DOI:10.5392/IJoC.2011.7.1.014

Combining Dynamic Time Warping and Single Hidden Layer Feedforward Neural Networks for Temporal Sign Language Recognition

Ngoc Anh Nguyen Thi
Department of Computer Science
Chonnam National University, Gwangju 500- 757, South Korea

HyungJeong Yang*
Department of Computer Science
Chonnam National University, Gwangju 500- 757, South Korea

SunHee Kim
Department of Computer Science
Chonnam National University, Gwangju 500- 757, South Korea

SooHyung Kim
Department of Computer Science
Chonnam National University, Gwangju 500- 757, South Korea

ABSTRACT

Temporal Sign Language Recognition (TSLR) from hand motion is an active area of gesture recognition research in facilitating efficient communication with deaf people. TSLR systems consist of two stages: a motion sensing step which extracts useful features from signers' motion and a classification process which classifies these features as a performed sign. This work focuses on two of the research problems, namely unknown time varying signal of sign languages in feature extraction stage and computing complexity and time consumption in classification stage due to a very large sign sequences database. In this paper, we propose a combination of Dynamic Time Warping (DTW) and application of the Single hidden Layer Feedforward Neural networks (SLFNs) trained by Extreme Learning Machine (ELM) to cope the limitations. DTW has several advantages over other approaches in that it can align the length of the time series data to a same prior size, while ELM is a useful technique for classifying these warped features. Our experiment demonstrates the efficiency of the proposed method with the recognition accuracy up to 98.67%. The proposed approach can be generalized to more detailed measurements so as to recognize hand gestures, body motion and facial expression.

Keywords: *Dynamic Time Warping; Sign language; Single hidden layer feedforward neural networks; Time series analysis; Extreme Learning Machine; Back Propagation.*

1. INTRODUCTION

Sign languages are the primary form of communication between members of the deaf community [7]. However, these languages are not widely known outside of these communities, and hence a communication barrier exists between deaf and hearing people. Therefore, an automatic recognition system of human hand sign is important for a wide range of computer applications, from human computer interaction to the study of an athlete's performance [8].

Sign language recognition systems may be divided into two stages: a motion sensing step, which extracts useful movement

* *Corresponding author: E-mail: hjyang@jnu.ac.kr*
Manuscript received Jan.28, 2011 ; accepted Mar.18, 2011

* *This research was supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency)" (NIPA-2011-C1090-1111-0008). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (2010-0017364).*

data from a signer's motion, and a classification process, which classifies the movement data as a sign [8]. Since sign languages can be represented as corresponding time series or time–variable signals, it is not possible to compare each sign signal in Euclidean space directly because of misalignments in times. Signals of temporal sign languages have different durations because they are recorded from different people or even a same person is not ever able to reproduce movements exactly, and sign sampling rates may also be different.

Sign language can be stretched or compressed depending on both the speed of the movement and the signer. There are no continuous row to row correspondences between data of similar signs as shown in Fig. 1. In addition, the sizes of the sign sequence database are very large. It is a very time consuming work to perform exhaustive comparison through sequence alignment algorithms, by which two sign sequences are classified into the same class if they have high homology in terms of feature patterns extracted.

Our main goal is to find a combination of features and a classifier to maximize the classification accuracy. In this paper, we adopt DTW for warping data to align the length of the time series data with the same size. SLFN trained by ELM is used to classify time series sign languages by using the warped features of DTW as input data. A comparative study on classification performance is conducted among Extreme Learning Machine (ELM) [1], Least Square Extreme Learning Machine (LS-ELM) [3], Regularized least–squares ELM (RLS-ELM) [2] and Back propagation Neural Network with sigmoid activation function. Experimental results show that ELM is the best optimal classification system for the sign language recognition with DTW. Also compared with classical Back Propagation neural network, ELM achieves better performance with much shorter learning time. Specifically, experimental results show that the training speed of ELM is approximately 91 times faster than that of BP. The classification accuracy of ELM goes up to 98.67%. Two other ELM approaches, LS-ELM and RLS-ELM also obtain better performance with faster learning speed than BP network. In addition, network architecture of ELM does not have any control parameters such as learning rate, learning epochs, momentum, stopping criteria, etc., to be manually tuned.

The remainder of paper is organized as follows. Section 2 gives a brief review of related literatures. Section 3 describes two methods which are proposed for time alignments and classification of a sign language in this paper. Section 4 experimentally demonstrates the effectiveness of our proposed approach for sign language classification tasks. Section 5 concludes this paper.

2. RELATED WORKS

To recognize sign languages, many sign language recognition systems have been proposed. Wilson and Anspach [9] classified video images of hand shapes into their linguistic counterpart in American Sign Language (ASL). The video images were preprocessed to yield Fourier descriptors that encode the shapes of the hand silhouettes, which were then used as inputs to the neural network that classifies the shapes. Classification was performed for 36 hand shape data and achieved 78% accuracy.

