'95 KFIS Workshop, Ik-San, Korea

On-line Korean Sign Language(KSL) Recognition using Fuzzy Min-Max Neural Network and Feature Analysis

Zeungnam Bien and Jong-Sung Kim
Dept. of Elec. Eng., KAIST 373-1 Gusung-dong Yusung-gu, Taejon, 305-701, KOREA ,
zbien@ee.kaist.ac.kr, joskim@ctrsys.akist.ac.kr

Abstract

This paper presents a system which recognizes the Korean Sign Language(KSL) and translates into normal Korean speech. A sign language is a method of communication for the deaf-mute who uses gestures, especially both hands and fingers. Since the human hands and fingers are not the same in physical dimension, the same form of a gesture produced by two signers with their hands may not produce the same numerical values when obtained through electronic sensors. In this paper, we propose a dynamic gesture recognition method based on feature analysis for efficient classification of hand motions, and on a fuzzy min-max neural network for on-line pattern recognition.

1 Introduction

Human-hand gestures have been used as a means of communication among people for a long time, being interpreted as streams of tokens for a language [1]. They may vary from the stylized lexicon of a traffic cop to the highly developed syntax of a natural language such as the sign language.

In this paper are considered static hand posture and dynamic hand gesture. The static hand posture is specific value or range of values of the degrees of freedom or subsets of the degrees of freedom of the hand, or specific values of the features of the hand. Dynamic hand gesture is derived from motion of the fingers and hand, or from continuous features [1].

The sign language which includes many hand gestures is a method of communication for hearing/speech impaired persons. It is understood by means of gestures of both hands and fingers [2].

This paper deals with a system which recognizes the Korean Sign Language(KSL) and translates it into a normal Korean speech.

According to a standard KSL dictionary [3], the Korean Sign Language with 45 years of history contains about 6,000 vocabularies. However, they are formed by combining a relatively small number of basic gestures.

The recognition of changing patterns of dynamic gestures in the time domain is essential to understand any KSL-based sentences. This means that the recognition of the KSL should be conducted in real-time. For our system, an electronic device, called Data-Glove Model 2+ system[4], is adopted as an input device in consideration of cost effectiveness of hardware versus real-time processing capability. It is also known that the pattern classes of KSL gestures are not linearly separable and that patterns tend to overlap with each other. Therefore, it is desirable to design a pattern classifier in such a way that the amount of mis-classification for those overlapping classes is minimum. Also, the system needs some form of learning capability due to the varying nature of the patterns to handle.

It is remarked that in [5] and [6; 7], neural network based methods were presented for recognition of the American Sign Language(ASL). In the work by [5] were used the back-propagation neural networks for recognition of simple gestures(derived from the ASL), but it seems that extensive training is required with explicit specification of the beginning and end of the gesture. For abrogating the need for extensive training in the neural network, a methods of relabeling a self organizing map(SOM) is proposed in [6]. This method is proposed to avoid the usual heavy work of retraining when new vocabulary words are added to the system though this method also requires to train the system for recognition of initial classes. Also if the number of new vocabularies to be added gets increased, the system may fail to

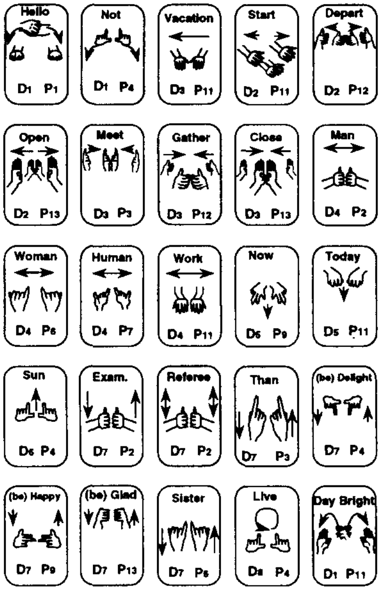


Figure 1: Examples of Korean Sign Languages

show the same success rate.

In this paper, we propose a dynamic gesture recognition method by employing feature analysis for efficient classification of hand motions, and applying a fuzzy min-max neural network [8] for on-line pattern recognition.

