HAND GESTURE RECOGNITION USING ORIENTATION HISTOGRAM AND BACKPROPAGATION NEURAL NETWORK

Khin Pa Pa Aung, Zin May Aye
University of Computer Studies (Mawlamyine), Myanmar
khinpapaaung.257@gmail.com

Abstract

This paper presents to recognize American Sign Language (ASL) hand gestures, based on a pattern recognition technique by using orientation histograms and backpropagation neural network (BPNN). The static ASL digitized images of English alphabets are used in this hand gesture recognition system. In this system orientation histogram is used as feature vectors because of its robustness in lighting changes conditions of images and also the position of the hand within the image should not affect the feature vector. The advantage of using BPNN is that it can perform a particular function by adjusting the values of the connections between elements, so input feature vector leads to specific target output. This system consists of Image Processing, Training and Testing phases in BPNN. The output of the system will be displayed the corresponding alphabet letter with corresponding input feature vectors. This system will be beneficial between the deaf and hearing communities problems.

Keywords: American Sign Language, Orientation Histogram, Backpropagation Neural Network, Feature Vector, Hand Gesture

1. Introduction

Hand gesture recognition is important for the communication between person with hearing ability and the hearing impaired as well as the research on the human – computer interface. Hand gesture recognition has various applications like computer games, machinery control, and thorough mouse replacement. Gesture recognition is a difficult task in computer vision due to the numerous degrees of freedom of a human hand. Detecting and understanding hand and body gestures is becoming a very important and challenging task in computer vision. The significance of the problem can be easily illustrated by the use of natural gestures that we apply together with verbal and nonverbal communication. Recognition of the shape of the hand is also important in some applications like Sign Language recognition and Human Computer Interaction. [2] Hand gestures can be classified in two categories: static and dynamic. A static gesture is a particular hand configuration and pose, represented by a single image. A dynamic gesture is a moving gesture, represented by a sequence of images. [3] The system is focused on the recognition of static hand gesture digitized images. Acquired images are transformed into Orientation Histograms to apply as a feature vector for neural network to recognize.

An artificial neural network, usually called "neural network", is a mathematical model or computational model that tries to simulate the structure and/or functional aspects of biological neural networks. Neural networks are non-linear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. Artificial Neural Networks are mostly applied in many types of applications such as pattern recognition, identification, classification, speech, vision and control system. ANN has supervised and unsupervised learning. [6] This system applied supervised learning back propagation neural network.

This paper is organized as follows. Section 1 presents introduction of the system. The related work for this system is described in Section 2. In Section 3, theory background related to this system is presented. Section 4, presents the design and implementation of this system. In Section 5, results and conclusions of this paper are described.

2. Related Work

Recently, there has been a surge in interest in recognizing human hand gestures. Where, some researcher interested in static (isolated sign language recognition) gesture and some researcher interested in continuous gesture recognition. [1] Peter O'Donovan analyzed Static Gesture Recognition by applying Restricted Boltzmann Machines (RBMs) to model the manifold of 3 human gestures: pointing, thumbs up, fingers spread.[5] Mahmoud Elmezain, Ayoub Al-Hamadi, and Jorg Appenrodt proposed an automatic system that recognizes both isolated and continuous gestures for Arabic numbers (0-9) in real-time based on Hidden Markov Model (HMM).[4] Christian Vogler and Dimitris Metaxas analyzed American Sign Language Recognition, using Parallel Hidden Markov Models to achieve parallel processes independently.[1]

This paper is developed to recognize static ASL hand gesture using orientation histogram with backpropagation neural network. Orientation histogram is used for feature vectors to train and test in the neural network.

3. Background

The related background theory of hand gestures recognition system with orientation histogram is described in this section.

3.1 American Sign Language

American Sign Language is the language of choice for most deaf people in the United States. It is part of the "deaf culture" and includes its own system of puns, inside jokes, etc. ASL is one of the many sign languages of the world. ASL consists of approximately 6000 gestures of common words with finger spelling used to communicate obscure words or proper nouns. Finger spelling uses one hand and 26 gestures to communicate the 26 letters of the alphabet.[3] Some ASL for English alphabets are shown in Figure 1.