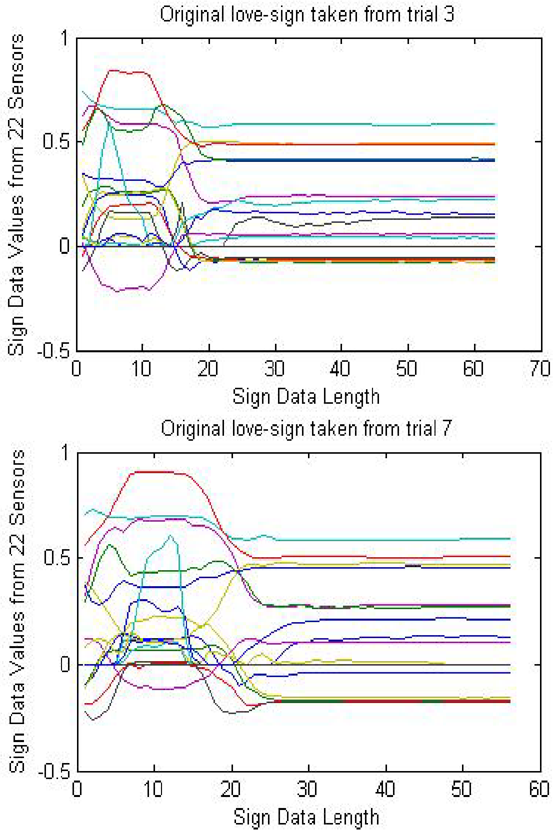


Fig. 1. Data for two love signs taken from trial 3 and trial 7. Same signs can have different lengths, and different corresponding attribute pairs can have different variations at different time.

Another approach was taken by Xiaoqing Weng and Junyi Shen who extracted feature of temporal sign languages by using two dimensional singular values decomposition [4]. They extended standard SVD by proposing a new approach called 2dSVD, the method captured explicitly the two dimensional nature of time series samples. The eigenvectors of row-row and column-column covariance matrices of time series samples were computed for feature extraction. After the feature matrix was obtained for each temporal sign sample, one nearest neighbor classifier was used for sign language classification. Experiment results gave 95% accuracy for classification.

Dynamic Time Warping (DTW) and Hidden Markov Models (HMMs) were two methods that simultaneously align signals and compute a likelihood of similarity [24], [10], [11], [12]. They have been applied successfully to recognize sign language, human gestures, speech, online or offline handwriting. In [13], a Multilayer Perception provided estimates of the emission probabilities for all phonemes of speech, subsequently used for matching HMM. In [14] a Neural Network classified the measurements of separate frames into a first and second guesses of a speech phoneme, and DTW used the phoneme matches with a template word as a distance measure. In [15], the measurements for a frame of a gestured command, recorded by a camera, were converted into a probability estimate of each state by a Radial Basis Function network. The resulting state emission probabilities were used for HMM. In [16], Chinese sign language was measured with data gloves. Signs that were not well separated by HMM alone were classified in an extra recognition step by a Support Vector Machine (SVM) using a DTW kernel. In [17], a sequential HMM was trained for each hand gesture measured from two cameras. HMM match result was split into five components which were used as features for a multiclass SVM classifier, trained by applying one HMM to all training gestures. The final classification was obtained by majority voting of the results of the HMM/SVM pairs for all gesture classes. A similar approach was chosen in [18] to classify online hand writing characters. The above works can be improved over HMM or DTW alone. However, these methods relied directly on the likelihoods obtained from DTW or HMM.

Instead, in [6] Jeroen F. Lichtenauer et al. developed a system to avoid conflicting likelihood modeling demands. They combined statistical DTW (SDTW) and two single model classification methods, namely Combined Discriminative Feature Detectors (CDFD) and Quadratic classification on DF Fisher Mapping (QDFFM), on a set of 120 different signs of the Dutch Sign Language. They also showed that SDTW provides a significant improvement over HMM. The system by combining method SDTW & DFFM gave an average accuracy of 92.3%.

In [19] Blaz Strle, Martin Mozina, and Ivan Bratko proposed a new algorithm, Qualitative Dynamic Time warping (QDTW) which was a modifier of DTW algorithm. They demonstrated how well QDTW performs in comparison to DTW and how different persistence settings affect classification accuracy using weighted k-nearest neighbor. Classification accuracy performed on ten signs from Australia sign language dataset was 0.72 using persistence is 0.1.

From these experiments above, this paper emphasizes on developing a classification system that identifies effectively sign language sequences. We first apply DTW on sign language sequences for performing the time alignment. We then use the latest machine learning ELM system for classifying based on complete warped features from DTW.

3. PROPOSED METHODS

This section describes the proposed method of dynamic time warping and ELM for warped feature extraction and sign classification.

3.1 Dynamic Time Warping

In this section, we briefly describe DTW for aligning duration of sign language sequences. DTW is a pattern adjustment algorithm that measures the similarity between two sequences nonlinearly in the time dimension. It finds similarities of certain nonlinear variations in the time dimension between two patterns [20]. Optimal alignment which is a minimum distance of warp path is obtained by allowing assignment of multiple successive values of one time series to a single value of the other time series. Therefore, DTW can also be calculated on time series of different lengths, shown in Fig. 2 [19].

By using DTW, optimal alignment is found among several different warp paths. This can be presented through two given time series arrays $A = (a_1, a_2, \ldots, a_n)$ of length n and $B = (b_1, b_2, \ldots, b_m)$ of length m, where $a_i, b_j \in R$ are arranged to form a n by m grid. Each grid point corresponds to an alignment between elements $a_i \in A$, $b_j \in B$. This is illustrated in Fig. 2. Therefore, a $n \times m$ matrix D is constructed, where the (i^{th}, j^{th}) element of the matrix contains the distance $d(a_i, b_j)$ between the two points a_i and b_j (with Euclidean distance, $d(a_i, b_j) = (a_i - b_j)^2$). A warp path $W = w_1, w_2, w_3, \ldots, w_k$ where w_K is a sequence of grid points, and each w_k corresponds to a point $w_k = (i, j)_k$. Warp path W maps elements of two time series A and B.

