International Journal of Advanced Culture Technology Vol.6 No.3 173-177 (2018)
DOI https://doi.org /10.17703//IJACT2018.6.3.173

IJACT 18-9-19

Image Comparison Using Directional Expansion Operation

Suk Won Yoo

SeoKyeong Univ., Dept. of Computer Science, Seoul, Korea
swyoo@skuniv.ac.kr

Abstract

Masks are generated by adding different fonts of learning data characters in pixel unit, and pixel values belonging to each of the masks are divided into 3 groups. Using the directional expansion operators, we expand the text area of the test data character into 4 diagonal directions in order to create the boundary areas to distinguish it from the background area. A mask with a minimum average discordance is selected as the final recognition result by calculating the degree of discordance between the expanded test data and the masks. Image comparison using directional expansion operations more accurately recognizes test data through 4 subdivided recognition processes. It is also possible to expand the ranges of 3 groups of pixel values of masks more evenly such that new fonts can easily be added to the given learning data.

Keywords: *Regional Boundary, Directional Expansion, Image Processing, Character Recognition, Boundary Expansion, Correlation, Accumulation, Image Analysis, Feature Extraction*

1. Introduction

Image processing technique uses an image created by an optical device as an input, and it might be regarded as a process of changing the input image in order to extract necessary information. The information obtained in the image processing process is used to analyze [1] the image in the image recognition process later. In order to recognize test data characters with a new font using given learning data characters, image processing [2] for test data characters as well as learning data characters is also important. In particular, the process of image comparison [3] between learning data characters and test data characters is an important part of many processes used for character recognition [4]. In order to obtain better recognition results, we use directional expansion operators to create a boundary area separating the background area and the text area by expanding the text area of the test data character into 4 diagonal directions.

2. The Related Works

As the industry develops, human uses a variety of information, and the amount of information processed by computer systems is increasing rapidly each year. Especially in the modern society, a lot of information is provided in the form of text and images, and the document automation system automatically processes them. What important thing in developing a document automation system [5] is the ability to recognize characters in documents or images [6]. As a process of recognizing characters, a method of comparing with learning data characters in case of using a specific font [7], and a method of extraction of character structures in case of

Manuscript received: July 17, 2018 / revised: July 31, 2018 / Accepted: August 20, 2018
Corresponding Author: swyoo@skuniv.ac.kr
Tel:+82-2-940-7181
SeoKyeong Univ., Dept. of Computer Science, Korea

handwritings have been studied. Nowadays, there are some other studies for character recognition using neural network theory [8]. However, this approach requires a lot of learning time and has difficulties in modifying the learning data or adding new data into the given learning data.

3. The Main Subject

Nowadays, due to the development of the 4th industry, recognition technology is used for unmanned automation in many fields. Especially, car plate numbers, mail items, and goods are automatically recognized and processed by using character recognition systems in parking lots, post offices, and markets. These recognition systems use specific fonts to increase the character recognition rate, but the recognition rate might vary greatly for various types of fonts. In order to resolve these problems, this paper introduces a technique to recognize characters with a new font using directional expansion operators.

3.1 The Basic Concept

What important thing in recognizing characters with various fonts is the process of dealing with the variety of fonts. There is a need for a rule that recognizes not only the unique characteristics of characters but also the various changes of fonts of characters. Thus, in order to recognize the changes of various fonts of the learning data, we use a method to newly create a boundary area between the background and the text area of test data character.

The 4 pixels constituting Direct Neighbor tends to be biased only in up-down or left-right direction although the connection with the center pixel is strong. On the other hand, the 4 pixels constituting Indirect Neighbor have weak connection with the center pixel, but they might create an adverse effect of blurring the image. In order not to use Indirect Neighbors and also not to bias too much in one direction, we use directional expansion operators to extend text area in 4 diagonal directions using two Direct Neighbors in this paper.

3.2 Image Comparison using Directional Expansion Operation

The image comparison using directional expansion operation will be described in detail by the following steps.

Step 1) Learning data consists of 10 different characters from 0 to 9, and each character has 10 different fonts. Test data is composed of 10 different characters from 0 to 9 and has a new font not used for learning data characters. Each learning data character and test data character is a black and white image with 0 and 1 of 15x15 size, where 0 represents background and 1 does text area.

Step 2) For each learning data character, create a mask by adding 10 different fonts in pixel unit. Because mask is created by adding 10 different fonts, it has a characteristic that emphasizes common parts among various fonts.

