A Hangeul Recognition Method Using Directional Edges in Open Captions

Seung Chul Jun, Myeong-Gyu Kang and Sung Han Park
Department of Computer Science & Engineering
HanYang University, Ansan, GyungGi-Do, Korea
Tel : +82-31-400-4109, Fax : +82-31-406-4673
{scjun, mgkang, shpark}@cse.hanyang.ac.kr

Abstracts: This paper proposes an efficient method to recognize Hangeul in video open captions. The open captions in news video can play an important role in the video indexing. The strokes of Korean character have a very strong horizontal and vertical directionality and some strokes appear repeatedly in each character. Based on this characteristics, in this paper, we propose an efficient algorithm to extract the character regions in open caption and recognize the characters based on these characteristics of Korean character. The simulation results demonstrate the efficiency of our algorithm in terms of computation time and recognition accuracy.

1. Introduction

As the amount of video contents increase rapidly, the automatic search of video data becomes an important issue[1]-[4]. Face, sound or captions are the most widely used information for the analysis of video data. Especially in news video, open captions frequently appear and so play an important role in the news video indexing. If texts in the open captions are extracted and recognized exactly, video contents can be efficiently stored and easily searched by the indexing key extracted in the process of the open caption recognition. Thus the exact recognition of the characters in the caption region is an important issue in the news video data storage and indexing.

For this purpose, we first need to extract the open caption regions within the video data. The conventional methods to extract the open caption regions are based on the characteristics of the open caption in which many vertical and horizontal line components are located[5]-[7].

The exact recognition of characters constitutes two steps. First we need eliminate the noise image residing in the open caption region[6]-[8] and then recognize the open captions. One method to remove the noise image is based on the characteristics that open caption exists in consecutive frames and has an intensity above a given value[8]. In this paper we consider only the recognition of Korean characters since the recognition of Hanguel is much different from that of the English. In general, there are two conventional methods of recognizing Hanguel[9]-[12]. One is based on the Korean phonemes (consonants and vowels) and the other on the syllables. The method based on phonemes has a merit of the quick recognition. However separating consonants and vowels from Korean character is a very difficulty job. The method based on the syllables has an advantage of recognition without separating syllables into the Korean phonemes but should compare with every Korean syllables.

To solve these problems this paper proposes an efficient method to recognize Korean characters without separating characters into consonants and vowels. Korean characters have characteristics of having several strong line components in its stroke[9]. We will use this property of Hanguel in our recognition algorithm. For this purpose, we define seven directional components in stroke.

In session 2, we deliver the methods to extract the text region in the video data and then the candidate open caption regions. The method to eliminate noise in these regions is explained. In session 3, the proposed Korean character recognition algorithm is provided. In session 4 the simulation results of the proposed algorithm are discussed and finally the conclusions will be followed in session 5.

2. Extraction of Caption Regions

2.1 Detecting Text Frames

In this paper we define a text frame as the frame in which text regions such as open caption exist. To extract the open caption from video frame, we first need to detect text frames from video sequences since a text region does not appear in all frames. The proposed method to find a text frame is shown in figure 1. We first assume the candidate text region in the frame of news video and search the text frame. In news video, text is generally assumed to appear on the right upper or the bottom area of the frame. The text region will show much more horizontal or vertical components compared with the news contents since Korean character has strong horizontal or vertical edge lines. To

This work was supported by grant No. 2000-2-303-005-3 from the Basic Research Program of the Korea Science & Engineering Foundation.

detect whether the frame is a text frame, we apply the horizontal differential filter (HDF) the candidate text region and then check the strength of the horizontal component. The horizontal strength is computed by counting the number of pixels of the HDF output. If frames with the strength greater than a predefined threshold value appear along several pixels corresponding to the character height vertically in consecutive, the frames are considered as text frames. In this paper we use 10 to 20 pixels as the height of Korean character in open caption. To make the text frame detection more reliable and simple, we perform a contrast stretching to the image data in the candidate text region and then convert the stretched gray image it to a binary image. In this paper we use 200 as the binary threshold intensity value since the intensity contrast of text is very high and greater than 200 in most of text regions. Figure 1 shows the proposed text frame detection algorithm.

