International Conference on Internet Studies,
July 18-19, 2015, Tokyo, Japan

A NOVEL OCCLUSION SIGN LANGUAGE RECOGNITION

Ching-Tang Hsieh,
Tamkang University, Taiwan
hsieh@ee.tku.edu.tw

Hsing-Che Liou,
Tamkang University, Taiwan
bobbysky654@yahoo.com.tw

Li-Ming Chen
Tamkang University, Taiwan
ms0071799@hotmail.com

Ting-Wen Chen
Tamkang University, Taiwan
alucardoom@hotmail.com

ABSTRACT

Sign language plays an important role in communicate with changers that hearing improved. However, the sign language in many countries and areas different and auto recognition system became the research way in recent year. In this paper, we devise a novel method for occlusion processing in Taiwan Sign Language recognition system. Our method employs adxl345 and Kinect to extract the feature of signer. Then the features are regulated by the dictionary of sparse coding. In final, the HMM model and result signs are recognized from the features that corrected by our method. In experimental result, we present the data that our employ. Then we describe closing test result and future work.

Keyword: Taiwan Sign Language (TSL), Sparse Coding, Occlusion

1. INTRODUCTION

Sign language is a body language that plays an important role in communicating with hearing changers. As speaking language, there are many different sign languages in every country. In Taiwan, the Taiwan Sign Language (TSL) is also used to communication with hearing changers. However, in our lives, the people who knowing sign language are not generally, and only some governments and hospitals have recognition system to transform sign language. How to fusion the sign language recognition system in people lives is one way for research in recent years. The auto sign language recognition system has image-based and sensor-based two main approaches [1]. In sensor-based systems, signers equip instrumented gloves with sensor to extract features that training model and recognizing signs. However, sensor-based systems have challenges that as uncomfortable equipment by the signer. In image-based systems, signers do not equip any device. Nevertheless, image-based systems need substantial computations in the preprocessing stage.

In related works, researchers proposed and discussed their scheme to recognize sign language by features that extracted. A classical image-based system has five processing step: image obtained, pre-processing, segmentation, features extracted and classify [2]. In addition the five steps, more problems as background, illumination, the segment of face and other noises. Wu et al. [3] proposed the image-based gesture recognition and discussed the applications. Moni et al. [4] employed color grove and HMM to recognize gesture, they commented different method to segment gesture sequence. In [5], the algorithms that sign language recognition are divided static and dynamic two classes by Kausar et al. A neuro-fuzzy system for sign language recognition was developed by Al-Jarrah et al. in [6]. Their system includes image obtained, filtering, segmentation, detecting the contour of hands, and feature extracted five steps. The accuracy is 93.6% in bare hands experiment. Al-Rousan et al. [7] used the scheme that reduced segmentation by color glove and the feature was extracted from hands area to construct an adaptive neuro-fuzzy inference system for alphabet sign recognition. And this scheme has 95.5% correct rate. In [8], six colors of gloves that five for fingertips and one for the wrist region were utilized to extract features. Then the sign of word was recognized by polynomial classifier. In [9], Al-Jarrah et al. designed a image-based system that does not use visual markings. They extracted a set of features that are translation, rotation, and scaling invariant from the image of bare hands processed. The 97.5% accuracy is obtained from 30 Arabic alphabet signs experiment in database. Maraqa et al. in [10] used recurrent neural networks for alphabet signs recognition. In their experiment, 900 samples where have 30 signs database by two signers and color glove as [8] are utilized. In particular, the 30 features are extracted from color glove, and obtained 95% correct rate in their experimental result [11]. In [12], a bare hand image of ArSL recognition system that has 91.3% accuracy was designed by El-Bendary et al. They assumed a small pause between two signs in feature segmentation step. At recognition step, a multilayer perceptron neural network and a minimum distance classifier were utilized. Hemayed et al. [13] discussed a sign language recognition system that transforms sign to speech. In [14], an alphabet signs recognition system based on image was developed by Naoum et al. Their system accuracy in bare hands, red, black and white glove experiments is 50%, 75%, 65% and 80% respectively. This system extracted contour from image histogram, then k-nearest neighbor algorithm was employed to classify it. A pulse-coupled neural network (PCNN) ArSL recognition system was proposed by Elons et al. [15], [16]. The system can able to compensate for background brightness and lighting non-homogeneity. Under geometrical transforms, bright background, and lighting conditions, their system showed invariance and achieving 90% accuracy in recognition result. In addition to traditional image-based system, new human-machine interaction systems have been introduced lately. In particular, Microsoft Kinect has attracted special attention. The Kinect has recently been used for action recognition with application in human-machine interaction [17].

