Environment. Technology. Resources, Rezekne, Latvia
Proceedings of the 11th International Scientific and Practical Conference. Volume II, 185-188

Latvian Sign Language Recognition Classification Possibilities

Aleksejs Zorins, Peter Grabusts
Faculty of Engineering, Rezeknes Academy of Technologies, Atbrivoshanas alley 115, Rezekne, LV-4601, Latvia, E-mail: Aleksejs.Zorins@ru.lv, Peteris.Grabusts@ru.lv

Abstract. There is a lack of automated sign language recognition system in Latvia while many other countries have been already equipped with such a system. Latvian deaf society requires support of such a system which would allow people with special needs to enhance their communication in governmental and public places. The aim of this paper is to recognize Latvian sign language alphabet using classification approach with artificial neural networks, which is a first step in developing integral system of Latvian Sign Language recognition. Communication in our daily life is generally vocal, but body language has its own significance. It has many areas of application like sign languages are used for various purposes and in case of people who are deaf and dumb, sign language plays an important role. Gestures are the very first form of communication. The paper presents Sign Language Recognition possibilities with centre of gravity method. So this area influenced us very much to carry on the further work related to hand gesture classification and sign's clustering.

Keywords: artificial neural networks, centre of gravity, classification, hand gesture, sign language recognition.

I. INTRODUCTION

Gesture is one of very important people's communication components, which allows expressing emotions and gives important information in addition to spoken language. For most people the gesture is an additional method for communication, for deaf this is the main way to express themselves. Deaf people are being integrated into society through the sign language, the part of which is the representation of the national alphabet gestures. These alphabets mainly consist of static signs; however, the Latvian sign language has several signs, which are shown in motion (see Table 1).

In Latvia there is a website of Latvian Deaf People Rehabilitation, which has a Sign Language Interpreters' Department. The main goal of this organization is to "facilitate the client's social integration, availability of necessary information and services, provide sign language interpreter's services for communication with other individuals and legal entities according to the client's perception and communication abilities" [1].

There is still a lack of automated Latvian sign language recognition system in our country and there is a strong need for it from. The ultimate goal of the authors is to develop the recognition system of Latvian Sign Language based on Artificial Neural Networks to help people in social rehabilitation and integration in modern society. This is the main purpose of this paper and the starting point of the authors' research is to classify Latvian Sign Language alphabet using Artificial Neural Networks.

Nowadays there is an increasing interest in automatic SLR using video or web cameras to classify gestures correctly and transform deaf people language into a text or speech.

Fig. 1. The symbols of Latvian sign language [1]

ISSN 1691-5402

http://dx.doi.org/ 10.17770/etr2017vol2.2653

There is a wide range of SLR methods, which could be mainly described using input data and sensor technology for sign recording:

- Using different markers (on hands);
- Specially developed data gloves;
- Infra-red sensor technology (LeapMotion, Microsoft Kinect etc.);
- Sign recognition through visual (video or photo) methods.

Two last methods allow to recognize the sign language in a real time through web cameras or Leap Motion controller, which price is now really affordable (less than 70 euros) [2].

Next chapters will give a review of data pre-processing and recognition solutions with artificial neural networks

II. ARTIFICIAL NEURAL NETWORKS AND CLASSIFICATION PROBLEM

The simplified biological neuron as a model on artificial neuron has been created and implemented in many types of ANN (see Fig. 2):

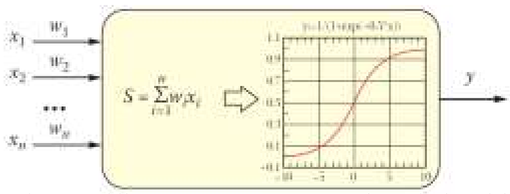

Fig. 2. Multilayer Perceptron neuron mathematical model [3]

In Figure 2 artificial neuron receives its inputs $x_1 \ldots x_n$ weighted with $w_1 \ldots w_n$. All inputs are calculated with a sum function and then propagated to activation function. After calculations neuron gives output value y, which is led to other neurons or is processed like a final output signal of the network. The artificial neural network itself is an interconnected neuron structure. These types of neurons are often used in back-propagation artificial neural networks.

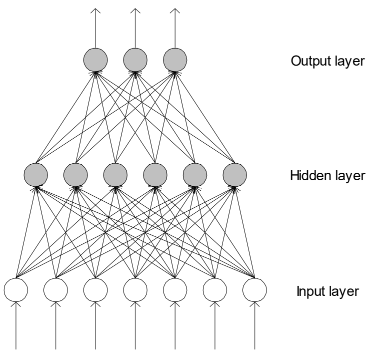

Fig. 3. The architecture of backpropagation network

A multilayer feed-forward network with an appropriate pattern of weights can be used to model some mapping between sets of input and output variables. Figure 3 shows an example of feed-forward network architecture, with three output units and one hidden layer, which can be trained using back-propagation. The shaded nodes in this figure are processing units. The arrows connecting input and hidden units and connecting hidden units and the output units represent weights.