A B C

Figure 1: Examples for ASL

3.2 Image Processing

The intention is to convey an idea of all the methodologies that can be applied to images for different purposes and possibly with different objectives. The image processing steps may not be the same for all types of applications according to the requirements and objectives of the systems. But there are certain common image processing steps encountered in the most of the image processing systems. Some of the image processing steps that are applied by this system are – image acquisition, resizing and gray scaling, thresholding, edge detection, creating orientation histogram.

3.3 Backpropagation Neural Network

The weight adapt learning process performed with the algorithm is called back-propagation learning. The back-propagation algorithm is the most widely used algorithm to perform training, particularly on large problems. Back-propagation is used for multi-layer networks and non-linear differentiable transfer functions. Input vectors and the corresponding target vectors are used to train a network until it can approximate a function, associate input vectors with specific output vectors, or classify input vectors in an appropriate way as defined by user. It is supervised learning. It consists of at least three layers: an input layer, at least one hidden layer, and output layer. The Backpropagation Neural Network Architecture for this system is shown in Figure 2. [6]

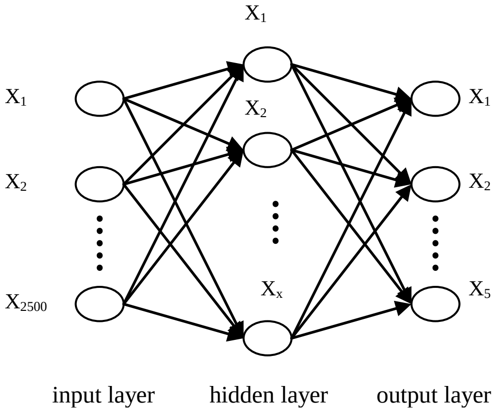


Figure 2. Backpropagation Neural Network Architecture

4. System Design and Implementation

This section presents the system design and the implementation of the system.

4.1 System Design

In this system it consists of two parts: Image Processing and Neural Network for recognizing process. In this paper Orientation Histogram (OH) is applied to achieve feature vectors for train and test in Neural Network. Figure 3 describes the system design of this system.

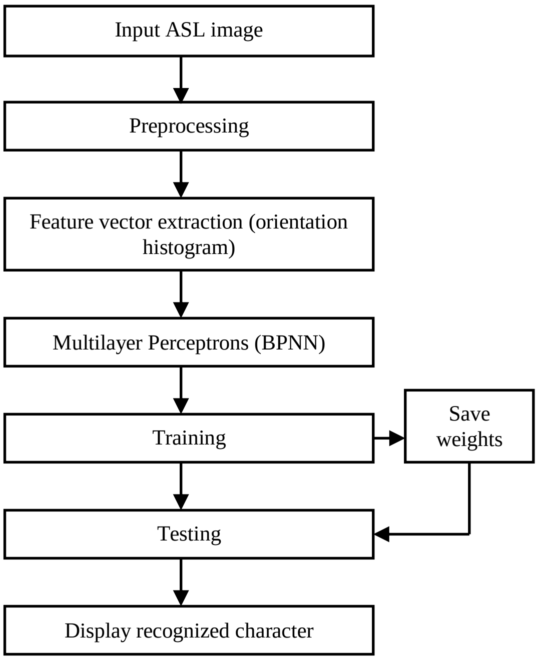


Figure 3. System Design

4.2 System Implementation

In this implementation of the system, it consists of two parts. First part is image preprocessing and the second is training and testing input feature vectors in BPNN.

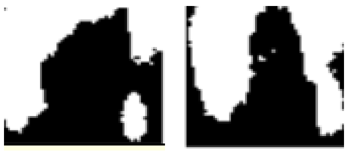
4.2.1 Image Preprocessing

All ASL digitized images are firstly resized and transformed into grayscale images. A grayscale image has no color information. Therefore, every pixel in a grayscale image has different shade of gray which is commonly represented by 8 bits. So, there are $2^8 = 256$ possible intensity values (shades of gray) for a grayscale image ranging from 0 to 255. Examples of grayscale images for ASL sign A are shown in Figure 4.