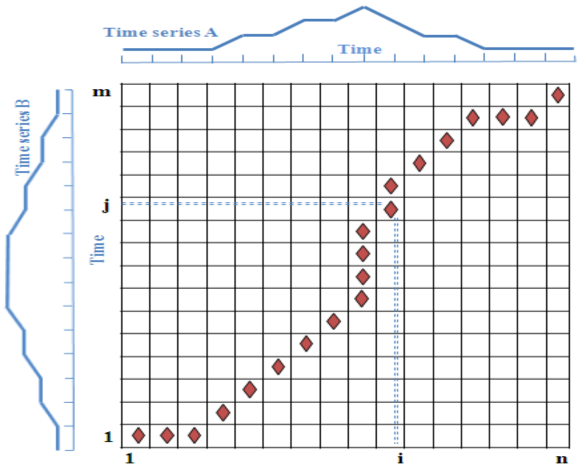


Fig. 2. An example of minimal warping path between two time series A and B.

The warping path must satisfy several constraints as follows:

Boundary conditions: $w_1 = (1, 1)$ and $w_k = (n, m)$. This requires the warping path to start in the first point of both sequences and end in the last point of both sequences.

Continuity: Given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a - a' \leq 1$ and $b - b' \leq 1$. This restricts the allowable steps in the warping path to adjacent cells (including diagonally cells).

Monotonicity: Given $w_k = (a, b)$ then $w_{k-1} = (a', b')$ where $a - a' \geq 0$, and $b - b' \geq 0$. This forces the points in W to be monotonically spaced in time.

The goal of DTW is to find the minimal distance of warp path between two time series A and B for an optimal duration. There are many warping paths that satisfy the above conditions. We are only interested in the path that minimizes the warping cost as in Eq. (1) [21]:

$$DTW(A, B) = \min\left\{\sqrt{\sum_{k=1}^{K} w_k}\right. \quad (1)$$

This path can be calculated using dynamic programming method as follows:

$$D(i,j) = d(i,j) + \min\begin{cases} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1) \end{cases} \quad (2)$$

Where $1 \leq i \leq n$, $1 \leq j \leq m$ and $d(i, j)$ is the distance of the two time series. D(1, 1) is initialized by d(1, 1).

The above approach deals with the stretching and compressing effects of temporal sign languages for alignment sign language duration.

3.2 Classification of warped features using Singular Hidden

Layer Feedforward Neural Network trained by ELM

After DTW extracts warped features from a sign matrix, SLFNs which are trained by ELM classify the features. All the feature vectors of training sign datasets are used as inputs to SLFNs classifiers for training. It is known that the problem facing in application of neural network is training algorithm. Traditionally all the parameters of the feedforward networks need to be tuned and thus there exists the dependency between different layers of parameters such as weights and bias. For past decades, gradient descent based methods have mainly been used in various learning algorithms of feedforward neural networks. However, it is clear that gradient descent based learning methods are generally very slow due to improper learning steps or may easily converge to a local minima. Besides, many iterative learning steps are required by such learning algorithms in order to obtain better learning performance.

Huang [1] proposed a new learning algorithm called Extreme Learning Machine (ELM) for Single hidden Layer Feedforward neural Networks (SLFNs) which are either additive neurons or kernel based schemes. In this paper, we use SLFNs with additive activation function such as sigmoid. For additive neuron based SLFNs one may randomly choose and then fix input weights which link the input layer to the hidden layer and the hidden neurons' biases and analytically determine the output weights which link the hidden layer to the output layer of SLFNs[22]. Input weights are the weights of connections between input neurons and hidden neurons. Output weights are the weights of the connections between hidden neurons and output neurons. After the input weights and the hidden layer biases are chosen arbitrarily, SLFNs can be simply considered as a linear system and the output weights which link the hidden layer to the output layer of SLFNs can be analytically determined through simple generalized inverse operation of the hidden layer output matrices. As analyzed by Huang, ELM tends to have good generalization performance and can be implemented easily. Unlike other tuning or adjustment methods which may neither be suitable for non differential activation functions nor prevent the troubling issues such as stopping criteria, learning rate, learning epochs, and local minima, the ELM algorithm can avoid these difficulties very well.

In generally, the architecture of neural networks is organized into an input layer, one or many hidden layers and one output layer. However, the single hidden layer feedforward neural networks in Fig. 3 [22] (SLFNs) can learn exactly N distinct observations with the proper choice of activation function and the number of hidden units [6]. The (SLFNs) form boundaries with arbitrary shapes and approximate any function with arbitrarily small error if the activation function is chosen properly. The architecture of SLFNs is shown in Figure 3.