In TSL, some of two hands signs that as Figure 1 are overlap or cross to mean something, and sometime it makes recognition mistake in image-based system [18]. In this paper, we present a sparse coding method for this problem which called occlusion in image processing.

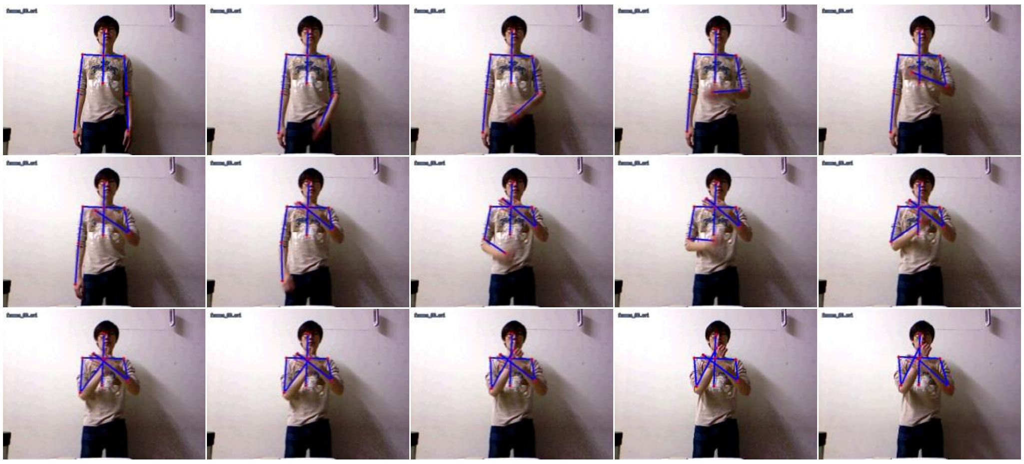


Figure 1. The frames of occlusion sign “only” in TSL

2. PROPOSED METHOD

In this section, we present our sign language recognition system with occlusion processing method. Since some of signs that have wrist rotating or not means difference in TSL [18], our system has not only Kinect but also adxl345 that assumed smart watch two sensors to extract training and recognizing features. The processing flow-chart as Figure 2 and Figure 3 shown.