Step 3) Find a maximum value for each of 10 masks. Then, divide the values of the corresponding mask into 3 groups: from 0 to 1/4 of the maximum, from 1/4 of the maximum to 3/4 of the maximum, and from 3/4 of the maximum to the maximum. After that, values belonging in the first group are changed to 0, values in the second group are to 1, and values in the third group are to 2. Among these 3 values of the changed mask, 0 corresponds to the background, 1 does to the variation of various fonts, and 2 does to the text area. For reference, the reason for separating into 3 groups using 1/4 of the maximum value and 3/4 of the maximum value is the result obtained by experimenting with various kinds of values in the calculation process using these directional expansion operators.

Step 4) For each of test data characters,

Step 4-1) Change pixel value 1 to 2. In this way, pixels with a value of 2 correspond to the text area, and pixels

with a value of 0 do to the background.

Step 4-2) Move pixels with value of 2 to the right direction and to the down direction respectively, and overlap with itself. Then, the result of expanding the test data character in right and down directions is obtained. At this time, a new area added to the test data character while expanding process is represented by pixel value 1. Let's say that the generated result is the expansion #0 of the test data character. In other words, the expansion #0 is the result obtained by expanding the test data character in two directions, to right and to down. In a similar fashion, expanding the test data character in two directions, to left and to down, is called expansion #1, expanding it in right and up directions is expansion #2, and let the result obtained by expanding it in left and up directions be expansion #3. Then, each test data character has 4 expanded modifications that are different from each other.

Step 4-3) For each of the 4 expansion results,

Step 4-3-1) Degree of discordance for each of the 10 masks is calculated. This means that the expansion result and the mask are compared in a pixel unit, and then it is regarded that a mismatch between the expansion result and the mask has occurred at the corresponding pixel if any condition of the following two cases happens: 1) the pixel value of the expansion result is 0 and the pixel value of the corresponding pixel of the mask is 2, or 2) the pixel value of the expansion result is 2 and the pixel value of the corresponding pixel of the mask is 0. As a result, since the expansion result and the mask are compared in a pixel unit, the degree of discordance to be calculated is the sum of the number of discordant pixels between the expansion result and the corresponding mask.

Step 4-4) Each test data character currently has 4 discordances for each of the 10 masks. For each mask, find the average of the 4 discordances that the mask has. Then, each test data character has 10 average values of discordances.

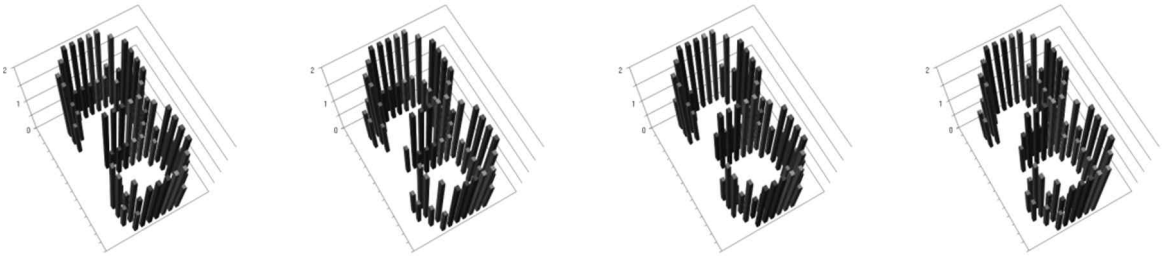
Step 4-5) Find a mask with the minimum discordance value among these 10 averages. The mask with the minimum discordance value is selected as the recognition result for the corresponding test data character.

3.3 The Results

The image comparison system has a Learning Data Set [9] composed of 100 number of characters and also a Test Data Set [9] composed of 10 characters, where test data has a new font that the learning data did not use. Each of learning data characters and test data characters is a black and white image of 15x15 size. The background has a pixel value of 0 in white, and the text area does a pixel value of 1 in black.

For each of the 10 learning data characters, add its 10 different fonts in pixel unit and find maximum value. Then, the result obtained by adding in pixel unit is divided into 3 groups: from 0 to 1/4 of the maximum, from 1/4 of the maximum to 3/4 of the maximum, and from 3/4 of the maximum to the maximum value. The pixel values belonging to these groups are modified to 0, 1, and 2 in order and then stored in the corresponding mask.

Figures 1(a) through 1(d) show the results obtained by running the directional expansion operators on the test data character 3 as an example. In Figure 1(a), the text area of character 3 is represented by pixel 2, and a new boundary area is created at the right and the down of the text area of character 3 by the directional expansion operator. The newly created boundary area is represented by pixel 1, and the text area of character 3 is wrapped with a thickness of 1 in size. This boundary area can be regarded as a result of expanding the text area of character 3 by moving it in right and down directions. Figure 1(b) shows the result of expanding the text area by moving it in left and down directions because the new boundary area is created at the left and down sides of the text area of the character 3 by the directional expansion operator. Figure 1(c) shows the result of expanding the text area by moving it in right and up directions because a new boundary area is created at the right and the up. Figure 1(d) is the result of expanding the text area by moving it in left and up directions.



