

# Recognition of Hand Written Hangeul Based on the Stroke Order of the Elementary Segment

\*Jeong-Young SONG, \*Kageo AKIZUKI, †Hee-Hyol LEE, †Won-Kyu CHOI

\*Akizuki Laboratory, Dept. of Electrical Engineering, Waseda university, 3-4-1 Ohkubo, Shinjuku-ku, Tokyo, 169 JAPAN

†Faculty of Engineering, Fukuoka Institute of Technology 3-30-1 Wajirohigashi, Higashi-ku, Fukuoka, 811-02 JAPAN

‡Center for Advanced Management Information System, 305-701 KAIST, Kusong-dong 373-1 Yusong-gu, Taejon, KOREA

## <ABSTRACT>

*This paper describes how to recognize hand written Hangeul character using the stroke order of the elementary segment.*

*The recognition system is constructed of 4 parts : character input part, segment disassembling part, character element extraction part and character recognition part.*

*The character input part reads the character and performs thinning algorithm. In the segment disassembling part, the input character is disassembled into elementary segments using the direction codes and the feature parameters. In the character element extraction part, we extract the character element using the stroke order and the knowledge rule. Finally, we able to recognize the hand written Hangeul characters by assembling the character elements, in the character recognition part.*

## 1. INTRODUCTION

Hangeul characters are composed of 14 consonants and 10 vowels as the basic character elements. Each of the character element is assembled from the following 7 kinds of the elementary segments.

- / ; LI(Left-Inclination)
- \ ; RI(Right-Inclination)
- — ; LR(Left-Right)
- | ; UD(UP-Down)
- ○, ⊙ ; C(Circle)
- ㄱ, ㅋ ; RD(Right-Down)
- ㄴ, ㄷ ; DR(Down-Right)

Furthermore, the segments are generated from combination of the basic strokes(/, \, |, —, ○).

Considering the foregoing special features of Hangeul, a hand written Hangeul character recognition system is proposed in this paper. The system takes into account of the stroke order of the elementary segments. The flow of the recognition procedure is shown in Fig. 1.

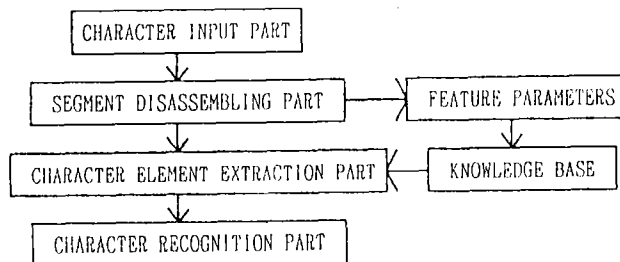


Figure 1: Recognition procedure

## 2. CHARACTER INPUT PART

A hand written Hangeul character is read by a camera and transformed to 82 by 82 white and black points. We perform the centering of the character and thinning process as a preprocessing. For example, the hand written character KIM'김' after the preprocessing is shown in Fig. 2. The numbers written in the character are the stroke order of the elementary segment.

## 3. SEGMENT DISASSEMBLING PART

In this part, Hangeul characters are disassembled into the 7 elementary segments. The el-

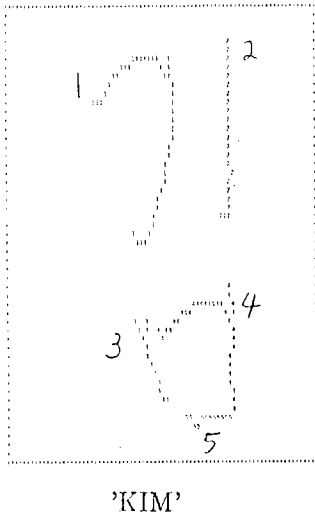


Figure 2: An example of input character (After the preprocessing)

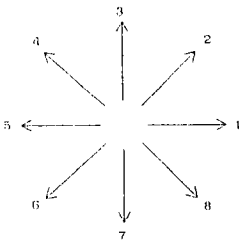


Figure 3: Direction codes

elementary segments are found by the following procedure. In the first place, we detect starting point of the stroke by searching from the upper-left corner. From the starting point and next point, we identify the direction codes shown in Fig. 3. The stroke is determined from the number of the series of the same direction code. When the point number is smaller than 10, it is rejected as a noise. After go through these steps, we extract the feature parameters; starting coordinates, ending coordinates, length of stroke and angle of stroke. And by following to the flow-chart in Fig. 4, we extract feature parameters. These parameters are stored into knowledge base. And then, the input character is disassembled to the elementary segments by the rule shown in Table 1.