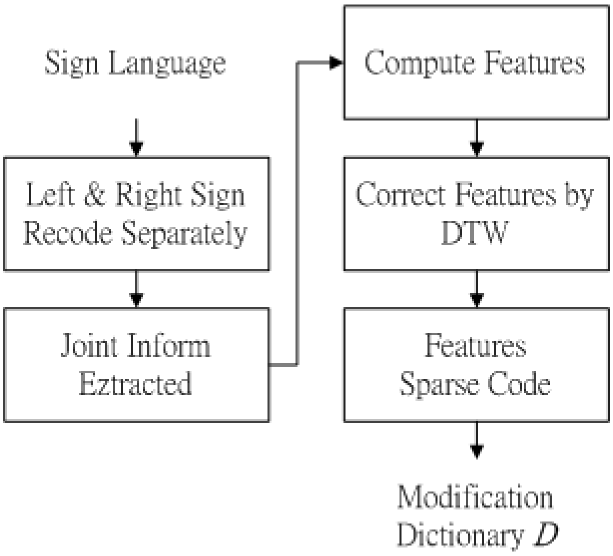


Figure 2. The flow-chart of modification dictionary $\boldsymbol{D}$ training

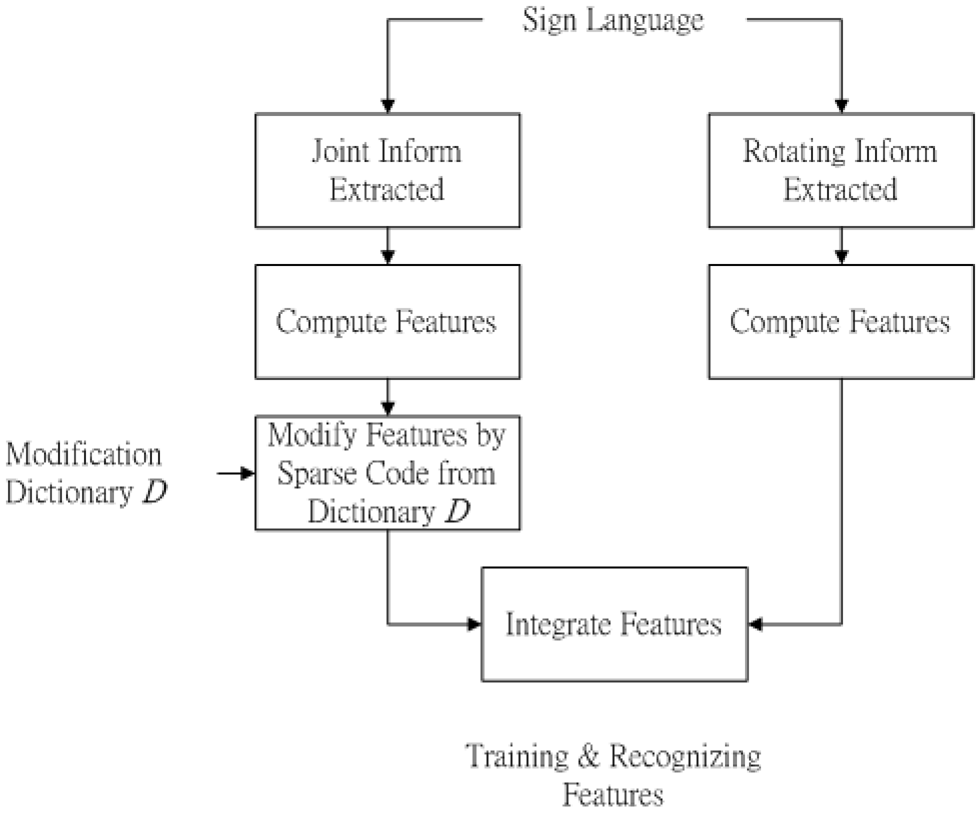
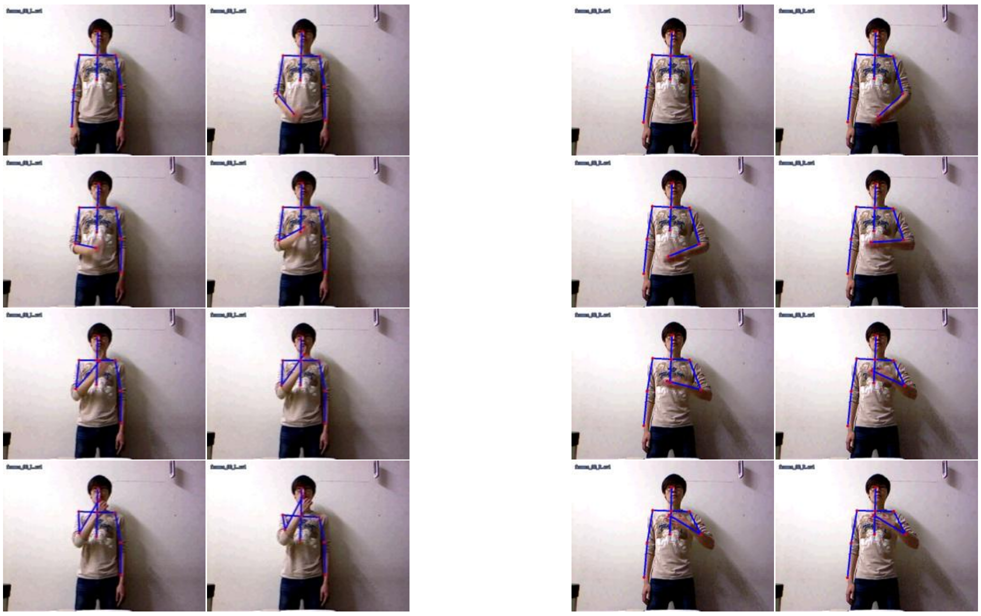