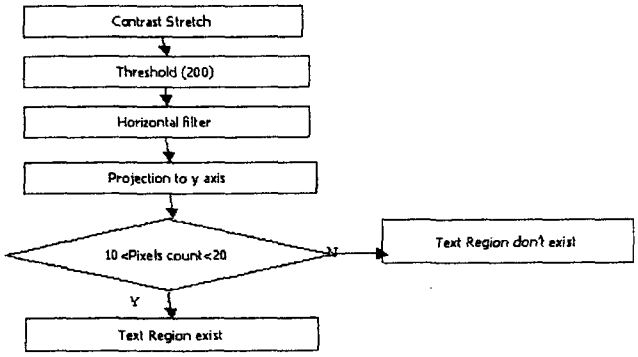

Fig 1. The Text Frame Detection Algorithm

2.2 Extracting Caption Regions from Text Frames

To recognize the Hanguel open caption, we first define the exact open caption regions. The caption region can be defined by the upper left and the bottom right positions. The upper and bottom positions are determined in the detection of the text frame. The starting and ending positions of the caption region can be determined by applying a vertical differential filter (VDF) to the candidate text region. We compute the number of pixel points of the VDF output. The starting position of the caption region is the first vertical line satisfying the condition such that the number is within the pixel value of Korean caption character, which is assumed to be 10 to 20 pixels. The ending position is the vertical line that the above condition is first violated.

An Open caption exists in several consecutive frames in same position without its content being altered. The time interval between frames having the same captions is called the text cut. To improve the recognition rate of captions with complex background at low resolution, we eliminate to some extent the noise in the caption region by averaging these consecutive text frames within the text cut in pixel domain. In addition, by converting the image in the caption image to a binary image and labeling the binary image, we improve the noise problem.

3. Recognizing Hangeul Captions Using Directional Edges

The Korean characters can be categorized into three groups shown in Figure 2 depending on the features of stroke of each character. As shown in the figure we can see that Korean characters have mainly horizontal and vertical edge components and consist of the breaking line-type edge segments. Based on this observation, we define seven stroke types shown in figure 3. The table 1 shows an example of how Korean character can be expressed with the seven strokes.

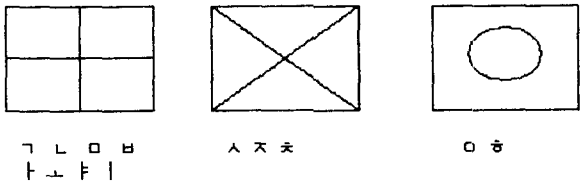

Fig 2. The Pattern Diagram of Korean Character

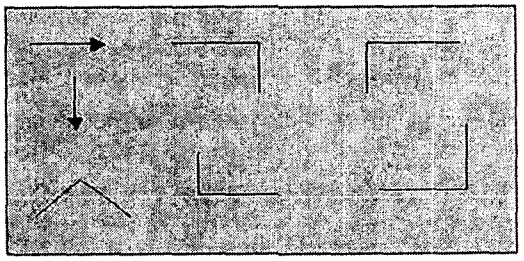

Fig 3. The Seven Directional Edges

Table 1. An Example of Representing Characters Using The Seven Strokes

	ㄱ	ㅂ	ㅏ	ㅑ
→	1	2	1	2
↓	1	2	1	1
∧	0	0	0	0
⌐	1	1	0	2
ㄴ	0	2	1	0
⌜	0	1	1	0
⌟	0	2	0	2

Given a character image, component of strokes constituting each character is determined by the following

three operations. To detect the vertical component among the seven strokes, we count the number of pixels along each vertical line and find the number of lines whose pixel number is larger than a threshold value. We call this number of lines as the vertical number. Similarly the horizontal number can be determined along the horizontal line. Figure 4 shows the number of pixels along horizontal and vertical lines of a Korean character "다" and the horizontal number is two and the vertical number three.

