A Novel Dynamic Character Grouping Approach Based on the Consistency Constraints

Pengfei Xu[1(✉)], Qiguang Miao[3], Ruyi Liu[3], Feng Chen[2], Xiaojiang Chen[1], and Weike Nie[1]

[1] School of Information Science and Technology, Northwest University, Xi'an 710127, China
pfxu@nwu.edu.cn
[2] Office of Science and Technology, Northwest University, Xi'an 710127, China
[3] School of Computer, Xidian University, Xi'an 710071, Shaanxi, China

Abstract. In optical character recognition, text strings are extracted from images so that it can be edited, formatted, indexed, searched, or translated. Characters should be grouped into text strings before recognition, but the existing methods cannot group characters accurately. This paper proposes a new approach to group characters into text strings based on the consistency constraints. According to the features of the characters in the topographic maps, three kinds of consistency constraints are proposed, which are the color, size and direction consistency constraint respectively. In the proposed method, due to the introduction of the color consistency constraint, the characters with different colors can be grouped well; and this method can deal with the curved character strings more accurately by the improved direction consistency constraint. The final experimental results show that this method can group the characters more accurately, and lay a good foundation for text recognition.

Keywords: Grouping characters · Topographic maps · Color information · Character expandability · Consistency constraint

1 Introduction

Recognizing individual characters separately fails to take advantage of the whole word context, and the recognition results cannot represent the meaning of the word[1-2]. Character grouping is a difficult task, and much of the previous methods can only work on specific cases[3]. In order to solve these problems, researchers proposed character grouping methods to group characters into text strings, then these strings can be recognized to represent the meaning of the words more accurately.

In 1999, Goto proposed a method called Extended Linear Segment Linking, which was able to extract text strings in arbitrary orientations and curved lines[4]. This method works on touching characters effectively, and requires that the sizes of the characters are similar. A bottom-up approach was proposed by Pal[5], but it cannot work on the curved text strings. In 2008, Roy proposed a method based on the foreground and background information of the characters to extract individual text strings from multi-oriented and curved text document[6]. In 2009, another method was presented by him to

H. Zha et al. (Eds.): CCCV 2015, Part II, CCIS 547, pp. 170–179, 2015.
DOI: 10.1007/978-3-662-48570-5_17

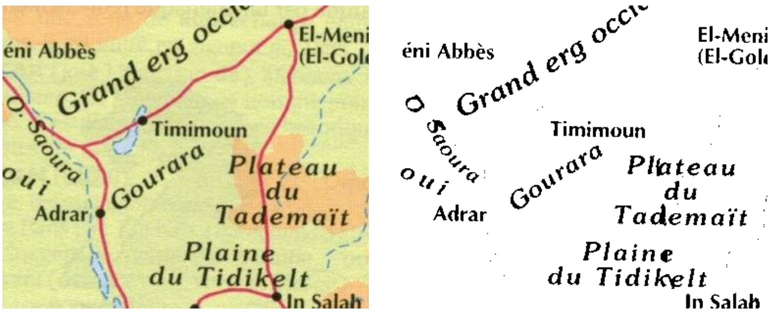
segment English multi-oriented touching strings into individual characters by using convex hull information[7]. These methods can deal with curved strings, but the directions of the strings were detected only in 4 directions. In 2004, a method for separating and recognizing the touching/overlapping characters was proposed by Velázquez[8]. In this method, OCR was applied to define the coordinate, size and orientation of the character strings, and four straight lines or curves were extrapolated to separate those symbols that were attached. In 2011, Aria Pezeshk grouped the individual characters into their respective strings using pyramid decomposition with Gaussian kernels[9,10], but this method cannot distinguish different text strings when they are nearby.

According to the analysis above, it is known that most researchers focused on the study of text separation and recognition, but the methods for grouping text strings are not researched deeply. Chiang has done lots of work on grouping characters[11] and text recognition[12], and a conditional dilation algorithm was presented for grouping characters into text strings[11]. Compared with other methods, Chiang's method can get better results. But there are still some problems, for example, the color information of characters is not considered, and the string curvature condition is not perfect. In order to solve these problems, this paper proposes a method to group characters into text strings based on the consistency constraints of the character color, size and directions.

The organization of the paper is as follows: In section 2, we make an analysis of the features of characters in maps. And the proposed method is described in section 3. In section 4, we compare the experimental results with Chiang's method. Finally, the concluding remarks are given in section 5.

2 The Analysis of Characters in Topographic Maps

In topographic maps, such as the map shown in Fig.1, the distribution of characters is very complex. The sizes of characters in different text strings are not the same, and the distance between some text strings is very small. In addition, there are broken and touching characters in the segmented maps. All these facts adversely affect the accuracy of the grouped text string.



