qquad (1)$$

If there is a hyperplane $\omega \bullet x + b = 0$, all samples can be correctly classified without error and the separation margin can be maximized, then the hyperplane is so-called optimal hyperplane.

The optimal separating hyperplane is computed as a decision surface of the form:

$$\mathrm{f(x)} = \mathrm{sgn}\{\sum_{i=1}^{m} \alpha_i^* y_i (x_i \cdot x) + b^*\} \qquad (2)$$

In this study, we perform LIBSVM for multi-class classification. The LIBSVM is a library for support vector machines which is available in [12]. By using the software, a cross-validation method was utilized to analyze the classification performance. All labeled samples were split into two subsets: 8 subjects are randomly selected for training and all others for testing. The trained SVM classifier predicts the action label for the testing actions. This procedure was repeated ten times, and the average results are listed in the Table 1.

III. EXPERIMENTAL RESULTS

For implementation, image set used in this work was obtained from the Moeslund gesture database. The database consists of 2040 images of 248×256 pixels that represent 24 static signs in gray scale. The signs A to F have 40 images for each class, while the signs G to Y have 100 images for each class. The sign J and Z are not used because these two signs both have motion included. Figure 3 shows some examples of database (one example for each sign).

In order to make the system simpler and faster, sign images are resized after segmentation to binary images. Their size is reduced to a resolution of 80×100 pixels in order to decrease computational cost. The calculation of Zernike moments and Fourier descriptor are based on small images. Figure 4 shows the pre-processing stage adopted in this work.

In segmentation step, a fast image segmentation method based on histogram information was used for each sign image. And then, a small image and its boundary were obtained after image resized. As for the order of the Zernike moments, we evaluated several orders (from 3 to 6) to find the best one for classification. Order 6 performs better in the experiments. For the Fourier descriptor features, the number of features used in our implementation is 15. In the training/classification step, a cross validation were perform by using support vector machine.

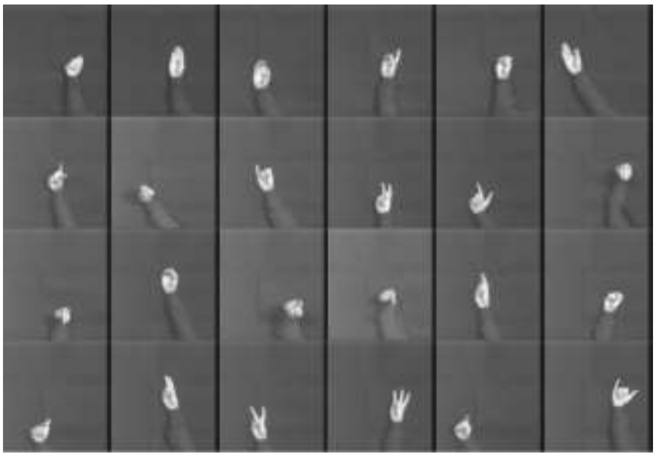
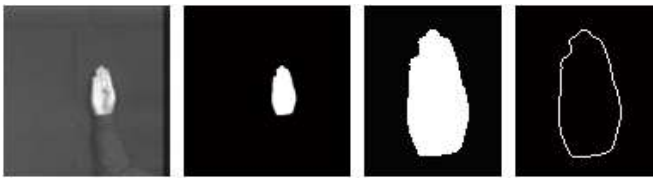


Figure 3: Static gestures from Moeslund Database (24 signs)



(a) alphabet b (b) segmented (c) resize (d) boundary

Figure 4: Pre-processing of the gesture images

Table 1 summarizes the precision and recall results obtained by the simulation process. We can see from the Table that the proposed signature outperforms the competing techniques. The experiment results indicate that the proposed method is suitable for sign language recognition.

Table 1. Precision and recall results of this work

	Zernike Moments	Zernike + ICDs
Precision	0.953	0.963
Recall	0.965	0.980

IV. Summary and Conclusions

In this work, we presented a robust system that is able to automatic translate the static gestures of alphabets. The proposed methods used the support vector machine classifier and features extracted from a novel signature, Inner-Centroid Distance. In the experiments, a comparison between the proposed methods and Zernike moments is performed by using a public sign database. Experiments show that the proposed approach is effective and simple to implement. The work done in this paper is actually a preliminary work for sign language recognition. Future research is needed to solve the confusing letters and the isolated signs with motions.

Acknowledgment

This work is supported by Top-notch Academic Programs Project of Jiangsu Higher Education Institutions (TAPP).

References

[1] Murthy G. R. S. and Jadon R. S., A review of vision based hand gestures recognition, International Journal of Information Technology and Knowledge Management, 2(2), pp. 405-410, 2009

[2] Kulkarni A. V. S., Lokhande A., Appearance Based Recognition of American Sign Language Using Gesture Segmentation, International Journal on Computer Science and Engineering, 2(3), pp. 560-565, 2010

[3] Amin M. A., Yan H., Sign Language Finger Alphabet Recognition from Gabor-PCA Representation of Hand Gestures, International Conference on Machine Learning and Cybernetics, IEEE, pp. 2218-2223, 2007

[4] Munib Q., Habeeb M., Takruri B., Al-Malik H.A., American sign language (ASL) recognition based on Hough transform and neural networks, Expert Systems with Applications, 32(1), pp. 24-37, 2007

[5] Ravikiran J., Mahesh K., Mahishi S., Finger Detection for Sign Language Recognition, Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2009), Hong Kong, IEEE, pp. 1-5, 2009

[6] Kelly D., McDonald J., Markham C., A person independent system for recognition of hand postures used in sign language, Pattern Recognition Letters, 31(1), pp. 1359-1368, 2010

[7] Chou F. H., Su Y. C., An Encoding and Identification Approach for the Static Sign Language Recognition, The International Conference on Advanced Intelligent Mechatronics, Taiwan, pp: 885-889, 2012

[8] Otiniano-Rodrıguez K. C., Cámara-Chávez G., Menotti D., Hu and Zernike Moments for Sign Language Recognition, International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'12), pp. 1-5, 2012

[9] Rajathi P., Jothilakshmi S., A Static Tamil Sign Language Recognition System, International Journal of Advanced Research in Computer and Communication Engineering, 2(4), pp. 1885-1891, 2013

[10] Hu Y., Li Zuoyong, An Improved Shape Signature for Shape Representation and Image Retrieval, Journal of Software, 8(11), pp. 2925-2929, 2013

[11] Ling Haibin, David W. Jacobs, Shape Classification Using the Inner-Distance, IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(2), pp. 286-299, 2007

[12] Chang C.C. and Lin C.J.: LIBSVM: a library for support vector machines. (2001). Software available at: http://www.csie.ntu.edu.tw/~cjlin/libsvm

YONG HU received his Ph.D. from school of Computer Science & Technology at Nanjing University of Science and Technology (NUST) in 2010. He is currently an Associate Professor at Jinling Institute of Technology (JIT). His main research interests include image processing, pattern recognition and machine learning.

Wei-yan SHEN received her Master's Degree from school of Computer Science & Technology at Nanjing University of Science and Technology (NUST) in 2007. She is currently a lecturer at Jinling Institute of Technology (JIT). Her main research interests include image processing, pattern recognition and machine learning.

Wei ZHENG received her Master's Degree from school of Computer Science & Technology at University of Electronic Science and Technology of China (UESTC) in 2007. She is currently a lecturer at Jinling Institute of Technology (JIT). Her main research interests include image processing, pattern recognition and machine learning.