Gesture Recognition and Machine Learning Applied to Sign Language Translation

Luis A. Estrada Jiménez[1], Marco. E. Benalcázar[2] and Nelson Sotomayor[1]

[1]Escuela Politécnica Nacional/Departamento de Automatización y Control Industrial, Quito, Ecuador
[2] Escuela Politécnica Nacional/Departamento de Informática y Ciencias de la Computación, Quito, Ecuador

Abstract— **In this paper we propose an intelligent system for translating sign language into text. This approach consists of hardware and software. The hardware is formed by flex, contact, and inertial sensors mounted on a polyester-nylon glove. The software consists of a classification algorithm based on the k-nearest neighbors, decision trees, and the dynamic time warping algorithms. The proposed system is able to recognize static and dynamic gestures. This system can learn to classify the specific gesture patterns of any person. The proposed system was tested at translating 61 letters, numbers, and words from the Ecuadorian sign language. Experimental results demonstrate that our system has a classification accuracy of 91.55%. This result is a significant improvement compared with the results obtained in previous related works.**

Keywords— **Sign language translation, gesture recognition, machine learning, pattern classification**

I. INTRODUCTION

Worldwide, at least 360 million people are likely to be deaf and dumb (or simply deaf) [1]. Most of these people do not have access to public services. In the case of hearing impairment, mutism has created the need of using sign language. A sign language consists of a set of manual gestures that use fingers, palms, arms, and even body movements to represent letters, numbers, and words. This type of language allows people who cannot speak or hear to communicate with others. However, most of the non-disabled people do not know this type of language. Therefore, for deaf people performing daily activities turns out to be hard, especially in public areas. Additionally, it is difficult and expensive to make non-disabled people learn sign language. For these reasons, automatic systems that translate sign language into text are required.

Some systems that recognize gestures and translate sign language into text have been proposed. Most of these systems use image processing techniques to perform the translation [2]. Other approaches use different kinds of sensors to acquire the orientation of the hand and fingers [3] [4]. The main drawback of all these systems is that they cannot distinguish similar gestures neither patterns that consist of movements (i.e., dynamic patterns).

In [3], the authors propose a sensorized glove combined with Euclidean classifiers capable of translating up to 10 static patterns (i.e., patterns that have no motion). A similar system is proposed in [4]. This system is able to recognize the Malaysian sign language. This approach is based on flex sensors and an accelerometer, and is able to recognize 9 static patterns. Even though these approaches show good classification accuracy, they were not developed to provide a wide range of recognition of words and expressions.

In [5], [6], and [7], prototypes of sign language translators based on the Kinect sensor are proposed. In [7], a system based on depth-estimates from the Kinect sensor is presented. To compute the angles of the fingers, this system estimates the position of both hands and some features of their contours. The dynamic time warping (DTW) algorithm is used for a pattern recognition task. This system is able to identify 52 static and dynamic gestures. Since these approaches are based on artificial vision systems, their performance is highly dependent on the light conditions where the system is used. In [8], the sign language translation is based on a sensor called leap motion controller. This approach uses decision trees (DT) and genetic algorithms to recognize 24 static characters with 82.7% of classification accuracy. Besides this relatively good performance, this approach is not capable of recognizing dynamic gestures.

Because of the problems described above, the field of gesture recognition and sign language translation is still open for research. In this paper, we propose an intelligent system to translate sign language into text. This approach can learn to recognize different gestures (i.e., static and dynamic) using machine learning and pattern recognition techniques.

The hardware of the proposed system was mounted on a polyester-nylon glove which has good properties in terms of durability, elasticity, and comfort. The hardware is composed of 3 types of sensors: flex, contact, and inertial. Combining the information of these sensors, we estimated the relative orientation and movement of the hand and its fingers. These estimates are used as the inputs of a gesture recognition system. This system is based on 3 classification algorithms: k-nearest neighbors (k-NN), decision trees, and the DTW algorithm [9]. The first two algorithms were used to recognize static gestures, whereas the DTW algorithm was used to identify dynamic gestures. We tested the proposed system at translating 61 gestures from the Ecuadorian sign language [10]: 30 letters, 10 numbers, and 21 expressions (Fig. 1).

The proposed system for gesture classification is based on a learning process. This process makes possible that our system can be trained to work with other sign languages different from the Ecuadorian one. Moreover, the proposed system could be trained to recognize signs and gestures from other domains different from the sign language

© Springer Nature Singapore Pte Ltd. 2017

I. Torres et al. (eds.), *VII Latin American Congress on Biomedical Engineering CLAIB 2016, Bucaramanga, Santander, Colombia, October 26th -28th, 2016*,
IFMBE Proceedings 60,
DOI: 10.1007/978-981-10-4086-3_59

translation. These domains can include gaming, robotics, and assistive technology.