(a) The topographic map (b) The text strings

Fig. 1. Text strings in topographic maps

At present, there are few methods for grouping characters into text strings, and these methods can only be used when the characters could be separated accurately. But some characters are mistakenly grouped into other text strings or leaved out when the characters in the map image have different sizes, directions, and colors, especially when some text strings are in a curved line.

3 Dynamic Character Grouping Based on the Consistency Constraints

According to the analysis in section 2, it is difficult to group the multi-oriented, multi-sized, and curved characters into text strings. This section gives a new method for grouping characters into text strings. The color information, which is not considered by the existing method[11], is used as an additional constraint due to that some characters are presented by different colors. The directional constraint is designed more perfect in the new method. So based on the consistency constraints of the character color, size and direction, the new character grouping method can be described by the following expression.

$$\mathrm{T} = \mathrm{G}\left(c,\ s,\ d\right) \tag{1}$$

Where, T is the grouped text string, $\mathrm{G}\left(\bullet\right)$ is the character grouping operation function. c, s and d are the color, size and direction consistency constraints, which means that the characters in one text strings have the similar color and size, and the centers of these characters are on a curved line，whose curvature is in a numerical range.

3.1 The Color Consistency Constraint

In topographic maps, some text strings are represented by different colors, so we can use this information to distinguish different text strings. And for a character α, its color feature is defined as:

$$C_\alpha = M\left(\alpha_{\mathrm{R}},\ \alpha_{\mathrm{G}},\ \alpha_{\mathrm{B}}\right) \tag{2}$$

Where C_α is the main color feature of the character α, which is obtained by color histogram. $M\left(\bullet\right)$ is the operation of color extraction, and the average value of RGB in the character area is used as the main color feature of this character. The color difference between the characters α_1 and α_2 can be obtained by Mahalanobis distance.

$$\mathrm{D}_{\alpha_1,\alpha_2} = \sqrt{\left(\mathrm{C}_{\alpha_1} - \mathrm{C}_{\alpha_2}\right)^{\mathrm{T}} S^{-1}(\mathrm{C}_{\alpha_1} - \mathrm{C}_{\alpha_2})} \tag{3}$$

Where S^{-1} is the inverse of the covariance matrix of the samples. The color consistency constraint of the characters means that $\mathrm{D}_{\alpha_1,\alpha_2}$ should be less than or equal to a minimum threshold T_c.

$$\mathrm{D}_{\alpha_1,\alpha_2} \le T_c \tag{4}$$

Where T_c is the average color difference in the area aound the current characters α_1 and α_2.

$$T_c = \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} D_{p_i,p_j}}{(\frac{N(N-1)}{2})} \tag{5}$$

Where N is the number of the pixels in the area, D_{p_i,p_j} is the color distance between the pixels p_i and p_j

If there are no color information, all the characters are viewed as that all of them have the same color. Further, they are grouped by the size consistency constraint and the direction consistency constraint.

3.2 The Size Consistency Constraint

For a character, its size is the max value of the length and the width of its bounding box, and the sizes of the characters in one text strings must be similar. So the size ratio of the characters must be smaller than a threshold T_s. According to the size features of the characters, such as the English letter '*f*' and 'a', we always choose $T_s = 3$ [11].

3.3 The Direction Consistency Constraint

For the text strings in the binary image, each connected component is a single character, and each character needs to be connected to at least one character. Assuming that there is a character α_1, we can get another character α_2 which is the closest character to α_1 according to the size consistency constraint and the color consistency constraint. Furthermore, based on the distances between different connected components, we try to get the neighbor character $\alpha_{1,\mathrm{N}}$ of α_1, and $\alpha_{2,\mathrm{N}}$ of α_2. In this way, at least two and no more than four characters, which may belong to the same text string, can be obtained.

There are three cases when we check α_1 and α_2 belong to the same text string or not.

(1). If there are no neighbor characters of α_1 and α_2, these two characters are grouped into a text string directly.

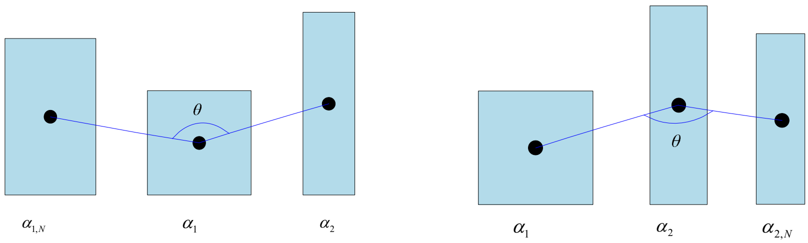
(2). If there is one neighbor character of α_1 or α_2, we need to check the direction consistency constraint of these three characters. Although the sizes of the characters of one text string meet the size consistency constraint, they are not the same, so the centers of these characters are on a curved line rather than on a straight line. The curved text grouping is similar with this case.

