A Comparative Study of Shape and Texture Features for Finger Spelling Recognition in Big Data Applications

Yong Hu, Jinling Institute of Technology, Nanjing, China

ABSTRACT

With a wide variety of big data applications, Sign Language Recognition has become one of the most important research areas in the field of human-computer interaction. Despite recent progresses, the task of classifying finger spelling is still very challenging in Sign Language Recognition. The visually similarity of some signs, the invisibility of the thumb and the large amount of variation by different signers are all make the hand shape recognition very challenging. The work presented in this paper aims to evaluate the performance of some state-of-the-art features for static finger spelling of alphabets in sign language recognition. The comparison experiments were implemented and tested using two popular data sets. Based on the experimental results, analysis and recommendations are given on the efficiency and capabilities of the compared features.

KEYWORDS


INTRODUCTION

Sign language recognition (SLR) began to appear in the 1990s. With a wide variety of big data applications, SLR 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 (Murthy & Jadon, 2009). 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 (Kulkarni & Lokhande, 2010).

In the article by Kelly et al. (2010), they present a user independent framework based on Support Vector Machine (SVM). An eigen-space 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. Dahmani and Larabi (2014) proposed a framework based on the combination of three shape descriptors: Discrete orthogonal Tchebichef moments applied on both internal and external outlines hand, Hu moments and a set of geometric features derived from the convex hull that encloses the hand shape taking into account the hand orientation. The proposed descriptors are combined in several sequential and parallel manners and applied on different datasets.

DOI: 10.4018/IJMCMC.2017070105

Copyright © 2017, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Otiniano-Rodriguez et al. (2012) proposed Sign Language Recognition methods by using the SVM classifier and features extracted from Hu and Zernike Moments. A comparison between the proposed methods using a database composed of 2040 images of 24 symbol classes is performed. Experiments show the Zernike moments features achieved a higher accuracy.

Hrůz et al. (2011) experiment the Local Binary Patterns (LBP) based features for Sign Language Recognition. The recognition performance of LBP features was compared with the geometric moments and their combinations. The experiments on a database consisting of 11 signers and 23 signs show that the Local Binary Patterns outperform the geometric moments. Chou et al. (2014) adopt the LBP operator to generate the hand texture to deal with the finger spelling not recognizable by the hand structure. The experimental results show that this system is an effective real-time recognition system with high accuracy.

Kim et al. (2013) developed a semi-Markov conditional random field (SCRF) approach to the unconstrained finger-spelling recognition problem. The concatenation of the Histograms of Oriented Gradients (HOG) descriptors were used as visual descriptor for a given hand region. Thippur et al. (2013) evaluate three different state-of-the-art visual shape descriptors, Hu Moments, Shape Context and HOG, which are commonly used for hand and human body pose estimation. Some recommendations were given based on the evaluation experiments. Escobedo Cardenas et al. (2015) use depth information to characterize hand configurations and calculated three histograms of cumulative magnitudes as input to SVM classifier.

Despite recent progresses, sign language recognition systems are still in its infancy. The visual similarity of some signs, the invisibility of the thumb and the large amount of variation by different signers all make the hand shape recognition very challenging. In this work, we focus on the problem of recognizing hand shapes from the ASL (American Sign Language) finger spelling alphabet. Five different state-of-the-art visual shape and texture descriptors, Hu moments, Zernike Moments, Local Binary Patterns (LBP), Gray-Level Co-occurrence Matrix (GLCM) and Histogram of Gradients (HOG) are evaluated. These features are commonly used for Finger Spelling Recognition recently. However, the existed approaches were conducted in different dataset by using different classifier. That makes it difficult to tell what kinds of features are suitable for the recognition task. Evaluation experiments and analysis are conducted by using Jochen-Triesch static hand postures database (Marcel, 2001) and Thomas Moeslund’s gesture recognition database (Moeslund, 2002).

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

SHAPE AND TEXTURE DESCRIPTORS

Hu Moments

Hu moments are a reformulation of the non-orthogonal centralized moments from the hand contour. The moments are different combinations of scale-normalized and centralized moments of the shape silhouette, which makes it scale and rotation invariant. The set of seven Hu moments are calculated from the 2D moments and central geometric moments by using the algebraic invariants. The definition of the seven invariant Hu moments is list as follows:
iPod as a Visitor’s Personal Guide
www.igi-global.com/chapter/ipod-visitor-personal-guide/17102?camid=4v1a

MagiThings: Gestural Interaction with Mobile Devices Based on Using Embedded Compass (Magnetic Field) Sensor
www.igi-global.com/chapter/magithings/133749?camid=4v1a