2 Overview of the Korean Sign Language

The Korean Sign Language(KSL) is a method of communication for hearing impaired persons. The KSL is understood by means of gestures of (both) hands and fingers. For example, gesture '$man(Nam - Ja\ in\ Korean)$' is expressed by unfolding the thumb and gesture '$human(Sa - Ram\ in\ Korean)$' is expressed by unfolding the thumb and a little finger with swing motion [9]. Fig. 1 shows examples of the KSL.

2.1 The Korean Sign Language(KSL)

The Korean standard Sign Language is based on Korean Grammar and systematized by many KSL researchers in 1991. Based on shape of objects and association, the KSL is communicated via specified hand gestures [2].

To express the KSL word or sentence, all of the two hands and ten fingers may be used. The right hand performs main motional actions while the left hand is often employed for auxiliary purpose. That is, most of the left hand motions are usually symmetrically the same as those of the right hand [10].

2.2 Analysis of the KSL

To recognize the KSL, characteristics of the KSL are investigated in pattern recognition aspects.

The KSL contains about 6,000 gestures [3] but most of the gestures are formed by combining some basic gestures. In this paper, we have selected important basic gestures shown in Fig. 1 from the KSL textbook [11] for our pilot study.

It is considered two principles for selecting gestures.

i) All of gestures are used in x-y/x-z plane.

ii) Compound words and gestures with multiple motions are not contained in this paper.

In this paper, dynamic hand gesture is expressed by

$$
\begin{aligned}
G_k &= (x_k, y_k, z_k, t_k, F_{ki}) \ , && (1) \\
& 1 \leq k < K \ : \ \text{time step [1/15 sec.]} && (2) \\
& F_{ki} \ : \ \text{flex angles} \ , \ i = 1, \cdots, 10 &&
\end{aligned}
$$

It is noted that the KSL contains two elements, hand motion and hand posture. Human checks the change of signer's hand motion and recognizes the shape of signer's hand at that time for recognizing the KSL gesture.

The KSL has globally about 4 direction types, horizontal, vertical, slant, and circular motion. After analyzing these gestures, one can easily conclude that these gestures contain 14 basic direction types of hand motion patterns shown in Fig. 2 based on above 4 direction types. And 14 types of basic hand postures are included as shown in Fig. 3. In this paper for recognizing gestures, a static hand posture contains only flex angles of signer. An orientation values are unnecessary to classify these gestures.

The FMMN network [8] is applied to classify 14 hand postures.

A classification of 14 basic direction type is done by feature analysis method. The features are the direction change in sampled latest 3 points, the cumulative length in hand motion, difference of left and right position, current phase, and cumulative direction change in radian.

3 KSL Recognition System

The VPL Data-Glove [4] put on a hand in motion generates 16 types of data. That is, 10 flex angles, 3 position data(x, y, z), and 3 orientation data(roll, pitch, yaw) are obtained from each of