As shown in Fig.2, if the angle θ of the two lines is less than or equal to the threshold T_d , these three characters are satisfied with the direction consistency constraint.

$$|\theta - 180°| \le T_d \tag{6}$$

$$T_d = 180° \pm \beta \tag{7}$$

Here a curvature parameter β is used to control T_d, according to the distances and the size difference between the characters, or on the basis of the empirical data, this curvature parameter β is set to 50° generally.



(a) Neighbor character of α_1 (b) Neighbor character of α_2

Fig. 2. One neighbor character of α_1 or α_2

(3) If α_1 and α_2 have a neighbor character respectively, we need to check the direction consistency constraint of these four characters, as shown in Fig.3. If either the angle θ_1 or θ_2 is less than or equal to a minimum threshold T_d, these four characters are satisfied with the direction consistency constraint.

$$(|\theta_1 - 180°| \le T_d) | (|\theta_2 - 180°| \le T_d) \tag{8}$$

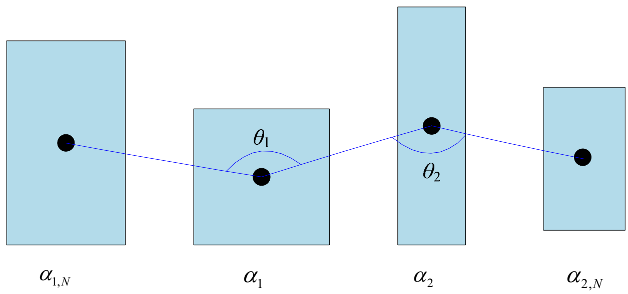


Fig. 3. Neighbor characters of α_1 and α_2

If the characters α_1, α_2 and neighbor characters are satisfied with the consistency constraint of the character color, size and direction, α_1 and α_2 are grouped into the same text string.

4 Experiments and Analysis

In this section, several experiments are made on artificial images and topographic maps to verify the accuracy of the proposed method. Because Chiang has made a comparison between the method proposed by him and several other methods in ref[3], and he came to the conclusion that his method has the best performances. therefore, in this paper, only Chiang's method[11] is chosen as a comparison method.

4.1 Experiments on Artificial Images

We made two artificial images which contain multi-color, multi-oriented, multi-sized, and curved strings to test our new method. In these images, the characters are seperated and unbroken. All the results are shown in Fig.4, the red pixels are the spreaded background pixles which connect to only one character, and each green pixel connects to two characters which would be grouped. By comparing the results, the proposed method can get more accurate results.

As shown in Fig.4 (c), Chiang's method can deal with the multi-oriented, multi-sized, and curved text strings well. But when there are two text strings in different directions, and the beginning or the end character of one string is close to one character of the other string. In this case, Chiang's method cannot deal with the beginning or the end characters of these two text strings, due to the disadvantages of Chiang's method in string curvature condition. In the proposed method, the direction consistency constraint is designed more perfect, so the results are better, as shown in Fig.4(b) .

(a) The artificial image

(b) the result acquired by the proposed method

(c) the result acquired by the Chiang's method

Fig. 4. the results obtained from artificial images

In other case, Chiang's method cannot handle the close text strings with different colors, due to that color information is not used in his method, as shown in Fig.5(c). In contrast, the color consistency constraint is used in the proposed method, so the characters with different colors are not grouped into the same string, as shown in Fig.5(b).

(a) The artificial image B

(b) the result obtained by the proposed method

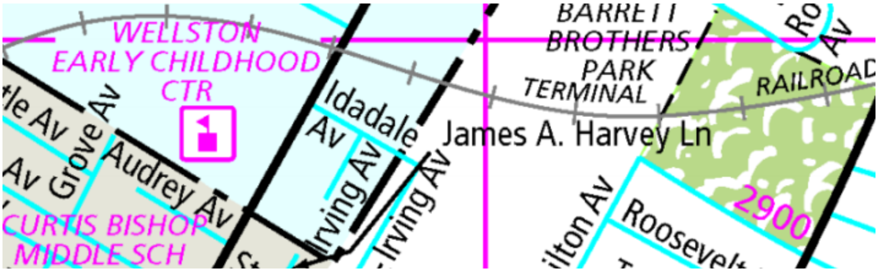
(c) the result obtained by the Chiang's method

Fig. 5. the results obtained from artificial images

4.2 Experiments on Topographic Map Images

In order to verify the accuracy of the proposed method in applications, two topographic maps are chosen as test images. Characters are separated based on color information and morphological features[2]. The grouped characters are shown in Fig.6.