For given S arbitrary distinct samples $S=\{(x_j, t_j)|\ j=1,2,\cdots,n\}$ where $x_j = [x_{j1}, x_{j2}, \cdots, x_{jd}]^T$ and $t_j = [t_{j1}, t_{j2}, \cdots, t_{jC}]^T$ are the jth input patterns and its target, respectively. Hidden layer outputs of Standard SLFNs with d input units, N hidden neurons, and C output units are mathematically modeled as below:

$$h_j=[f(w_1 \cdot x_j+b_1), f(w_2 \cdot x_j+b_2), \ldots, f(w_N \cdot x_j+b_N)]^T \qquad (3)$$

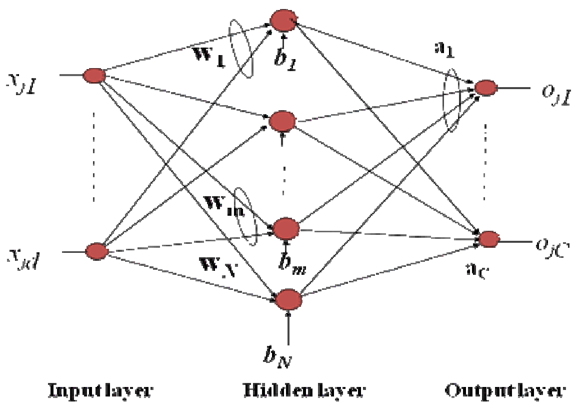


Fig. 3. Architecture of Single hidden layer feedforward neural network

Where $w_m = [w_{m1}, w_{m2}, \ldots, w_{md}]^T$ is the weight vector connecting the mth hidden neuron and the input neurons, b_m is its bias and $f()$ is the activation function, $w \bullet x = \langle w,x \rangle$ denotes the inner product of two vectors w and x.

The ith output corresponding to the jth input patterns x_j is given by Eq. (4).

$$o_{ji} = h_j \bullet a_i, \qquad (4)$$

Where $a_i = [a_{i1}, a_{i2}, \ldots, a_{iN}]^T$ is the weight vector connecting from the hidden unit to the ith output neuron, h_j is hidden layer output vector determined above.

The main purpose of training process is to determine network parameters that minimize error function defined in Eq.(5):

$$E = \sum_{j=1}^{n}(o_j - t_j)^2 = \sum_{j=1}^{n}\sum_{i=1}^{C}\left(h_j \bullet a_i - t_{ji}\right)^2 \qquad (5)$$

This is equivalent to the solution of linear system in Eq.(6):

$$FA = T, \qquad (6)$$

Where F is called the hidden layer output matrix defined as in Eq. (7) and Eq. (8)[2].

$$\begin{aligned} F &= [h_1 h_2 \ldots h_n]^T \\ &= \begin{bmatrix} f(w_1 \cdot x_1 + b_1) & \cdots & (w_N \cdot x_1 + b_N) \\ \vdots & \ddots & \vdots \\ f(w_1 \cdot x_n + b_1) & \cdots & (w_N \cdot x_n + b_N) \end{bmatrix} \end{aligned} \qquad (7)$$

$$A = [a_1\, a_2 \ldots a_C] \quad T = [t_1\, t_2 \ldots t_n]^T \qquad (8)$$

In ELM, the input weights and hidden biases are randomly generated instead of tuned. Therefore, the nonlinear system is converted to a linear system shown in Eq. (6).

The determination of the output weights is as simple as finding the least square solution to the given linear system. The minimum norm least square (LS) solution to the linear system in Eq. (6) is as follows:

$$A = F^{\dagger}T$$

Where $F^{\dagger}$ is the Moore Penrose generalized inverse of matrix F. The minimum norm LS solution is unique and has the smallest norm among all the LS solutions. As analyzed by Huang, ELM using such Moore Penrose inverse method tends to obtain good generalization performance with dramatically increased learning speed.

4. PERFORMANCE EVALUATION

In this paper, the experiment is carrying out on real world dataset Australia Sign language [25]. Below gives a brief description of dataset.

AUSLAN dataset contains 95 different sign languag words such as alive, all, answer, boy, building, by, change, col come and so on [19]. Each of which is repeated 27 times whi mean 27 samples per sign. Hence there are totally 2565 sampl in AUSLAN dataset. Samples of each sign were captured from native AUSLAN speaker using 22 sensors on the Cyber Glov Each sample can be regarded as a multivariate time series wi 22 variables. The average length of each sample is around (time points. 27 samples per sign can be regarded as one clas As an example, time series describing the movements computer sign language on different samples taken from sens 3 are shown in Fig. 4. As we see, although the same sign, tl lengths of this sign on different samples are different. Tl lengths of computer sign on the 1st, 7th, 9th, 12th and 17th samp are 53, 71, 62, 58 and 81, respectively. In this study, dataset divided into two parts: training set and testing set. The trainii set is 70% of the pattern data set and the remaining 30% is us as testing set. The training set is used to build an optimal neu network. The test set is used to evaluate the performance of this network for recognition sign patterns.