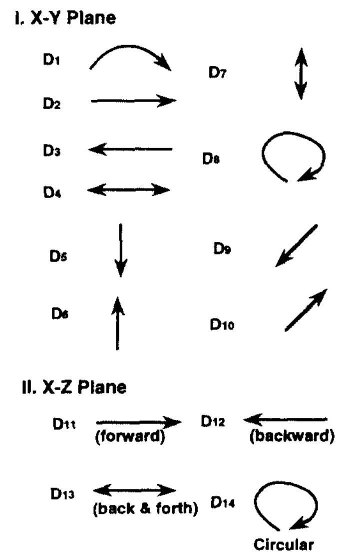


Figure 2: 14 Basic Direction types

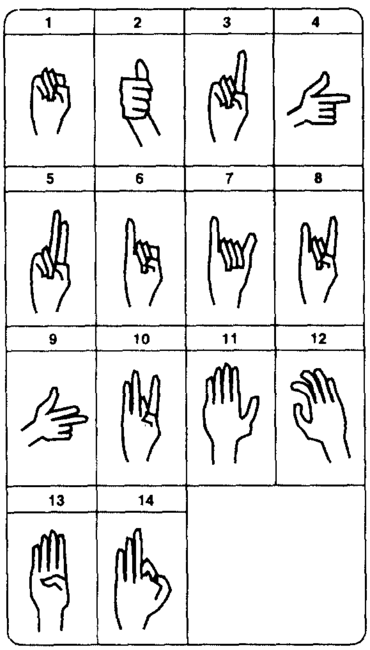


Figure 3: 14 Postures

the two Data-Gloves. Two sets of the 16 parameters from the right and left hand Data-Gloves, however, are often heavily contaminated by external signals generated during electronic sensing and/or caused by signer's unnecessary motions including unconscious hand trembling. It is desired to suppress those irregularly generated external signals.

For recognition of the KSL, these kinds of raw data are processed to generate some features of the hand motion.

In the following, we describe the process of recognizing the KSL in a more detailed manner.

3.1 Setting of Initial Position

To initiate any dynamic gesture, there should be some starting point in space. Therefore, it is needed for the recognition system to be independent of the initial position of a gesture. For this, the initial x and y-axis data are recorded from the Data-Glove and all the subsequent position data of the Data-Glove are calibrated by continuously subtracting the initial position data from the current position data(see Eq. 3.)

$$(\ x(k),\ y(k),\ z(k)\ )\ -\ (\ \mathbf{x}_s,\ \mathbf{y}_s,\ \mathbf{z}_s\ ) \quad (3)$$
$$k = 1, 2, \cdots\ :\ \text{time step [1/15 sec.]}$$
$$(\mathbf{x}_s, \mathbf{y}_s, \mathbf{z}_s)\ :\ \text{starting position}$$

It is remarked that our coordinate system is determined in reference to the mechanism of the Data-Glove Model 2+ system [4] and, thus, the vertically up/down hand motion is called an y-axis motion while the x-axis motion occurs horizontally left/right, and backward/forward motion is appeared along the z-axis(refer to Fig. 7)

3.2 Classification of Hand Motion by Feature Analysis Method

As mentioned earlier, the input data obtained by the gloves are heavily contaminated by distortion due to ferrous metallic substances in polhemus 3-D sensor system and signer's unconscious hand trembling.

1) Reduction of hand trembling.

An input position within small deviation(1 inches) of the previous input position is discarded.

$$
\begin{aligned}
l1(k) &= \sqrt{dx(k)^2 + dy(k)^2} \qquad (4)\\
l2(k) &= \sqrt{dx(k)^2 + dz(k)^2}
\end{aligned}
$$
$$:\ \text{length of path-segment}$$
$$dx(k) = x(k) - x(k-1)$$
$$dy(k) = y(k) - y(k-1)$$

$$dz(k) = z(k) - z(k-1)$$

$$\text{if} \quad l(k) < \text{Threshold value}(\delta_1 = 1"),$$
$$x(k) = x(k-1),\ y(k) = y(k-1),$$
$$z(k) = z(k-1)$$

2) Reduction of magnetic interference

An change of direction within small deviation(1 inches) of the previous input position is discarded.

$$\text{if } |dx(k)| < \delta_1(1"),\ x(k) = x(k-1) \quad (5)$$
$$\text{or, if } |dy(k)| < \delta_1(1"),\ y(k) = y(k-1)$$
$$\text{or, if } |dz(k)| < \delta_1(1"),\ z(k) = z(k-1)$$

After these preprocessings, x and y-axis data are inputted to system every time step(1/15 sec.). As explained earlier, we extract 5 types main features from hand motion in the KSL.

(1) Sampled latest 3 points

For checking of direction change(up, or down, slant motion), this system has two registers($D_{x,y,z}(1), D_{x,y,z}(2)$) which contain sampled latest 3 points.

$$D_x(1) = dx(m),\ D_x(2) = dx(m-1) \quad (6)$$
$$D_y(1) = dy(m),\ D_y(2) = dy(m-1)$$
$$D_z(1) = dz(m),\ D_z(2) = dz(m-1)$$

, where m is sampled time step which all data (time step[1/15 sec.]) are preprocessed. An m is same to 1 time step[1/15 sec.] - no preprocessing case, or greater than 1 time step.