From the results, the proposed method has advantages over Chiang's method. The former can deal with color and direction information better. From Fig.6(a_2) and （a_3), we can see that the proposed method can distinguish text strings with different colors, but Chiang's method cannot. Many characters, which is the beginning or end characters of the text strings, are not grouped into the corresponding text strings in the Chiang's results.



(a$_1$) The topographic map M_2

(a$_2$) the result obtained by the proposed method and the detailed image

(a$_3$) the result obtained by the Chiang's method and the detailed image

(b$_1$) The topographic map M_3

Fig. 6. The results obtained from topographic maps

(b$_2$) the result obtained by the proposed method and the detailed image

(b$_3$) the result obtained by the Chiang's method and the detailed image

Fig.6. (*Continued*)

5 Conclusion

It is well known that it is difficult to design a perfect method to group all the characters accurately in topographic maps. This paper proposes an algorithm to group characters into text strings based on the designed consistency constraint of the character color, size and direction. In this method, color features are introduced, and the direction consistency constraint is designed more perfect. Experimental results show that the proposed method can get better results. However, there are still many works to do. The proposed method cannot deal with the characters in the strings with big character spacing, so new methods should be studied.

Acknowledgments. The work was jointly supported by the National Natural Science Foundations of China under grant No. 61472302, 61272280, U1404620, 41271447, 61373177, and 61272195; The Program for New Century Excellent Talents in University under grant No. NCET-12-0919; The Fundamental Research Funds for the Central Universities under grant No. K5051203020, K50513100006, K5051303018, JB150313, and BDY081422; Natural Science Foundation of Shaanxi Province, under grant No.2014JM8310 ; The Creative Project of the Science and Technology State of xi'an under grant No. CXY1441(1); The State Key Laboratory of Geo-information Engineering under grant No.SKLGIE2014-M-4-4. The Science Foundations of Northwest University under grant No.14NW25, 14NW27, 14NW28.

References

1. Adam, S., Ogier, J.M., Cariou, C., et al.: Symbol and character recognition: application to engineering drawings. International Journal on Document Analysis and Recognition **3**(2), 89–101 (2000)
2. Pezeshk, A., Tutwiler, R.L.: Extended character defect model for recognition of text from maps. In: 2010 IEEE Southwest Symposium on Image Analysis & Interpretation (SSIAI), pp. 85–88. IEEE (2010)
3. Chiang, Y.Y., Adviser-Knoblock, C.A.: Harvesting geographic features from heterogeneous raster maps. University of Southern California (2010)
4. Goto, H., Aso, H.: Extracting curved text lines using local linearity of the text line. International Journal on Document Analysis and Recognition **2**(2–3), 111–119 (1999)
5. Pal, U., Sinha, S., Chaudhuri, B.B.: Multi-oriented English text line identification. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 1146–1153. Springer, Heidelberg (2003)
6. Roy, P.P., Pal, U., Lladós, J., et al.: Multi-oriented English text line extraction using background and foreground information. In: The Eighth IAPR International Workshop on Document Analysis Systems. DAS 2008, pp. 315–322. IEEE (2008)
7. Roy, P.P., Pal, U., Llados, J., et al.: Multi-oriented and multi-sized touching character segmentation using dynamic programming. In: 10th International Conference on Document Analysis and Recognition. ICDAR 2009, pp. 11–15. IEEE (2009)
8. Velázquez, A., Levachkine, S.: Text/graphics separation and recognition in raster-scanned color cartographic maps. In: Lladós, J., Kwon, Y.-B. (eds.) GREC 2003. LNCS, vol. 3088, pp. 63–74. Springer, Heidelberg (2004)
9. Pezeshk, A., Tutwiler, R.L.: Automatic feature extraction and text recognition from scanned topographic maps. IEEE Transactions on Geoscience and Remote Sensing **49**(12), 5047–5063 (2011)
10. Burt, P., Adelson, E.: The Laplacian pyramid as a compact image code. IEEE Transactions on Communications **31**(4), 532–540 (1983)
11. Chiang, Y.Y., Knoblock, C.A.: Recognition of multi-oriented, multi-sized, and curved text. In: 2011 International Conference on Document Analysis and Recognition (ICDAR), pp. 1399–1403. IEEE (2011)
12. Chiang, Y.Y., Knoblock, C.A.: An approach for recognizing text labels in raster maps. In: 2010 20th International Conference on Pattern Recognition (ICPR), pp. 3199–3202. IEEE (2010)