A Original **A** Grayscale

Figure 4. Origin and Grayscale of ASL for A

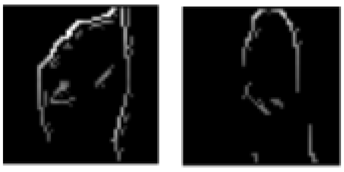
The second step in image processing is thresholding. Thresholding is a non-linear operation that converts a gray-scale image into a binary image where the two levels are assigned to pixels that are below or above the specified threshold. The input to a thresholding operation is typically a grayscale or color image. In the simplest implementation, the output is a binary image. Threshold images for A & B of ASL signs are described in Figure 5.



A **B**

Figure 5. Threshold Image

Thirdly the system used C# library for Canny Edge Detection algorithm. It consists the steps of Smoothing, Finding gradients, Non-maximum suppression, Double thresholding, and Edge tracking by hysteresis. Figure 6, showed edges for A & B signs.



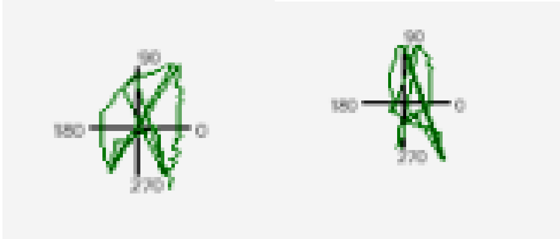
A **B**

Figure 6. Edge Detection

Fourth, the orientation histogram algorithm is applied to extract feature vectors.

Orientation Histogram

In this process x and y derivative filters are used to calculate gradients of x and y, then the gradient direction is calculated with arctan(dx,dy), and the contrast is $\mathrm{sqrt}(dx^2+dy^2)$. Feature vectors are extracted from this orientation histogram. All feature vectors are then trained in the NN for training and testing. Figure 7, show orientation histograms for A & B ASL signs.



A B

Figure 7. Orientation Histogram for A & B

4.2.2 Training Phase

In the training phase, the weight and the required parameters are initialized. The input images are accepted and preprocessed. Then orientation histograms are used as feature vectors for the network to be trained. Then the network calculates the net input and actual output for each hidden neuron. And also it calculates the net input and actual output for each output layer neurons. Then the network computes error between target and actual output from the network. In this system the target error is set to 0.01. If the error is less than the maximum error, weight will save in the network. If the error is greater than the maximum error caculate the error for output and hidden layers and update weights by backpropagation. Training program continues until the maximum error is less than the sum squared error.