The back-propagation learning algorithm [3], [4] is formulated as a search in the space of the pattern of weights W, in order to find an optimal configuration W*, which minimizes an error or cost function E(W). The pattern of weights will then determine how the network will respond to any arbitrary input. The error or cost function is defined by (1):

$$E = \frac{1}{2}\sum_i \sum_p (t_{ip} - o_{ip})^2 \qquad (1)$$

This function compares an output value o_{ip} to a desired value tip over the set of p training vectors and i output units. The gradient descent method is used to search for the minimum of this error function through iterative updates:

$$W(k+1) = W(k) - \eta \nabla E \qquad (2)$$

where η is the learning rate, and ∇E is an estimate of the gradient of E with respect to W.

The algorithm is recursive and consists of two phases: forward-propagation and backward-propagation. In the first phase, the input set of values is presented and propagated forward through the network to compute the output value for each unit. In the second phase, the total-squared error calculated in the first phase is propagated from the output units to the input units. During this process, the error signal is calculated recursively for each unit in the network and weight adjustments are determined at each level. These two phases are executed in each iteration of the back-propagation algorithm until the error function converges.

A very important step in each neural network application is data pre-processing. In Sign Language Recognition problem this task has also a big necessity. This paper will not go into details with this issue, additional information could be found in our earlier research paper [5].

There are different architectures and learning algorithms used in classification approach. One of them is Modular Neural Network presented in [5]. These networks have improved generalization due to decomposition of complex function into simpler ones. The main idea is a natural decomposition of a function of large complexity into simple functions and realization of each function by a separate neural network [6]. The modular neural network with an output decomposition is shown in Fig. 4.

In this particular case a task is divided into several subtasks, each of which can be solved individually using an independent neural network.

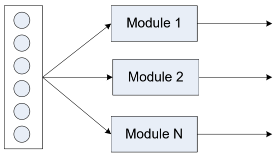

Fig. 4. Decomposition of the outputs of different modules

Another example of classification neural network is Kohonen "freezing" learning algorithm developed by author [7]. The standard Kohonen self-organizing maps are trained in unsupervised and in supervised manner. This type of network uses grid of neurons or a topological structure among the cluster units. There are m cluster units, arranged in a one- or two-dimensional array: the input signals are n-dimensional. Figure 5 shows architecture of a simple self-organizing network, which consists of input and Kohonen or clustering layer.

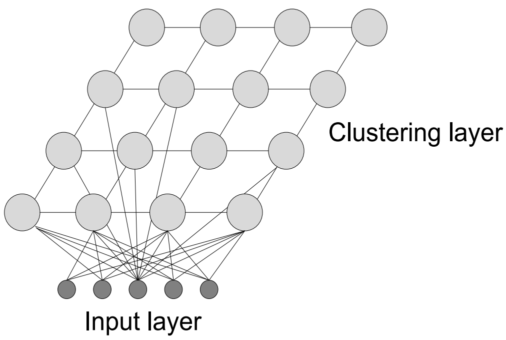

Fig. 5. The architecture of the Kohonen self-organizing map

The shadowed units are processing units. Usually one clustering unit for each class is not enough; therefore each clustering layer neuron in the Figure 4 consists of several neurons.

When a self-organizing network is used, an input vector is presented at each step. These vectors create the "environment" of the network. Each new input produces an adaptation of the parameters. If such modifications are correctly controlled, the network can build a kind of internal representation of the environment [6].

Consider the problem of charting an n-dimensional space using a one-dimensional chain of Kohonen units [3]. The units are all arranged in sequence and are numbered from 1 to m (see Fig. 6).

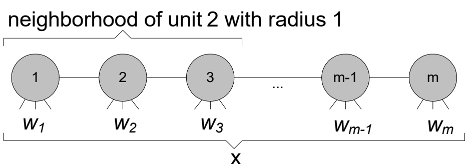

Fig.6. A one-dimensional lattice of computing units

The n-dimensional weight vectors w_1, w_2, …,w_m are used for the computation. The objective of the charting process is that each unit learns to specialize on different regions of input space. When an input from such a region is fed into the network, the corresponding unit should compute the maximum excitation. Kohonen's learning algorithm is used to guarantee that this effect is achieved.

There is a modified "freezing" algorithm developed by author allows splitting network learning process into some stages, when each part of the network is trained individually [7]. In the first learning step neural network is divided into number of clusters of neurons, where each of the clusters is associated with a definite class. In this way we obtain training with teacher. In the second stage each cluster is trained accordingly to standard Kohonen learning algorithm. Each of neuron clusters is trained individually, while others are "frozen" and do not take part in the training. After completion of individual cluster training the network is "de-frozen" and learning process ends [7].

III. PROPOSED HAND GESTURE RECOGNITION SYSTEM FOR CLASSIFICATION

There are different sign language recognition methods described in analytical research papers. One of them is recognizes hand gestures are classified as static and dynamic gestures (see Fig. 7.) [8]. Static hand gestures are those in which the hand position is unchanged during the gesturing period.