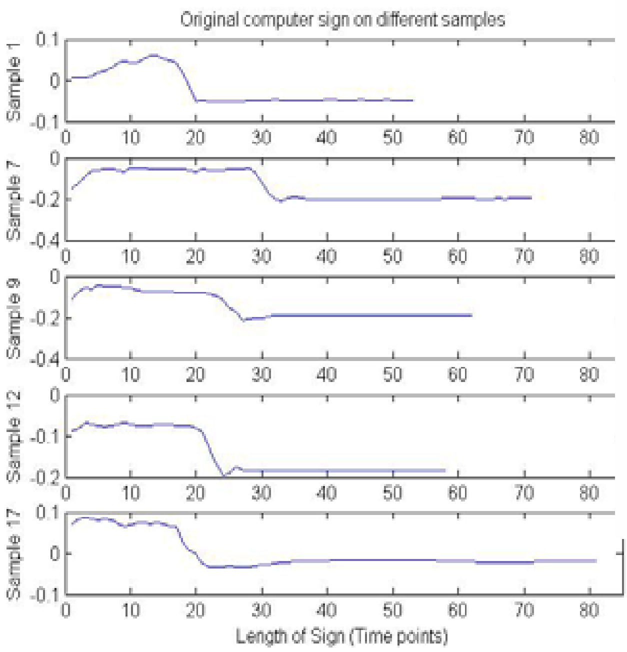
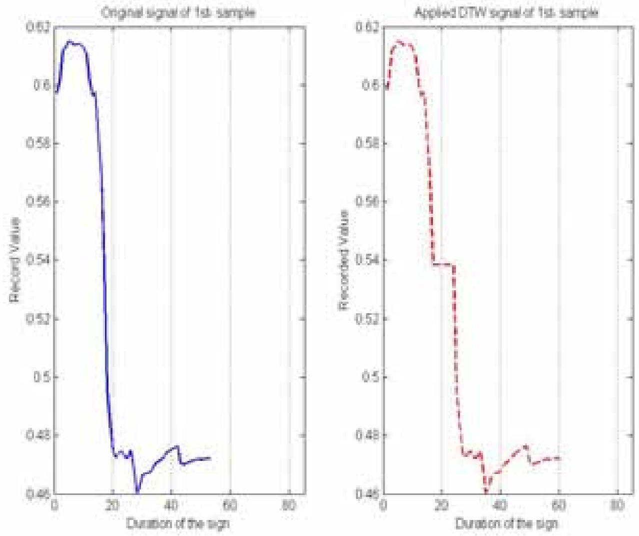
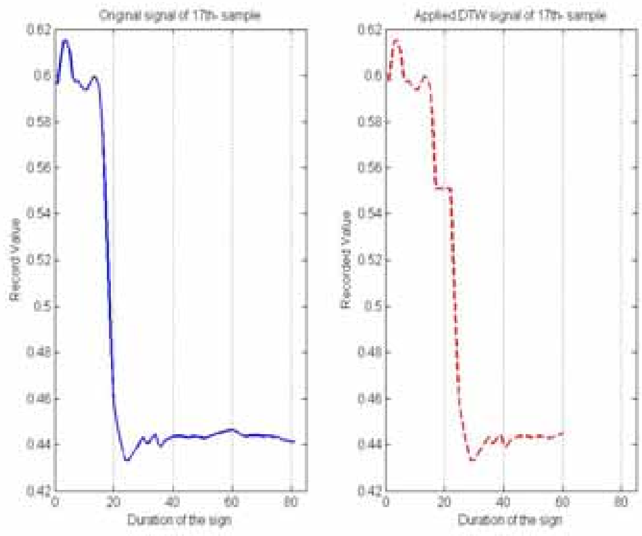


Fig. 4: Original computer sign taken from different samples.

On this study, the proposed DTW method for aligning different durations of the sign languages is described below. Taken Computer sign languages on sensor 4 only as an example, there are 53 time ticks in the first sample of this sign and 81 time ticks in the 17th sample shown on Fig. 5. In Fig. 5 (a) and (c) are original Computer signal and (b) and (d) are Computer signal after applying DTW on the first sample and 17th sample. As we see, although the same sign, the durations of the same Computer signs are different. Therefore, after using DTW to align the duration of Computer sign language data to be equal length for each other, the duration is 60 time ticks. Fig. 5 shows that there's no significant change in shape before and after applying DTW. Therefore, DWT solves the time scaling problem of different durations of sign language dataset.



(a) (b)



(c) (d)

Fig. 5. An example by applying DTW on a sign to align the duration from 53 to 60 and 81 to 60, respectively.

In the classification stage, the training algorithms are implemented on Matlab 7.8.0 environment with version 2009a. All algorithms of SLFNs use sigmoid activation function in expression (9)

$$f(x) = \frac{1}{1+e^{-x}} \quad (9)$$

All simulation results are averaged over 50 trials. It was found during our simulations that BP required large memory and for this application so that it would be out of memory when more than 20 hidden neurons are assigned to the BP network.

Table 1. Accuracy of ELM classifier

Training Time(s)	**Testing Time (s)**	**Training Accuracy (%)**	**Testing Accuracy (%)**	**Number of hidden Neurons**
5.296	0.0600	96.911	95.933	500
8.093	0.1041	97.933	95.733	600
10.828	0.1202	98.511	97.866	700
15.640	0.1250	99	98	800
20.062	0.1666	99.244	98.066	900
27.015	0.1771	99.288	98.667	1000
42.468	0.2968	99.600	96.933	1200

Table 2. Accuracy of BP classifier

Training Time (seconds)	**Training Accuracy (%)**	**Testing Accuracy (%)**	**Number of hidden Neurons**
1180.656	88.756	87.533	10
1586.984	92.222	89.666	13
1989.343	95.067	91.667	15
2309.125	95.266	93.466	17
2460.421	95.933	94.200	18
2670.250	96.177	94.200	19
Out of memory			20

Table 3. Performance comparisons of different classifiers

Me-thod	Training time (s)	Testing time (s)	Accuracy (%)		# hidden nodes
			Training	Testing	
ELM	27.015	0.1771	99.288	98.667	1000
RLS-ELM	40.171	0.4062	98.533	96.866	1200
LS-ELM	39.109	0.3437	96.400	94.600	1200
BP	2460.4	4.8482	95.933	94.200	18

The performance comparison between training algorithms for singular hidden layer feedforward neural network such as ELM, LS-ELM, RLS-ELM and BP are reported in Table 1, Table 2, and Table 3. In singular hidden layer feedforward neural network, the activation function sigmoid is used. The input features were normalized into range [-1 1].