(a) (b) (c) (d)

Figure 1. (a) Expansion #0 of Learning Data 3 in right and down directions
(b) Expansion #1 in left and down directions
(c) Expansion #2 in right and up directions
(d) Expansion #3 in left and up directions

With the condition that pixel values in each mask are divided into 3 groups using 1/4 and 3/4 of the maximum value, Table 1 shows discordance values with 10 masks using directional expansion operators for each of the 10 test data characters, the minimum discordance value, and the selected mask number as the finally recognized result for the corresponding test data character. As a result, directional expansion operators with 3 groups using 1/4 and 3/4 of the maximum values produce a 100% recognition result for the 10 test data characters.

Table 1. Discordance between MASK and Test Data with Range Values 1/4 and 3/4

	M#0	#1	#2	#3	#4	#5	#6	#7	#8	#9	MIN	M#
TD #0	3.00	33.25	24.75	21.50	34.75	16.50	14.25	33.75	15.75	14.75	3.00	0
TD #1	48.00	6.50	37.00	35.50	26.75	38.00	40.50	24.50	43.00	43.00	6.50	1
TD #2	29.00	23.25	10.00	20.25	39.50	28.00	34.25	19.75	21.50	20.25	10.00	2
TD #3	18.00	27.50	16.00	9.00	36.75	12.25	18.00	25.25	9.75	9.50	9.00	3
TD #4	31.50	38.50	36.50	33.00	5.25	37.00	36.25	38.50	31.25	36.25	5.25	4
TD #5	21.50	28.50	27.75	18.25	36.00	2.75	16.75	29.50	16.25	21.25	2.75	5
TD #6	15.25	33.50	27.50	17.75	36.00	7.25	5.75	32.00	6.50	17.00	5.75	6
TD #7	39.50	17.75	19.50	23.25	34.00	32.25	43.00	2.50	34.00	32.50	2.50	7
TD #8	13.50	33.50	23.00	14.50	34.75	8.50	8.50	30.00	3.00	15.75	3.00	8
TD #9	14.00	33.50	20.50	17.75	43.00	16.50	18.50	31.00	16.25	4.25	4.25	9

3.4 The Pros and cons of the proposed Algorithm

Similar to the conventional character recognition techniques, the proposed algorithm also has advantages and disadvantages. Advantages include: 1) To compute the boundary area surrounding the text area of test data characters, we use an expansion method subdivided into 4 different diagonal directions. Thus, test data characters can be recognized more accurately. 2) As the recognition rate increases, the width of the 3 areas of the masks can be expanded more evenly. So, more fonts might be added to the learning data. The disadvantage is that it takes more time to compute expansions in 4 different directions in order to calculate more accurately boundary area surrounding the text area of the test data character, as compared to other general area expansion methods.

4. Conclusion and Future work

Image comparison using directional expansion operation proposed in this paper recognizes not only the

characteristics of learning data but also the change of various fonts using 4 different directional expansion operators. By this process, it is possible to recognize characters with new fonts that have not been learned. In the future research, we are studying a method to solve the problem that test data is abnormally damaged due to the mechanical error of the optical devices.

References

[1] K. Jang, S. Lee, "Analysis of Digital Video with a Focus on the Documentary Colored Hearing", Journal of Convergence on Culture Technology(JCCT), Vol.2, No.2, pp.23-27, 2016.

[2] A. Koschan, M. Abidi, *Digital Color Image Processing*, Wiley, pp.37-67, 2008.

[3] S. Annadurai, R. Shanmugalakshmi, *Fundamentals of Digital Image Processing*, Pearson, pp.103-126, 2007.

[4] L. Jain, B. Lazzerini, *Knowledge-based Intelligent Techniques in Character Recognition*, CRC Press, pp.65-90, 1999.

[5] S. Ao, X. Huang, O. Castillo, *Intelligent Automation and Computer Engineering*, Springer, pp.79-92, 2010.

[6] M. Petrou, C. Petrou, *Image Processing*, Wiley, pp.47-73, 2010.

[7] W. Pratt, *Introduction to Digital Image Processing*, CRC Press, pp.139-154, 2013.

[8] A. Galushkin, *Neural Networks Theory*, Springer, pp.37-51, 2007.

[9] S. Yoo, "Character Recognition Algorithm using Accumulation Mask", International Journal of Advanced Culture Technology(IJACT), Vol.6, No.2, pp.123-128, 2018.