As an example of the UD, IF the series of the same direction code greater than 10 and the

angle of the stroke is between 70 and 100 degrees THEN this stroke is UD.

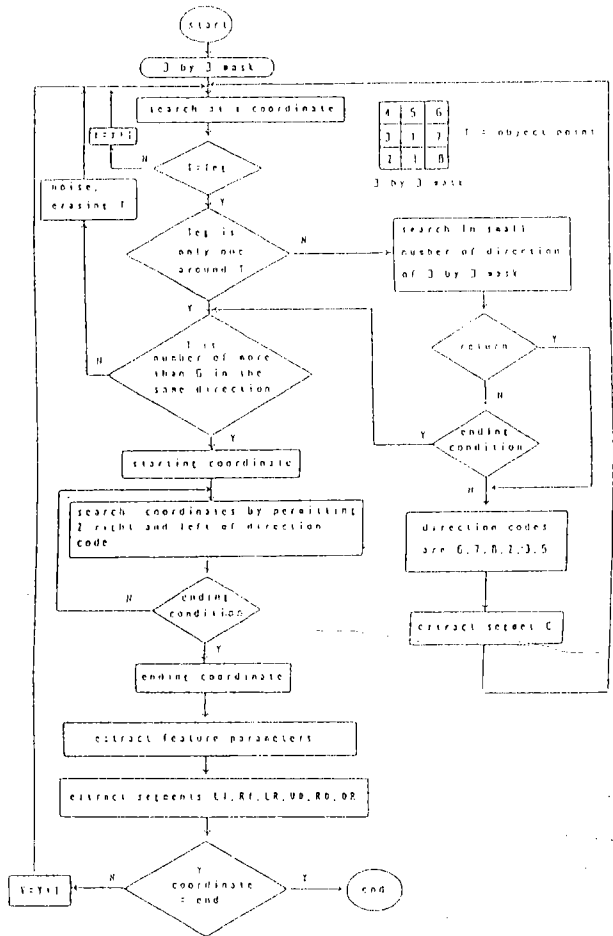


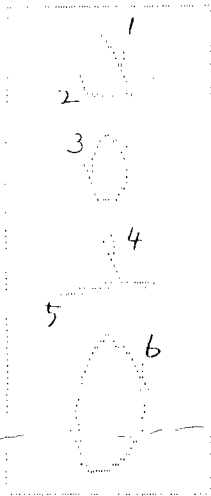
Figure 4: Flow-chart of extraction of the feature parameters

	The series of the same direction code greater than 10	The angle of the stroke
LI	6	40° ~ 70°
RI	8	110° ~ 140°
LR	1	170° ~ 210°
UD	7	70° ~ 100°
C	6, 7, 8, 2, 3, 5	250° ~ 360°
RD	1, 7	170° ~ 210°
		70° ~ 100°
	1, 6	170° ~ 210°
DR	7, 1	40° ~ 70°
		70° ~ 100°
	7, 2	170° ~ 210°
		70° ~ 100°
		210° ~ 240°

Table 1: The rules of elementary segments

The feature parameters of the input character

HONG'  $\frac{\text{홍}}{\text{ㅇ}}$ , and the extracted segments are shown Fig. 5 as example. Where # is the number of the stroke order, X0 and Y0 are the starting coordinates, X1 and Y1 are the ending coordinates, L is the length of the stroke and THE is the inclination of the stroke respectively. And we define the length of a circle as zero.