As shown in Table 3, training time, testing time and the number of hidden neurons for four different training methods are presented, respectively. ELM obtains the best average performance 99.288% and 98.667% for training and testing, respectively. It shows the second fastest training time among four algorithms of ANN which is 27.015 while BP gives the slowest performance with 2460.421 seconds of learning time. The accuracies are 95.933% for training and 94.2% for testing. The graphs in Fig. 6 and Fig. 7 show the number of hidden neurons versus the accuracies for testing of BP and ELM classification. The relationship between learning time and the number of hidden neurons for ELM and BP are shown in Fig. 8 and Fig. 9, respectively.

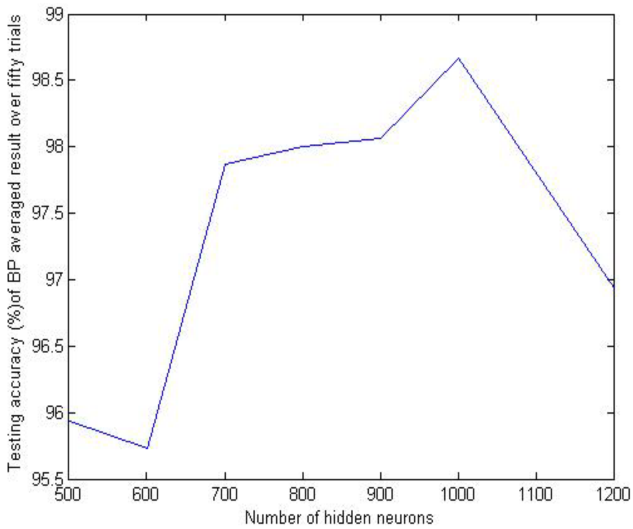


Fig. 6. Accuracy vs. the number of hidden neurons of ELM classification

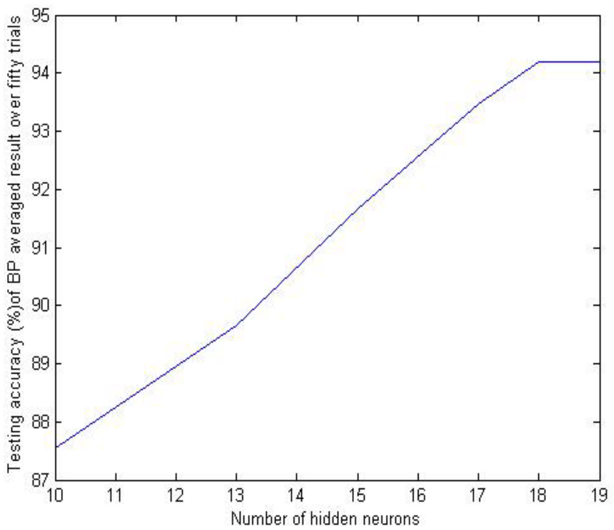


Fig. 7. Accuracy vs. the number of hidden neurons of BP classification.

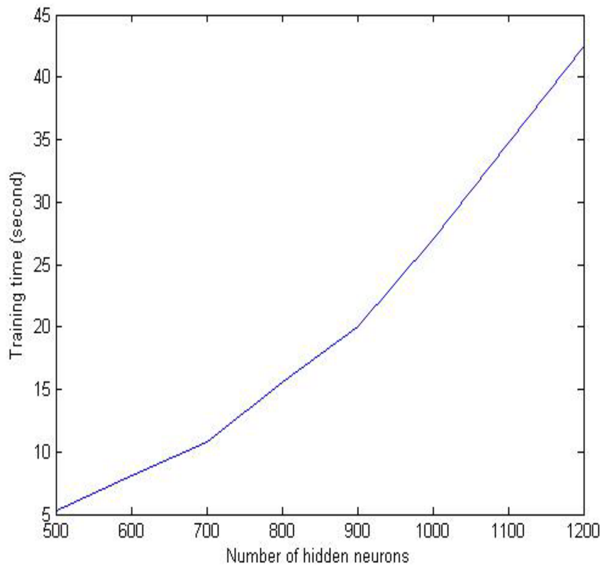


Fig. 8. Training time vs. the number of hidden neurons of ELM.

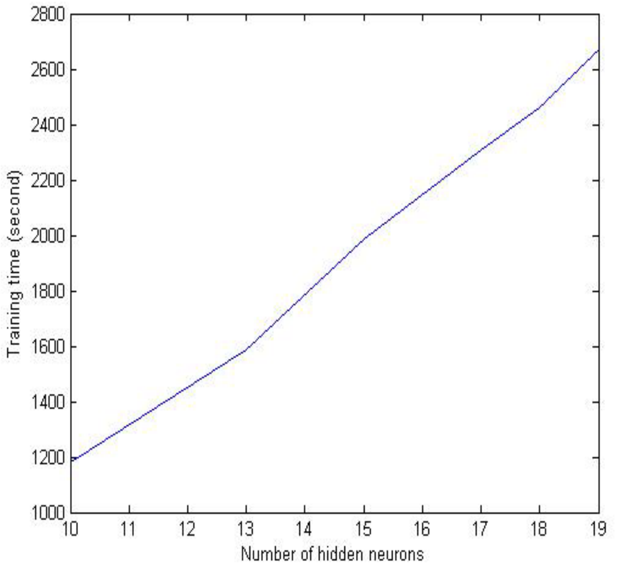


Fig. 9. Training time and the number of hidden neurons of Back propagation.