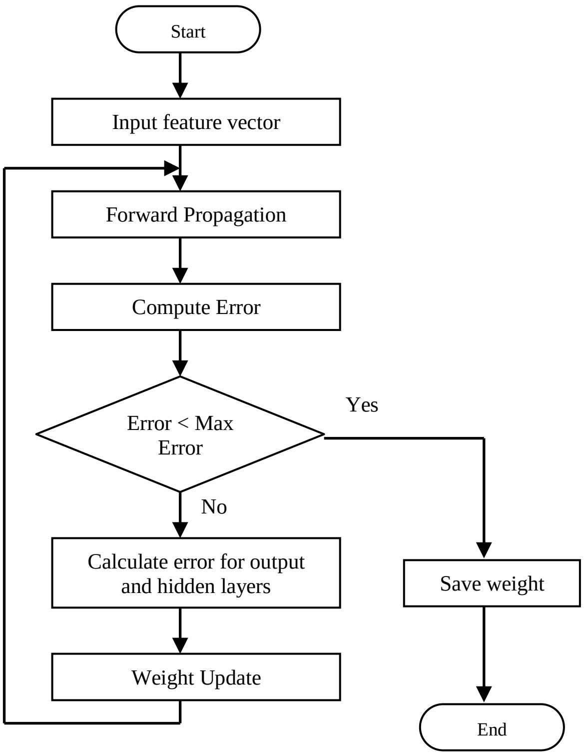


Figure 8. Training Phase of the System

4.2.3 Testing Phase

In this testing phase, the inputs image from the user are accepted and feature vectors are extracted by orientation histogram and applied them as testing data. The testing network consists of only feed-forward propagation. The testing network uses training network's weight as it's knowledge. As the network is a testing network, it does not need target output. Figure 9 depicts the testing phase of the system. If the user gives the testing data set to the network, the output of the network will show the corresponding character for the input feature vector.

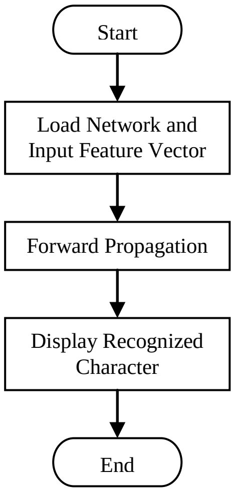


Figure 9. Testing Phase of the System

User interface form for testing phase is shown in Figure 10. In this figure firstly, test image (ASL sign) is loaded and then it is preprocessed with C# programming. Orientation Histograms are used as feature vectors for testing. Neural network has previously trained and has a save optimal weight and it is used in testing for the neural network to recognize the testing data. The result will display the recognized character corresponding with the input feature vector of ASL image.

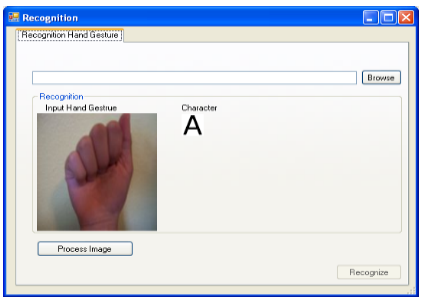


Figure 10. Output Result for ASL sign A

5. Results and Conclusions

In this section, results and conclusions of the hand gesture recognition system is described.

5.1 Results

The results for training and testing of hand gesture for English alphabets are shown in Table 1.

Table 1. Recognition Rate

	Data set	Recognized data set	Recognition Rate (%)
Training	390	360	92.3
Testing	260	200	76.9
Total	650	560	86.2

From this table 390 data set are trained and 260 data set are tested. According to the figures neural network can more recognized training data than testing data set. Some testing data set (feature vectors of orientation histogram) are nearly similar such as in ASL signs of (A, M, N, T,S), (G,H), (K,U,V,W,B), and (C,O). Recognition rate is lower in testing these images. Training data set are fully recognized in this BPNN and higher recognition rate than testing data set.

5.2 Conclusions

This system is developed by hand gestures recognition for English alphabets using static ASL images. Orientation histogram and backpropagation neural network (BPNN) are applied to implement the recognize system. Feature vectors of the input images are extracted by using orientation histogram to train and test in the BPNN. Training and testing results are also analysed in this system to know how network can recognize the input data set. The output of this system showed the recognized characters with respect to input feature vectors. This system can further be extended for real-time ASL and can also extend in the applications of facial expressions to distinguish between statements, questions and directives.

References

[1] C.Vogler and D.Metaxas, "Parallel Hidden Markov Models for American Sign Language Recognition, September 1999.

[2] F.Dadgostar,A.L.C.Barczak, "A Color Hand Gesture Database for Evaluating and Improving Algorithms on Hand Gesture and Posture Recognition",2005,Vol 7, pp 127-134.

[3] K.Symeonidis, "Hand Gesture Recognition Using Neural Networks", August 23,2000, School of Electronic and Electrical Engineering

[4] M.Elmezain, A.Hamadi, "A Hidden Markov Model-Based Continuous Gesture Recognition System for Hand Motion Trajectory" J¨org Appenrodt.

[5] P.O'Donovan, "Static Gesture Recognition with Restricted Boltzmann Machines".
[6] http:// www.Wikipedia.com/ Artificial neural network.