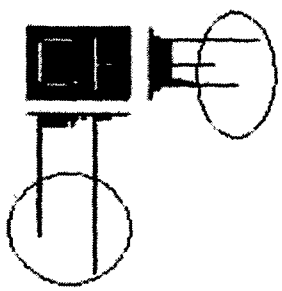

Fig 4. An Example of Vertical and Horizontal Number of Korean Character, "다"

The remaining five strokes can be determined using morphological filters. Figure 5 shows five morphological filters used in this paper to detect five strokes. Each morphological filter is used to detect one stroke. The results of this operation are compared with the predefined patterns shown in table 1.

In many Korean characters, consonants appear both in the "initial" and "last" of its orthographical syllable. And the "last", in most open caption, appears at the position below 50% of the character height. So for exact recognition of Korean characters, we need to separate the "last" from the other parts of Korean characters, since for example "강" and "악" may not be differentiated. The proposed method to isolate the "last" from a character is shown in figure 6.

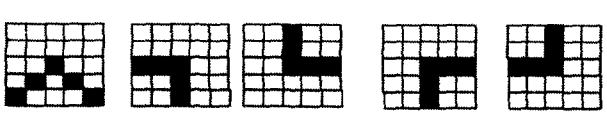

Fig 5. The Morphological Filters For Five Different Breaking Strokes

The overall procedure to recognize the Korean characters is given in figure 7. We first determine whether a character has a "last" or not. If there is a "last", it is separated from the character. Then the seven strokes of the "last" is detected and compared with the pattern in Table 1. The other part of the character is recognized in the same way. Figure 8 shows the results of an example of Hanguel recognition.