The comparison among ELM and two other ELM approaches which are LS-ELM and RLS-ELM is shown in the Fig. 10. Unlike [2][3], on this study LS-ELM and RLS-ELM cannot improve ELM algorithm on both accuracy and learning time. However, both of them demonstrate the accuracy are better than BP algorithms with 94.6% and 96.8%, respectively. They also can run approximately 61 time faster than BP algorithm in the case when the best generalization performances are obtained for all of them. The comparison of relationship between performance and the number of hidden neurons of three methods ELM, LS-ELM and RLS-ELM for testing accuracy is shown in Fig. 10.

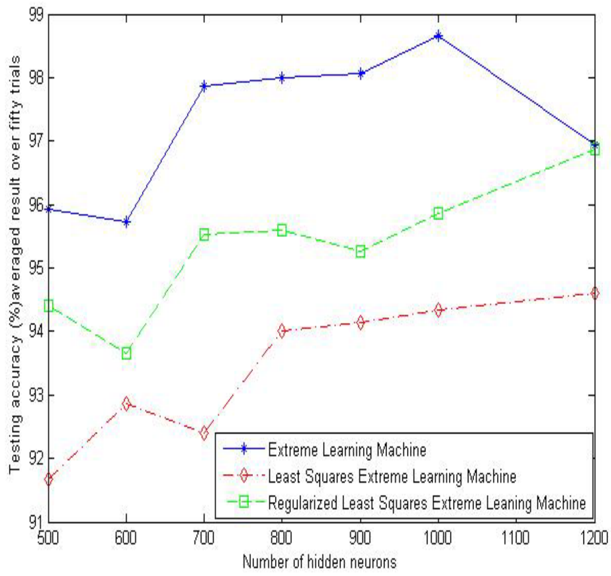


Fig. 10. Comparison of the classification accuracy according to the number of hidden neurons of three methods.

Among four algorithms of SLFNs demonstrated above, ELM algorithm works generally best for temporal sign language recognition. ELM network system shows that the accuracy for sign language classification is up to 98.667% and training times and testing time are just about 27.015 seconds and 0.1771 seconds. Besides, referring to related works section, our proposed method demonstrates a reasonable accuracy in classification which is higher than previous reported classification results as Xiaoqing Weng [4] demonstrated the accuracy for classification on 25 signs of Australia sign language dataset is 95%.

5. CONCLUSION

In this paper, Extreme Learning Machine (ELM) combining with DTW has been employed for the recognition of temporal sign languages. The study demonstrates the advantage of using DTW is to normalize time variable signals of sign language as a fixed size feature set for feature extraction. After the warped features are obtained for each temporal sign language, ELM classifier is used for classification. The experimental results show that the sign language recognition via temporal classification can be carried out with an accuracy rate as high as 98.67% with learning time is just about 27.015 times in second. We compared the performance of ELM together with three popular approaches of ANN, namely BP, LS-ELM and RLS-ELM on classification of temporal sign language with twenty super families to find the best optimal classification system. This study demonstrates that ELM needs much less training time compared to three remainder networks. The performance of ELM is better than LS-ELM, RLS-ELM and BP in this application on both classification accuracy and learning time. In addition, compared with BP, ELM can be implemented easily since there is no parameters to be tuned expecting for network size which is common to BP.

REFERENCES

[1] G. B. Huang, Q. Y Zhu and C. K. Siew, "Extreme learning machine: Theory and application," Nero computing, vol. 70, May. 2006, pp. 489 – 501.

[2] H.T. Huynh, Y. Won, and J.J. Kim, "An improvement of extreme learning machine for compact single hidden layer feedforward neural networks," International journal of neural systems, vol. 18, no. 5, 2008, pp. 433- 441.

[3] H.T. Huynh and Y. Won, "Small number of hidden units for ELM with two-stage linear model," IEICE Trans. On Information and Systems, vol. E91.D, Issue. 4, 2008, pp. 1042- 1049.

[4] X.Weng, Junyi Shen, "Classification of multivariate time series using two dimensional singular value decomposition," Knowledge-Based Systems, vol. 21, Issue 7, 2008, pp. 535-539.

[5] G. B. Huang, "Learning capability and storage capacity of two hidden layer feedforward networks," IEEE Transactions on Neural Networks, vol. 14, no. 2, 2003, pp. 274- 281.

[6] J. F. Lichtenauer, Emile A. Hendriks, and Marcel J. T. Reinders, "Sign Language Recognition by combining Statiscal DTW and Independent Classification," IEEE transactions on pattern analysis and machine intelligence, vol.30, no. 11, Nov. 2008, pp. 2040- 2046.

[7] P. Vamplew, "Recognition of sign language gestures using neural network," Proc. 1st Euro. Conf. Disability Virtual Reality & Assoc. Tech, Maidenhead, UK, 1996, pp. 27- 33.