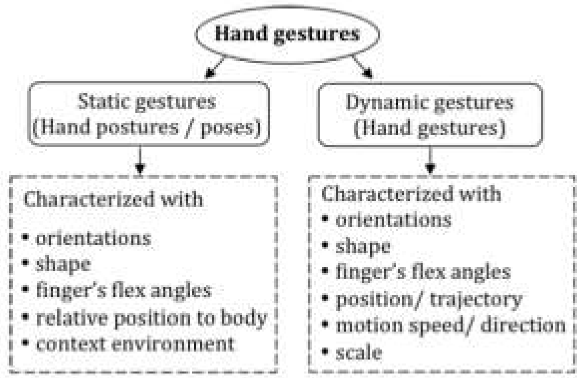

Fig. 7. Classification of hand gestures

The author's research focuses on static gestures only. Internet and other resources give us several sign language recognition applications with different methods [9], [10], [11].

Our hypothesis claims that a gesture symbol could be recognized and classified with a center of gravity method (COG). This method allows splitting alphabet symbols into several groups or clusters, thus simplifying recognition task (see Fig. 8).

Fig. 8. Center of gravity in human palm.

The COG is a common feature to check the movement route of a hand gesture - the difference between two COGs provides information about the movement speed and direction.

The hand gesture recognition procedure is divided into two stages. The first stage is preprocessing and the second stage is the classification one [11].

The first stage starts with_reducing the complexity of feature extraction for hand gestures. The output image of hand portion extraction process is converted into binary image.

There are various different kinds of distinguished features that can be extracted from an image. One of the possibilities is to use an active and in-active finger which is represented by 1 and 0 respectively. As a result an image of finger portion was obtained using only singly connected pixels, the co-ordinate values for the finger tip was stored accordingly.

After the pre-processing phase is complete it is possible to calculate a centroid of hand using the following equation [11]:

$$\bar{X} = \frac{\sum_{i=0}^{k} x_i}{k} \quad \bar{Y} = \frac{\sum_{i=0}^{k} y_i}{k} \qquad (3)$$

where $(\bar{X}, \bar{Y})$ represents the centroid of the hand;

x_i and y_i are x and y coordinates of the i^{th} pixel in the hand region;

k denotes the number of pixels that represent hand area.

Then distance between the centroid and the finger tip is calculated using Euclidean distance according an equation below:

$$D = \sqrt{(x_2 - x_1)^2 - (y_2 - y_1)^2} \qquad (4)$$

where (x_1,x_2) and (y_1, y_2) represent two coordinate values.

In the second phase of recognition an artificial neural network with back-propagation or Kohonen algorithm could be implemented.

IV. CONCLUSIONS

Most of sign language applications focus on feature extraction of the hand gesture for performance of classification. The main goal of the preprocessing stage is hand gesture extraction from an image, removal of noise and unwanted regions, processing the extracted image to a form of a binary image and extraction of the significant features from an image for further classification.

There is a number of successful applications in sign language recognition using Artificial Neural Networks, however, no one still has been developed similar solution for Latvian sign language. The ultimate goal of the authors is to develop Latvian Sign Language recognition system based on artificial neural networks.

This analytical paper gives an insight into sign language recognition technology. The authors are doing several experiments with Latvian Sign language at the moment, and one of the possible solutions is center of gravity method and artificial neural network.

We are in the beginning of our research hoping to help people with special needs and advance in a very interesting Artificial Intelligence field.

REFERENCES

1. The Latvian Sign Language Development Department. [Online]. Available: http://zimjuvaloda.lv/lv/alphabet/daktilo-zimju-alfabets [Accessed: March 23,2017] (in Latvian).
2. The Leap Motion Store. [Online]. Available: https://store-eur.leapmotion.com/products/leap-motion-controller [Accessed: March 23, 2017].
3. R. Rojas, *Neural networks. A systematic approach*. Springer, Berlin, 1996.
4. L. Fausett, *Fundamentals of Neural Networks. Architectures, algorithms and applications*. Prentice Hall, Inc., 1994, pp.169-187.
5. A. Zorins and P. Grabusts, "Review of Data Preprocessing Methods for Sign Language Recognition Systems based on Artificial Neural Networks," International Scientific Journal of Riga Technical University, vol. "Information Technology and Management Science", December, 2016, vol. 19, pp. 98–103.
6. A. Konar, *Computational intelligence: principles, techniques and applications*. Springer-Verlag, London, 2005.
7. A. Zorins, *Improvement possibilities of interval value prediction using Kohonen neural networks. Riga Technical University conference proceedings*, Vol. 31, 8.-16. Riga, Latvia, 2007.
8. P. K. Pisharady, M. Saerbeck, "Recent methods and databases in vision-based hand gesture recognition: A review," Computer Vision and Image Understanding 141:152-165, December 2015.
9. J.R. Pansare, S.H. Gawande, M. Ingle, "Real-time static hand gesture recognition for American Sign Language in complex background," Journal of Signal and Information Processing,V.3., pp. 364-367, 2012. http://file.scirp.org/pdf/JSIP20120300010_94521693.pdf
10. C.V. Ng, S. Ranganath, "Real-time gesture recognition and application, Image and Vision Computing," 20, pp. 993-1007, 2002.
11. M.K. Ghose, A. Pradhan, "A hand gesture recognition using feature extraction," International Journal of Current Engineering and Technology, 2012.