(2) Cumulative Length

From start of hand gesture, the cumulative length is increased and recognizable region is decided by this feature. When multiple hand motion case($\mathbf{D}_4$, $\mathbf{D}_7$, $\mathbf{D}_{13}$) is recognized, this value can be effective.

$$L1(m) = \sum_{m=1}^{M} ||l1(m)|| \quad (7)$$
$$L2(m) = \sum_{m=1}^{M} ||l2(m)||$$

(3) Difference of both position(Left & Right)

In Fig. 1, there are some gestures which left hand's motion is static or both hand's motion is symmetric. This feature can be shown those motions.

$$dbx(m) = rx(m) - lx(m) \quad (8)$$
$$dby(m) = ry(m) - ly(m)$$
$$dbz(m) = rz(m) - lz(m)$$
$$rx(m),\ ry(m),\ rz(m) \ :\ \text{right hand}$$
$$lx(m),\ ly(m),\ lz(m) \ :\ \text{left hand}$$

(4) Current Phase

We can know current phase in 2-D plane(x-y/x-z plane) from this feature.

$$CP1(m) = \tan 2^{-1}\left(\frac{y(m)}{x(m)}\right) \quad (9)$$
$$CP2(m) = \tan 2^{-1}\left(\frac{z(m)}{x(m)}\right)$$

(5) Cumulative Direction Change in radians

In circular hand motion($\mathbf{D}_1, \mathbf{D}_8, \mathbf{D}_{14}$), there are many direction changes and sum of direction change is very large.

$$\theta_1(m) = \tan 2^{-1}\frac{y(m)-y(m-1)}{x(m)-x(m-1)} - \tan 2^{-1}\frac{y(m-1)-y(m-2)}{x(m-1)-x(m-2)} \quad (10)$$
$$\theta_2(m) = \tan 2^{-1}\frac{z(m)-z(m-1)}{x(m)-x(m-1)} - \tan 2^{-1}\frac{z(m-1)-z(m-2)}{x(m-1)-x(m-2)}$$

Using these 5 features, we can classify 14 direction types on-line. All features are inputted to system every time step and current hand motion is checked by some rules. Fig. 4 shows features which used in the classification method about 14 direction types and features for each category. There is some delay in final recognition of hand motion for differentiating multiple hand motions ($\mathbf{D}_4$, $\mathbf{D}_7$, $\mathbf{D}_{13}$) from single motions ($\mathbf{D}_2$, $\mathbf{D}_5$, $\mathbf{D}_6$, $\mathbf{D}_{11}$, $\mathbf{D}_{12}$).

3.3 Classification of Hand Posture by FMMNN classifier

In our gestures(Fig. 1), 14 types of basic hand postures are included as shown in Fig. 3. It is proposed in this paper that each hand postures be recognized by applying the technique of Fuzzy Min-Max Neural Network [8].

The two flex angles of each finger(linearly scaled to lie between zero and one) are inputted to FMMN network shown in Fig. 5. We use the same function to describe a hyperbox as suggested in [8]. The membership function of the hyperbox is

