Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 60 (2015) 371 – 380

Procedia Computer Science

19th International Conference on Knowledge Based and Intelligent Information and Engineering Systems

Depth Sensor Based Automatic Hand Region Extraction by Using Time-Series Curve and Its Application to Japanese Finger-spelled Sign Language Recognition

Katsufumi Inoue[a,*], Takami Shiraishi[a], Michifumi Yoshioka[a], Hidekazu Yanagimoto[a]

[a]*Graduate School of Engineering, Osaka Prefecture University, 1-1, Gakuencho, Naka, Sakai, Osaka, 599-8531 Japan*

Abstract

Hand sign recognition is one of most challenging issues in computer vision and human computer interaction, and many researchers tackle this issue. In this research, we focus on JFSL (Japanese Finger-spelled Sign Language) which is one of hand signs. The tasks for achieving high performance of JFSL recognition as well as other hand signs are how to extract hand region precisely and how to recognize hand signs accurately. To deal with the former task, in this paper, we propose an automatic hand region extraction method with a depth sensor. The characteristic points of our proposed method are to utilize Time-Series Curve, which is one of contour features, and to extract hand region accurately without wearing landmark object such as a color wristband. On the other hand, to tackle the latter task, in this research, we focus on a deep neural network based recognition method since such a method is reported that it allows us to achieve high performance for various recognition tasks. Therefore, in this paper, we investigate JFSL recognition performance with a deep neural network approach compared to that with the conventional image recognition method (HOG+SVM). From the experimental results with 8 subjects, we have confirmed that our proposed method allows us to extract hand region accurately regardless of subjects and JFSL signs. In addition, from the experimental results with a deep neural network based recognition method for JFSL recognition, we have achieved at least average recognition rate over 88%.

Peer-review under responsibility of KES International

Keywords: Japanese Finger-spelled Sign; Automatic Hand Region Extraction; Time-Series Curve; Depth Sensor; Deep Neural Network

1. Introduction

Recently rapid aging leads to increasing number of people with hearing difficulties and this has become a serious problem. To solve this problem, HSL (hand sign language) recognition has been splotlighted as a key technology to assist the communication with them. Although many researchers have already proposed the HSL recognition methods[1] and they have reported that their methods allow us to achieve high recognition performance, the number of HSL words that they can recognize is restricted in their research. Since HSL recognition is one of most challenging tasks in the some research fields such as computer vision and human computer interaction etc., it is still difficult to

* Corresponding author. Tel.: +81-72-254-7279.
E-mail address: inoue@cs.osakafu-u.ac.jp

Peer-review under responsibility of KES International
doi:10.1016/j.procs.2015.08.145

realize practical recognition system. To solve this problem, in this research, we focus on finger-spelled sign language, which is one of HSL, since the number of signs to master it for communication is much less than that of signs of HSL. From this viewpoint, the purpose of our research is to make a finger-spelled sign language recognition system, especially Japanese finger-spelled sign language recognition system[2], for the first step of smooth communication with the hearing-impaired people.

JFSL (Japanese Finger-spelled Sign Language) is a representation method of Japanese syllabary characters called Hiragana. Although signer needs to represent the characters one by one, JFSL is often utilized for representation of proper nouns. Each sign of JFSL is basically represented as the shape of a gesturing hand and some signs are represented by moving a gesturing hand while keeping the shape of hand for their base sign. The meaning of these signs is changed depending on the movement direction of hand. Since the number of Hiragana is greater than that of alphabet, JFSL recognition is more difficult and challenging task compared with ASL (American Sign Language) recognition. To deal with JFSL, the recognition system is divided into 3 processes; 1) hand region extraction, 2) hand shape recognition, 3) tracking a gesturing hand and recognition of its moving direction. For the first step to realize this system, in this paper, we mainly mention the first process and secondarily mention the second process.

For the first process, recently, depth sensor has been focused on as a key device for hand region extraction[3–5] because the methods with depth sensor enable us to robustly extract hand region for some noises, e.g. illumination change and cluttered background, compared with the color based methods[2,6–8]. However, in order to extract hand region accurately, some of them require signers to represent signs within the specific distance range from the depth sensor and require their hand to be the front-most object from the depth sensor. To relax these restrictions, a method using a depth sensor and a color wristband has also been proposed[4]. However, in such a method, wearing a color wristband on a gesturing hand's wrist is necessary to extract hand region precisely, which is rather inconvenient for real world applications. To solve these problems mentioned above, in this paper, we propose a new depth sensor based hand region extraction method by utilizing TSC (Time-Series Curve)[9], which is one of contour features. The characteristic point of our proposed method is to extract hand region automatically without wearing a landmark item such as color wristband. From the experimental results with 8 subjects for JFSL signs, we have confirmed that our proposed method allows us to extract hand region accurately regardless of subjects and JFSL signs. Additionally, compared with the conventional method with a depth sensor and a black wristband, we have achieved similar extraction results without wearing it on a gesturing hand's wrist. Moreover, in this research, we have investigated JFSL signs recognition accuracy with the hand region images extracted by our proposed method. From the experimental results with a deep neural network based method[10–12], we have achieved at least average recognition rate over 88%.

Although the contributions of our proposed method are little in the research area such as computer vision and human computer interaction, compared with existing work, our proposed method has following contributions.

1. In our proposed method, we employ TSC for extracting hand region precisely with depth sensor. As shown in Fig. 8, by using TSC, we have achieved similar extraction result for hand region without wearing a color wristband compared with the conventional method using a color wristband.
2. To our knowledge, this is the first research of JFSL signs recognition utilizing deep neural network based method. In some of JFSL signs recognition methods, HOG (Histogram of Oriented Gradient)[13] and SVM (Support Vector Machine) are utilized. Compared with these methods, in this research, we utilize simple deep neural network based method, i.e. the network includes convolution, pooling, normalization and activation layers.

The rest of this paper is organized as follows. In Section 2, we introduce related work of hand region extraction and finger-spelled sign language recognition. Next we explain our proposed method in Section 3. In Section 4, we show the experimental results. Finally, we conclude this paper in Section 5.

2. Related Work

Finger-spelled sign language recognition is one of most challenging tasks in computer vision and human computer interaction. In the research of finger-spelled sign language recognition, many researchers mainly deal with AFSL (American Finger-spelled Sign Language)[3,6,7], BFSL (British Finger-spelled Sign Language)[8] or FN (Finger-spelled Number)[4,5,14]. For these work, especially, we handle JFSL (Japanese Finger-spelled Sign Language) because JFSL