Information of kakijun 1  
# X0 Y0 X1 Y1 L THE  
1 5 36 16 49 17 139  
RIGHT-INCLINATION RI

Information of kakijun 2  
# X0 Y0 X1 Y1 L THE  
1 13 29 17 47 18 167  
RIGHT-INCLINATION RI

Information of kakijun 3  
# X0 Y0 X1 Y1 L THE  
1 23 39 23 39 0 360  
CIRCLE C

Information of kakijun 4  
# X0 Y0 X1 Y1 L THE  
1 39 41 48 44 9 108  
UP-DOWN UD

Information of kakijun 5  
# X0 Y0 X1 Y1 L THE  
1 49 40 49 61 21 180  
LEFT-RIGHT LR

Information of kakijun 6  
# X0 Y0 X1 Y1 L THE  
1 57 39 57 39 0 360  
CIRCLE C

'HONG'

Figure 5: An example of the feature parameters and elementary segments of HONG'  $\frac{\text{홍}}{\text{ㅇ}}$

#### 4. CHARACTER ELEMENT EXTRACTION PART

In this part, the character element is extracted by the procedure shown in Fig. 6.

In this procedure, the number of the segments starts from the 4 segments, because a character elements of Hangeul gets 4 segments as its maximum.

The order of the segments in writing Hangeul character is shown in Table 2.

On the other hand, some segment have same stroke order as the other, such as the vowels 'ㅏ' and 'ㅑ' have UD, LR stroke order. For these we apply the elements knowledge base.

Other Hangeul character vowels which have the knowledge base shown in Table 3. Where superscripts of feature parameters X and Y are the stroke order of the elementary segment. The subscripts are the starting coordinates.

Another example is the knowledge rules for consonant character elements 'ㅃ' and 'ㅆ' as

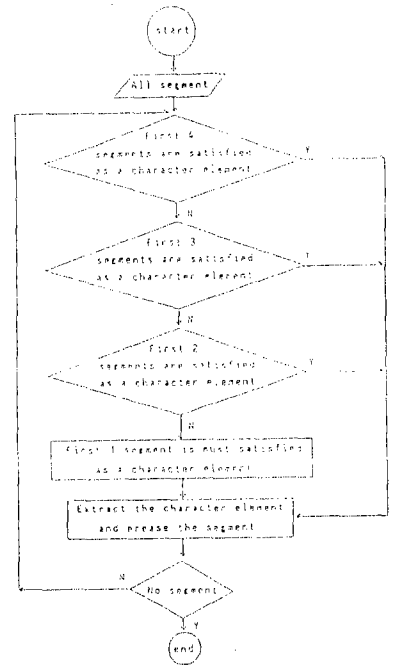


Figure 6: The procedure of character elements extraction

shown in Fig. 7.

	The first segment	The second segment	The third segment	The fourth segment
ㄱ	RD	*	*	*
ㄴ	DR	*	*	*
ㄷ	LR, LI	DR	*	*
ㄹ	RD	LR, LI	DR	*
ㅁ	UD, RI	RD	LR, LI	*
ㅂ	UD	UD	LR	LR
ㅅ	LI	RI	*	*
ㅇ	C	*	*	*
ㅈ	RD	RI	*	*
ㅊ	LR, UD, LI	RD	RI	*
ㅋ	RD	LR	*	*
ㅌ	LR, LI	DR, LR, LI	LR, DR	*
ㅍ	LR, LI	UD	UD	LR, LI
ㅎ	UD, LR, RI	LR, RI	C	*
ㅏ	UD	LR, LI	*	*
ㅑ	UD	LR, LI	LR, LI	*
ㅓ	LR, LI	UD	*	*
ㅕ	UD, RI	LR, LI	*	*
ㅗ	UD, RI	UD, RI	LR, LI	*
ㅛ	LR, LI	UD	*	*
ㅜ	LR, LI	UD	UD	*
ㅠ	LR, LI	*	*	*
ㅡ	UD	*	*	*
ㅝ	LR, LI	LR, LI	UD	UD
ㅞ	LR, LI	UD	UD	*
ㅟ	UD	*	*	*
ㅠ	LR, LI	LR, LI	UD	UD
ㅡ	LR, LI	UD	UD	*
ㅢ	UD	LR, LI	UD	*