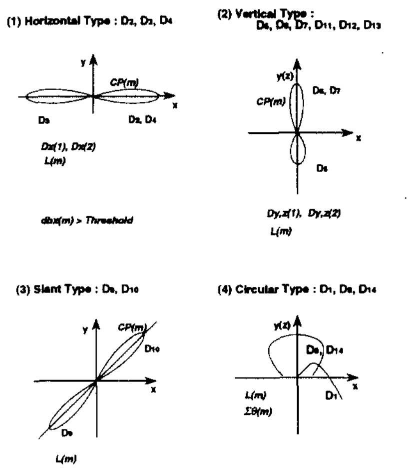


Figure 4: 4 categories in 14 directions classification algorithm

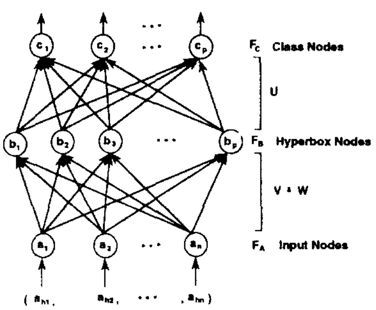


Figure 5: Structure of FMMN network

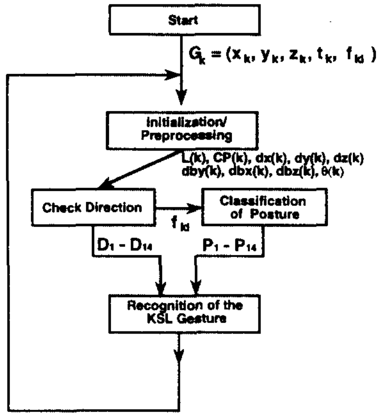


Figure 6: The Flow chart of the KSL recognition algorithm

expressed as follows:

$$b_j(F_h, V_j, W_j) = \frac{1}{20}\sum_{i=1}^{10}[\max(0,\ 1 - \max(0,\ \gamma\min(1,\ F_i - w_{ji}))) + \max(0,\ 1 - \max(0,\ \gamma\min(1,\ v_{ji} - F_i)))] \quad (11)$$

And $F_j, j = 1, 3, \cdots, 9$ is the flex angle of the inner joints of fingers and $F_j, j = 2, 4, \cdots, 10$ is the flex angle of the outer joints of fingers. If γ increases, the membership function becomes more crisp. In this system, γ of $F_j, j = 1, 3, \cdots, 9$ is smaller than γ of $F_j, j = 2, 4, \cdots, 10$ because $F_j, j = 2, 4, \cdots, 10$ is more sensitive to the change of flex angles. Given an input posture, the output of this network is the membership function values for 14 posture classes and the class with the maximum value of membership is classified as the designated posture if the value is above a given threshold(δ_2). If the maximum value is below δ_2, no decision is made.

3.4 Recognition of KSL

When a signer performs a motion for a gesture, many data are generated from the Data-Glove and inputted to the system. Feature analysis method transforms these raw data set into a set containing small number of data and this small number of feature data are used to recognize basic 14 direction classes. An input gesture is identified as one of the given direction classes, and then, the hand shape of that motion is recognized by the posture recognition method. With these two stages, the input pattern is recognized as one of the gesture

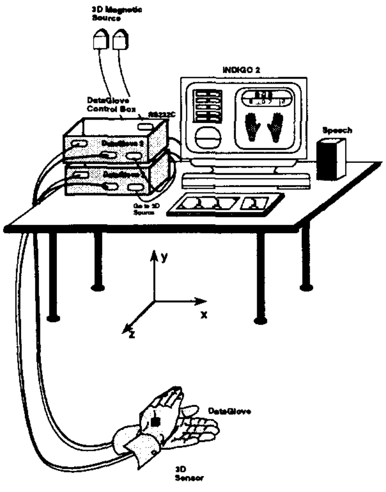


Figure 7: The Configuration of the KSL Recognition System

Figure 8: The Picture of the KSL Recognition System

classes on-line. Fig. 1 shows direction and posture type of each gesture.

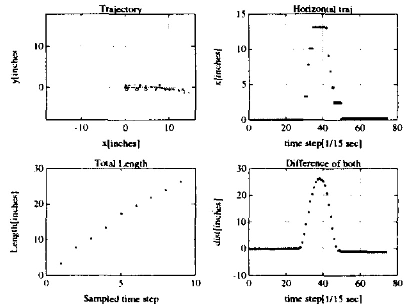
The flow chart of the recognition algorithm is given in Fig. 6.

4 Experimental Results

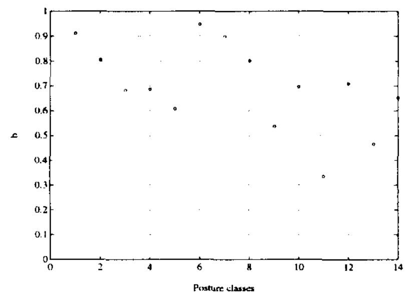
Fig. 7 shows the configuration of the KSL recognition system organized for our experiment and Fig. 8 shows real picture of our the KSL recognition system. All the functions are run on a INDIGO2 as main computer and the Data-Glove sends 32 data sets to the main computer via RS232C with 9,600 baud rates.