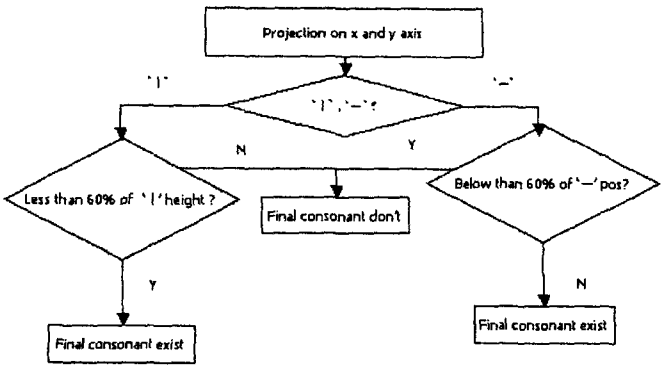

Fig 6. The Method Of Separating "last" Consonant From The Korean Character

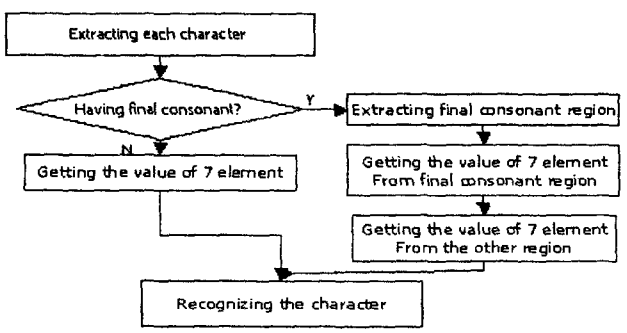

Fig 7. The Overall Procedures of Recognizing Hangeul Open Caption

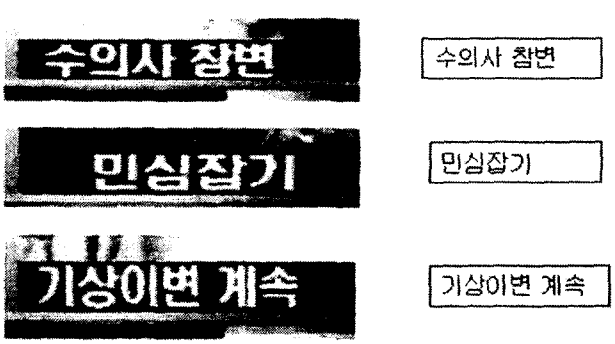

Fig 8. An Example of Hangeul Recognition

4. Simulation Results

We simulate our method using the main news video of three Korean broadcasting stations. The news video sequence is encoded in MPEG-1. The structure of the GOP is 15 frames and IBBP. Captions appeared in news data are mainly a " Kulim" and " Kodik" Korean fonts. The table 2 shows the simulation results. The correct recognition rate is approximately 70%, compared with the 60% accuracy of the conventional method. Our method also shows 4ms faster in processing time than the other methods. For the

scanned data of randomly typed 500 Korean characters, the proposed method shows the correct recognition rate of 100%.

Table 2. The Simulation Results

		Method	Right	False	Recog (%)	Time (msec)
No Noisy	Kulim font	This	297	3	99	350
		B	282	18	94	380
	Dodum font	This	294	6	98	410
		B	279	21	93	430
Noisy	MPEG1 News	This	102	30	77	21
		B	98	40	71	25

5. Conclusions

In this paper, we propose an efficient algorithm to recognize Hanguel open captions. We only separate the "last" consonant instead of separating consonants and vowels of Korean characters and define seven stroke components. In this way we only need to compare with very small range of predefined patterns instead of comparing with every Korean characters used in the conventional methods. The simulation results show the performance of our proposed method. This paper only considers the Korean characters, so we need to improve our methods to include the alpha-numeric characters.

References

[1] Thomas Sikora, "MPEG Digital Video-Coding Standards", *IEEE Signal Processing*, September 1997, pp. 82-100

[2] The Informedia Project information : Digital Video Library Proposal URL (http://www.informedia.cs.cmu.edu/)

[3] H.D. Wactlar, T.Kanade, M.A. Smith and S.M. Stevens, "Intelligent Access to Digital Video : Informedia Project", *IEEE Computer Magazine*, May 1996, pp. 46-52

[4] Toshio Sato, Takeo Kanade, E. K. Hughes, "Video OCR: Indexing Digital New Libraries by Recognition of Superimposed Caption", *Content-Based Access of Image and Video DataBase, 1998. Proceedings., IEEE Internation Workshop on*, pp. 52-60, 1998

[5] Y. Zhong, H. Zhang, A. K. Jain," Automatic Caption Localization in Compressed Video", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 385-392,2000

[6] Somyung Kim, Yeongwoo Choi, Kyusik, "Video Caption Extraction and Image Enhancement", *Proceedings., The 12th Workshop of Understanding and Processing Images*, Jan. 2000, pp. 342-346

[7] R. Lienhart, "Automatic Text Recognition for Video Indexing", *Proc. ACM Multimedia, Boston, MA*, Nov.1996, pp. 11-20

[8] H. Li, D. Doermann, "Text Enhancement in digital video using multiple frame intergration",*ACM Multimedia*, 1999

[9] Kil Nam Oh, Seung Ho Shin, Yong Ohk Chin,"A Recognition Algorithm of Hangeul Alpahbet Using 2-D Digital Filtering", *Journal of The IEEK*,1984, pp. 55-59

[10] Yong Ryul Kim,Sung Sub Kim, Young Shik Moon, "Study on Video Chrararcter extraction and Recognition", *The Proc. Of IEEK summer conference* ,vol. 24, pp. 141-144

[11] Young-Kyu Park, Sung-Guk Kim,WonYoung Yoo,Joon-Cheol Kim, Joon-Whoan Lee,"Extraction and Recognition of Character from MPEG-2 News Video Image", *The Journal of KISS*, vol. 6,1999, pp. 1410-1418

[12] Jee Ho Jung, Tae Young Choi, "Recognition of Printed hangeul Text Using Circular Pattern Vectors", *Journal of The IEEK*, May. 2001, pp. 33-45