Figure 3. The flow-chart of proposed system



(a) Left hand sign (b) Right hand sign

Figure 4. Respective sign from left and right hand

For occlusion problem, we employ sparse coding [17] to train a modification dictionary $\boldsymbol{D}$. For solving occlusion, the sign of left and right hand sequences $\boldsymbol{S}_L$ and

$\mathbf{S}_R$ are collected individual as Figure 4. Next, we extract the features $\mathbf{Y}_L$ and $\mathbf{Y}_R$ from first order differential $\mathbf{S}_L$ and $\mathbf{S}_R$. And then the frames of features are corrected through dynamic time warping (DTW) [19], and concatenate features $\mathbf{Y} = [\mathbf{Y}_L, \mathbf{Y}_R]$. The dictionary $\boldsymbol{D}$ is learned by sparse coding. The flow-chart shows as Fig.2.

In model training and recognizing, the features of double hands signs are collected as normal action. In Figure 1, the sequences $\boldsymbol{S}$ are collected by Kinect as Figure 3 from left way. After first order differential makes $\boldsymbol{S}$ to features $\mathbf{Y}'$, our method employs dictionary $\boldsymbol{D}$ which we trained to simulate the features of occlusion signs in real. At the same time, the rotating features are extracted from adxl sensor. As left way to first order differential, the rotating features are integrated with joints features that were extracted and modified by Kinect and dictionary $\boldsymbol{D}$. Finally, we use these features training the HMM model and recognizing signs.

3. EXPERIMENTAL RESULT

In this study, we employ 45 signs and 5 signers to test our proposed system. And each sign is collected 10 times per signer. For comparing the features that obtained from Kinect and adxl, the 45 signs that have 30 normal and 15 wrist rotating signs were selected. Some of sings have occlusion when the action of signers. For simulating smart watch to extract rotating features, the adxl sensor as Figure 5 shown is worn on the wrist. In closing test, the high accuracy is obtained in dependent test by our method.

Figure 5. Simulating smart watch equipment by adxl sensor

4. CONCLUSION AND FUTURE WORK

In this paper, we proposed a novel method for sign language recognition to solve occlusion problem when the actions of signers. And in closing test, the sign can be recognized accurately. In future work, we would test our method in independent test, add the words to model and compare with other methods. Then the human-machine interface and changer translator are implemented next work.

Acknowledgments: This work was supported by the National Science Council under grant number MOST-103-2632-E-032-001-MY3 and MOST 103-2410-H-032 -052.

5. REFERENCES

[1]. M. Mohandes, M. Deriche, and J. Liu. Image-Based and Sensor-Based Approaches to Arabic Sign Language Recognition. *IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS*, VOL. 44, NO. 4, AUGUST 2014.

[2]. R. Azad, B. Azad, and I. T. Kazerooni, "Real-time and robust method for hand gesture recognition system based on cross-correlation coefficient," *Adv. Comput. Sci., Int. J.*, vol. 2, no. 5/6, pp. 121–125, Nov. 2013.

[3]. Y. Wu and T.S. Huang, “Vision-based gesture recognition: A review,” Lecture Notes Comput. Sci., vol. 1739, pp. 103–115, 1999.