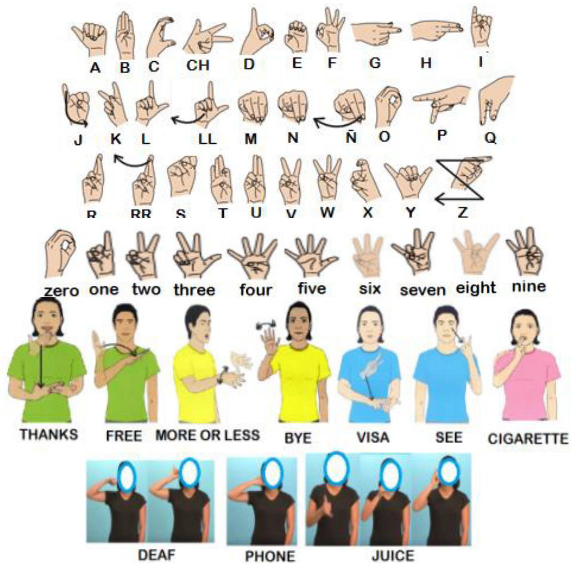

Fig. 1 Patterns and gestures that represent letters, numbers, words, and expressions of the Ecuadorian Sign Language [10]

This paper is organized as follows. In the first section, we have presented an introduction about the problem of sign and gesture recognition applied to sign language translation. The second section provides information about the materials and methods we used to develop the proposed approach. In the third section, we present and discuss the results obtained. Finally, in the fourth section, we present some concluding remarks of this work.

II. Materials and Methods

A. *Types of hand movements*

This work is based on the theory developed by Rouviere and Delmas [11]. These authors studied the following 4 types of movements of the hand: flexion and extension of the fingers, flexion and extension of the wrist, supination and pronation of the forearm, and radial and ulnar deviation of the hand. Fig. 2 shows some of these movements.

Fig. 2 Movements of the hand according to Rouviere and Delmas [11]

B. *Description of the Developed Hardware*

In this section, we describe the structure of the developed prototype including the sensors and the processing stage of their signals.

All the sensors used in this work were mounted on a polyester-nylon glove and their distribution is according to the graph that is shown in Fig. 3. The flexion and extension signals of the fingers were acquired using flex sensors. Two flex sensors were attached to each finger, except in the case of the thumb where we used only one sensor. We used two flex sensors because the two phalanges are needed for representing the movement of a finger [11]. The hand movements were measured by an inertial measurement unit (IMU) placed on the back of the hand. The IMU used for this work is the MPU-6050. This sensor measures the pitch, roll, and yaw angles (i.e., Euler angles). These angles were used to estimate the orientation of the hand in a 3D space.

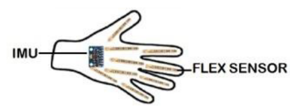

Fig. 3 Distribution of the sensors on the glove

The signals from the flex sensors were preprocessed by using a Wheatstone bridge followed by a differential amplifier. Then, all these signals were digitalized, with 10 bits of resolution, using an Arduino mega board. The data from the MPU-6050 IMU sensor was sent to the Arduino board using I^2C communication. Then, this information was processed in the Arduino to obtain the corresponding Euler angles. In this work, we only used the pitch and roll angles. We did not use the yaw angle since it is required a magnetometer to provide an accurate estimate of this angle. An aluminum contact sensor was placed between the middle and index fingers. This was done because the flex sensors and the MPU-6050 did not allow differentiating accurately the letters U and V. For each gesture, we obtained 12 digital signals: 9 from the fingers, 1 from the contact sensor, and the 2 Euler angles. Finally, all these signals were sent through a radio-frequency module to a personal computer for the classification task.

The proposed system is able to classify the movements and gestures of the hand into 61 different classes. Some gestures can only be differentiated from others based on the context where they are used. An example of this situation is number 2 and letter U. Because of this reason, the set of gestures to be classified was divided into 3 groups: numbers, letters, and words. Before testing the proposed system, the user needs to select the group inside which the classification will be performed.

C. *Training Process*

The proposed system needs to learn to recognize the movements and gestures for each user. For this process, we used a 12-dimensional feature vector containing the information from all the sensors on the glove. For training the system, the user must perform 5 repetitions for each gesture to be recognized. This value was selected as a tradeoff between the minimum number of samples for