[8] E. Jung Holden, Geoffray G. Roy, Robyn Owens, "Hand movement Classification an adaptive fuzzy expert system," Intl. J. Expert Systems, 1996, pp. 465- 480.

[9] E. Wilson, & G. Anspach, "Neural networks for sign language translation," SPIE: Applications of Artificial Neural networks, pp. 589- 599.

[10] D. Gavrila and L. Davis, "Towards 3D Model Based Tracking and Recognition of Human Movement: A Multi View Approach," Proc. IEEE Int'l Workshop Face and Gesture Recognition, June 1995, pp. 272- 277.

[11] A. Corradini, "Dynamic Time Warping for OffLine Recognition of a Small Gesture Vocabulary," Proc. IEEE ICCV Workshop Recognition, Analysis, and Tracking of Faces and Gestures in RealTime Systems (RATFGRTS '01), July 2001, pp. 82- 89.

[12] S. Yang and R. Sarkar, "Gesture Recognition Using Hidden Markov Models from Fragmented Observations," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR '06), 2006, pp. 766- 773.

[13] N. Morgan and H. Bourlard, "Continuous Speech Recognition Using Multilayer Perceptrons with Hidden Markov Models," Proc. Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP '90), 1990, pp. 413- 416.

[14] Y. Matsuura, H. Miyazawa, and T. Skinner, "Word Recognition Using a Neural Network and a Phonetically Based DTW," Proc. IEEE Int'l Workshop Neural Networks for Signal Processing (NNSP '94), Sep. 1994, pp. 329- 334.

[15] A. Corradini and H. Gross, "Camera Based Gesture Recognition for Robot Control," Proc. Int'l Joint Conf. Neural Networks (IJCNN '00), July 200, pp. 133- 138.

[16] J. Ye, H. Yao, and F. Jiang, "Based on HMM and SVM Multilayer Architecture Classifier for Chinese Sign Language Recognition with Large Vocabulary," Proc. Third Int'l Conf. Image and Graphics (ICIG '04), Dec. 2004, pp. 377- 380.

[17] O. Aran and L. Akarun, "Recognizing Two Handed Gestures with Generative, Discriminative and Ensemble Methods via Fisher Kernels," Proc. Int'l Workshop Multimedia Content Representation, Classification and Security (MCRCS '06), Sep. 2006, pp. 159- 166.

[18] C. Bahlmann, B. Haasdonk, and H. Burkhardt, "Online Handwriting Recognition with Support Vector Machines- A Kernel Approach," Proc. Eighth Int'l Workshop Frontiers in Handwriting Recognition (IWFHR '02), 2002, pp. 49- 54.

[19] B. Strle, Martin Mozina, Ivan Bratko, "Qualitative approximation to Dynamic Time Warping similarity between time series data," 23rd Annual Workshop on Qualitative Reasoning, QR– LJUBLJANA SLOVENIA, June 2009, pp. 22- 24.

[20] S.H. Kim, Hyung Jeong Yang, Kam Swee Ng, "Temporal Sign Language Analysis Based on DTW and Incremental Model," MICC 2009, Dec. 2009, pp. 14- 16.

[21] E. Keogh, "Exact Indexing of Dynamic Time Warping", Very Large Database, 2002, pp. 406- 417.

[22] H.T. Huynh and Y. Won, "Hematocrit Estimation from compact single hidden layer feedforward neural networks trained by evolutionary algorithm," IEEE Congress on Evolutionary Computation, 2008, pp. 2962- 2966.

[23] G.B, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," Proc. of International Joint Conference on Neural Networks (IJCNN2004), Jun. 2004, pp.489- 501.

[24] T. Starner, Visual Recognition of American Sign Language Using Hidden Markov Models, master's thesis, Massachusetts Inst. of Technology, Media Arts and Sciences, Jan. 1995.

[25] http://archive.ics.uci.edu/ml/datasets.html

Ngoc Anh Nguyen Thi

She received the B.S, in Faculty Mathematics-Informatics from Da Nang Education University, Viet Nam in 2006. She is currently a master student at Dept. of Electronics and Computer Engineering, Chonnam National University, Korea. From 2006 to 2008, she worked as a lecturer and researcher at Faculty Mathematics-Informatics of Da Nang Education University, Viet Nam. Her research interests focus on the intelligent computing in many applications such as pattern recognitions, bioinformatics, data analysis of data mining and machine learning.

HyungJeong Yang

She received the B.S, M.S and Ph. D from Chonbuk National University, Korea. She is currently an assistant professor at Dept. of Electronics and Computer Engineering, Chonnam National University, Gwangju, Korea. Her main research interests include Multimedia data mining, Pattern Recognition, Artificial Intelligent, e- Learning and e- Design.

SunHee Kim

She received the B.S, M.S in Dongguk University, Korea. She is currently a Ph.D. student at Dept. of Engineering and Computer Engineering, Chonnam National University, Korea. Her research interests focus on data mining, sensor mining and stream mining.

SooHyung Kim

He received the B.S at Dept. of Computer Engineering, Seoul National University, M. S and Ph. D at Dept. of Computer Science, Korea Advanced Institute of Science and Technology, Korea. He is currently a professor at Dept. of Electronics and Computer Engineering and a vice-Dean of Engineering College, Chonnam National University, Gwangju, Korea.