[4]. M. A. Moni and A. B. M. S. Ali, “HMM based hand gesture recognition: A review on techniques and approaches,” in Proc. 2nd IEEE Int. Conf. Comput. Sci. Inform. Technol., 2009, pp. 433–437.

[5]. S. Kausar and M.Y. Javed, “A survey on sign language recognition,” in Proc. Frontiers Inform. Technol., 2011, pp. 95–98.

[6]. O. Al-Jarrah and A. Halawani, “Recognition of gestures in Arabic sign language using neuro-fuzzy systems,” Artif. Intell., vol. 133, no. 1–2, pp. 117–138, 2001.

[7]. M. Al-Rousan and M. Hussain, “Automatic recognition of Arabic sign language finger spelling,” Int. J. Comput. Appl., vol. 8, pp. 80–88, 2001.

[8]. K. Assaleh and M. Al-Rousan, “Recognition of Arabic sign language alphabet using polynomial classifiers,” EURASIP J. Appl. Signal Process., vol. 2005, pp. 2136–2145, 2005.

[9]. O. Al-Jarrah and F. A. Al-Omari, “Improving gesture recognition in the Arabic sign language using texture analysis,” Appl. Artif. Intell., vol. 21, no. 1, pp. 11–33, 2007.

[10]. M. Maraqa and R. Abu-Zaiter, “Recognition of Arabic sign language (ArSL) using recurrent neural networks,” in Proc. 1st Int. Conf. Appl. Digit. Inform. Web Technol., 2008, pp. 478–481.

[11]. M. Maraqa, F. Al-Zboun, M. Dhyabat, and R. Abu Zitar, “Recognition of Arabic sign language (ArSL) using recurrent neural networks,” J. Intell. Learn. Syst. Appl., vol. 4, pp. 41–52, 2012.

[12]. N. El-Bendary, H. M. Zawbaa, M. S. Daoud, A. ella Hassanien, and K. Nakamatsu, “ArSLAT: Arabic sign language alphabets translator,” in Int. Conf. Comput. Inf. Syst. Ind. Manag. Appl., 2010, pp. 590–595.

[13]. E. E. Hemayed and A. S. Hassanien, “Edge-based recognizer for Arabic sign language alphabet (ArS2V-Arabic sign to voice),” in Proc. Int. Comput. Eng. Conf., 2010, pp. 121–127.

[14]. R. Naoum, H. H. Owaied, and S. Joudeh, “Development of a new Arabic sign language recognition using k-nearest neighbor algorithm,” J. Emerging Trends Comput. Inf. Sci., vol. 3, no. 8, pp. 1173–1178, Aug. 2012.

[15]. A. S. Elons, M. Abull-ela, and M. F. Tolba, “Neutralizing lighting non-homogeneity and back-ground size in PCNN image signature for Arabic sign language recognition,” Neural Comput. Appl., vol. 22, no. 1, pp. 47–53, May 2013.

[16]. A. S. Elons, M. Abull-Ela, and M. F. Tolba, “Pulse-coupled neural network feature generation model for Arabic sign language recognition,” IET Image Process., vol. 7, no. 9, pp. 829–836, 2013.

[17].S. R. Fanello, I. Gori, G. Metta, “Keep It Simple And Sparse: Real-Time Action Recognition,” Journal of Machine Learning Research, vol.14 pp. 2617-2640, 2013.
[18].Tsay, Jane, James H.-Y. Tai and Yijun Chen. 2015.Taiwan Sign Language Online Dictionary. 3rd Edition. Institute of Linguistics, National Chung Cheng University, Taiwan. http://tsl.ccu.edu.tw/web/browser.htm.
[19].E. Kostadinova, V. Boeva, L. Boneva, “An Integrative DTW-based Imputation Method for Gene Expression Time Series Data,” in Proc. 6th IEEE International Conference Intelligent Systems (IS), pp. 258 - 263, 2012.