Table 2: The order of the segments when writing Hangeul character

### 5. CHARACTER RECOGNITION PART

In this part, we recognize a character by assembling to the categories of character elements. As an example, the recognition process of hand written Hangeul HONG'  $\frac{\text{H}}{\text{O}}$ ' is shown in Fig. 8. This character is written in the order of consonant, vowel and consonant. The consonant '  $\frac{\text{H}}{\text{O}}$ ' is disassembled to the segments RI, RI and C and recognized as '  $\frac{\text{H}}{\text{O}}$ ' . The vowel '  $\text{O}$ ' is disassembled to the segments UD and LR and recognized as '  $\text{O}$ ' . The segment of the consonant '  $\text{O}$ ' is disassembled to C and recognized as '  $\text{O}$ ' . Combining these elements, we can recognize the character as '  $\frac{\text{H}}{\text{O}}$ ' .

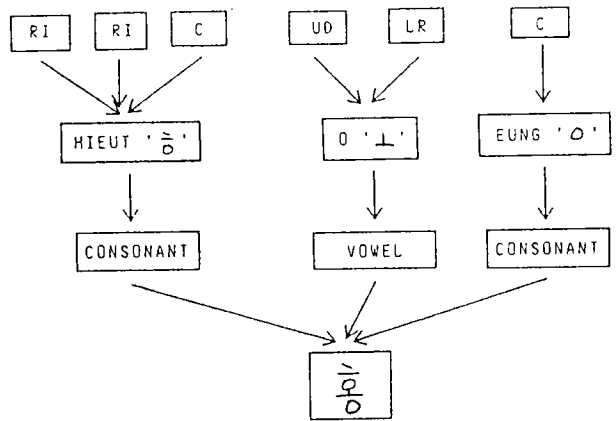


Figure 8: Recognizing process of hand written Hangeul '  $\frac{\text{H}}{\text{O}}$ '

	Number of stroke order	The rules of feature parameters
ㅏ	2	$X_0^1 < X_0^2$
ㅑ	3	$Y_0^2 < Y_0^1, Y_0^3 < Y_0^1$
ㅓ	2	$Y_0^2 < Y_0^1$
ㅕ	3	$X_0^1 < X_0^3, Y_0^1 < Y_0^2$
ㅗ	2	$X_0^2 < X_0^1$
ㅛ	3	$X_0^2 < X_0^1$
ㅜ	2	$Y_0^1 < Y_0^2$
ㅠ	3	$Y_0^1 < Y_0^2, Y_0^1 < Y_0^3$
ㅡ	1	*
ㅣ	1	*
ㅞ	4	$X_0^1 < X_0^3, X_0^2 < X_0^3$
ㅟ	3	$X_0^1 < X_0^4, X_0^2 < X_0^4$
ㅠ	3	$Y_0^2 < Y_0^1, Y_0^3 < Y_0^1$
ㅡ	3	$X_0^2 < X_0^1$
		$80^\circ < \theta^3 < 110^\circ$

Table 3: Knowledge base of the vowels

### 6. EXPERIMENTAL RESULTS

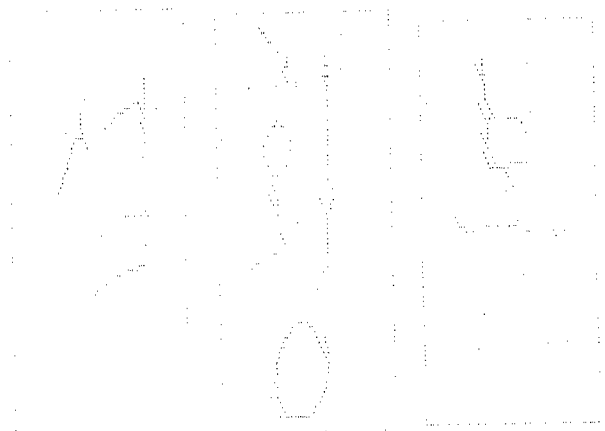
We applied our recognition system to hand written Hangeul characters which are written by 7 men and 6 women. These characters are the names randomly taken from "Korean Spiritual and Cultural Institute" which is the Korean standard name book. In this experiment, 95.3 percent of 280 characters are recognized correctly. The examples succeeded in recognition are shown in Fig. 9.