In this system, we use 10 flex angles and x, y, z-axis position data of both hand every time-step[1/15 second].

Fig. 9 shows a recognition result of gesture 'woman', for which the two hands are moving to make a right/left motion of hands with unfolded



(a) Direction $\mathbf{D}_4$



(b) Posture $\mathbf{P}_6$

Figure 9: Experimental Result: gesture 'woman'

little finger. In Fig. 9(a) is shown the motion trajectory along x and y-axis and some features sampled at every 1/15 second. '*' is sampled data and 'o' is sampled data after recognition of hand motion. In this case, the direction class is determined to $\mathbf{D}_4$ at the 45th time step by feature analysis method as described at section 3.2. After the direction class is decided, the membership values about 14 postures are obtained as shown in Fig. 9(b) by inputting flex angles to the FMMN network at 45th time step. In this Fig. 9(b), the number of class(horizontal axis) is assigned to the corresponding posture numbers in Fig. 3. In the network, $\theta = 0.2$, $\gamma_1 = 4.0$, $\gamma_2 = 8.0$, and $\delta_2 = 0.93$ are used. As shown in Fig. 9(b), the 6th posture has the maximum value of $b_j(11)$ which is greater than δ_2 and therefore, we conclude that the result of recognition is the gesture 'woman'.

Many experiments have been conducted with different sign languages and we have found that the success rate of our method in classification reaches up to almost 75 % of the given words. Taking into account the fact that deaf-mute who use gestures often misunderstand each other [2], we may say this success rate seems acceptable as a pilot study. We find that abnormal motions in the gestures and postures, and errors of sensors are partly responsible for the observed misclassification.

5 Conclusion

In this paper, a dynamic gesture recognition method is proposed, for which a new technique for efficient classification of motions is employed, and a fuzzy min-max neural network is applied for on-line posture recognition. The characteristics of human's hand gesture are investigated and on-line recognition system for Korean sign language gestures are implemented.

This system can be available for communication between doctors and patients in hospital, or between a police officer and a suspected person who impaired one's hearing in a court of justice.

References

[1] D. J. Sturman, *Whole-Hand Input.* PhD thesis, MIT, Feb. 1992.

[2] Korea Association for the Hearing Impaired, ed., *A handbook of Hearing Impaired(Written in Korean).* Special Education Press, 1992.

[3] S.-G. Kim, ed., *The Standard Korean Sign Language Dictionary (written in Korean).* Oh-Sung Press, 1993.

[4] VPL Research Inc., *Data-Glove Model 2+ : Operation Manual.* CA, 1992.

[5] S. S. Fels and G. E. Hinton, "Glove-talk : A neural network interface between a data-glove and a speech synthesizer," *IEEE Trans. Neural Networks*, vol. 4, pp. 2 – 8, Jan. 1993.

[6] M. B. Waldron and S. Kim, "Increasing manual sign recognition vocabulary through relabelling," in *Int. Joint Conf. on Neural Networks '94*, pp. 2885–2889, Jun. 1994.

[7] M. B. Waldron and S. Kim, "Isolated asl sign recognition system for deaf persons," *IEEE Trans. Rehabilitation Engineering*, vol. 3, pp. 261–271, Sept. 1995.

[8] P. Simpson, "Fuzzy min-max neural networks-part 1 : Classification," *IEEE Trans. Neural Networks*, vol. 3, pp. 776 – 786, Sep. 1992.

[9] The Korea Association for the Hearing Impaired, ed., *Korean Sign Language For Beginners(written in Korean).* Suh-Hyum Press, 1994.

[10] Dong-il Suk, *An Linguistic Analysis of the Korean Sign Language (written in Korean).* PhD thesis, Daegu Univ., Korea, Dec. 1989.

[11] Seoul Sun-Hee School, ed., *The Korean Sign Language I for middle school (written in Korean).* The Ministry of Education, 1992.