Some of the mistakes come from the segment disassembling part, because their character elements are connected each other.

The examples of failed in recognition are shown in Fig. 10.

<p>IF(( the first segment = UD .and.  the second segment = UD .and.  the third segment = LR .and.  the fourth segment = LR ) .AND.  ( <math>Y_0^1 &lt; Y_0^3 &lt; Y_0^4</math> .and. <math>X_0^2 &gt; X_0^1</math> .and.  <math>70^\circ &lt; \theta^1 &lt; 100^\circ</math> .and. <math>70^\circ &lt; \theta^2 &lt; 100^\circ</math> ) )  THEN 'ㅞ'</p> <p>IF((( the fifth segment = LR or RI ) .and.  the second segment = UD .and.  the third segment = UD .and.  ( the fourth segment = LR or RI ) ) .AND.  ( <math>X_0^1 &lt; X_0^2 &lt; X_0^3</math> .and. <math>X_0^3 &gt; X_0^2</math> .and.  <math>70^\circ &lt; \theta^2 &lt; 100^\circ</math> .and. <math>70^\circ &lt; \theta^3 &lt; 100^\circ</math> ) )  THEN 'ㅡ'</p>
--

Figure 7: Knowledge rules for character elements '  $\frac{\text{H}}{\text{O}}$ ' and '  $\text{H}$ '



'SEOL' 'HOANG' 'BONG'  
Figure 9: Patterns succeeded in recognition

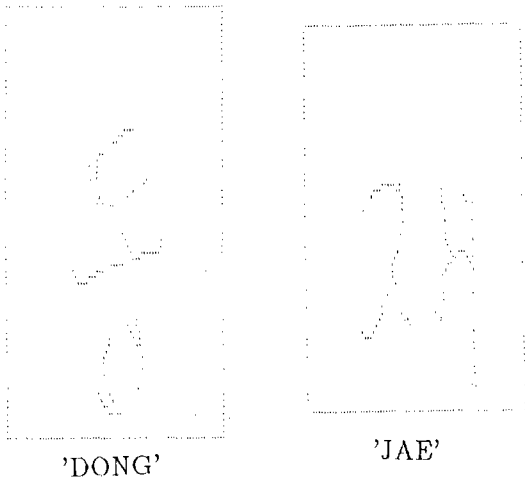


Figure 10: Patterns failed in recognition

## 7. CONCLUSION

In this paper, we propose a recognition system of hand written Hangeul which takes into account the stroke order of the elementary segments. The effect of this recognition method was confirmed by experimental studies. Now, we are improving the systems to recognize a character in running style.

## REFERENCES

- (1) J.Y.Song, H.H.Lee, W.K.Choi, K.Akizuki, A Recognition of Hand Written Hangeul by Structure Information, pp.715-716, SICE'92.
- (2) J.Y.Song, H.H.Lee, W.K.Choi, K.Akizuki, Recognition of Hangeul in a Simplified Form, pp.21-24, ISCIE 1992.
- (3) J.Y.Song, H.H.Lee, W.K.Choi, K.Akizuki, A Recognition of Hand Written Hangeul by Fuzzy Valuation of Segment, T.IEE Japan, Vol.113-C, No.4, '93.
- (4) J.Y.Song, H.H.Lee, W.K.Choi, K.Akizuki, Recognition of Hand Written Hangeul by the Back Propagation Algorithm with Moment Method, ISCIE, 1993.