Feature Point Based Stroke Section Extraction for Chinese Characters

Lin, Zhen Yong
School of Electrical & Electronic Engineering
Nanyang Technological University
Nanyang Avenue
Singapore 639798

Abstract-Structural method is a very powerful approach in Chinese character recognition in which strokes and their relations are used as the features to describe a Chinese character. Thus stroke extraction performance will affect the recognition rate seriously. However, stroke extraction in Chinese characters is a difficult task. In this paper, a feature point based stroke section extraction method is presented. It can extract stroke sections between various feature points with high accuracy. This method is robust to stroke shape changes, various noises or small holes on or between strokes, artistic effects and inherent structural noise. It has comparable performance to a thinning method in correct stroke section extraction accuracy but avoids most undesirable side effects of thinning methods with much lower stroke section extraction error rate.

Key words: Stroke extraction, Thinning, Primitive stroke, Chinese character recognition

1 Introduction

It is well known that Chinese characters contain rich structural information, and the basic elements of Chinese characters are strokes. The types and numbers of strokes and the relationships among the strokes are essential structural features of a Chinese character. The ability to extract structural information depends on accurate extraction of strokes. However, extracting the strokes of a Chinese character correctly is a major problem because the shapes of Chinese characters are very complex, there are many kinds of strokes, the thickness of a stroke may vary dramatically, the noise may appear near by the contour of strokes, strokes may be crossed or connected in various manner, and strokes may be distorted. The research on the stroke extraction can be divided into two categories: stroke extraction with thinning preprocessing[1, 2, 9, 10, 13] and stroke extraction without thinning preprocessing [2, 5, 6, 7, 8, 11, 12]. In the first category, some thinning algorithms are employed to obtain the skeleton of a character. The strokes of the character are then deduced from the thinned character. However, because most thinning algorithms will produce undesirable side effects or distortion on strokes, the first approach makes the task of stroke extraction difficult. Therefore, some non-thinning approaches to stroke extraction were investigated, such as line template method, polygonal approximation method, contour following method, propagation method, shrinking method, pen movement stroke extraction method, and knowledge-based stroke extraction method. For detailed description or comparison of these methods, please refer to [7, 8] . In this paper, a new stroke extraction approach without thinning is proposed. It is based on detection of all feature points, such as break points, end points, bend points, and joints from the original character pattern, then stroke sections between these feature points are tested and extracted. The information obtained for stroke sections can be used to form complete strokes, the same use as thinning result, or directly used for Chinese character recognition as in [4]. The detailed description of the algorithm is given in subsequent sections.

2 Feature Point Detection

The type of feature points to be considered are break point where stroke width has abrupt change, end point which is the termination point of a stroke section, bend point where a stroke changes its direction dramatically, and joint point where between three to eight stroke sections intersect each other. The algorithm is based on the following observations: for the end point of a stroke, the width of the stroke will become zero beyond one side of this end point; for bend point of a stroke, its stroke width will have abrupt change at the both sides of the bend point, or its stroke gradient will change sign at the both sides of the bend point; for joints, all strokes near the joint will have abrupt change of their stroke width. Figure 1 depicts this observation.

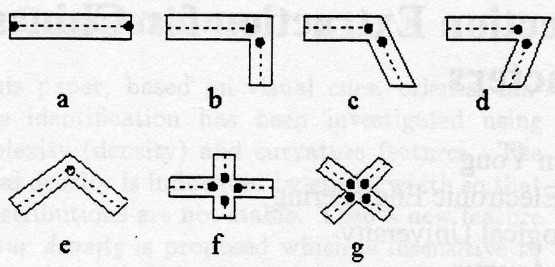


Figure 1. Strokes with abrupt width changes

In Figure 1, strokes are represented by the solid boundary lines; the doted lines indicate the middle points of the stroke. The collection of the middle points of a stroke is called the spine of the stroke. In Figure 1a, along the spine, stroke width at the right of the right black dot and at the left of the left black dot become zero, these black dots are called end points. An end point is defined as a middle point of a stroke which has zero stroke width at one side. Figure 1b-d show three bend stroke types where two stroke sections intersect at the same position with right, obtuse and acute angle. Traveling along the spine, stroke width before and after these black dots have abrupt changes. These black dots are called brake points. A break point is defined as a middle point which has stroke width change at its two sides more than a threshold value T. A bend point exists between the two break points. The bend point is defined as the intersect point of the two spines between the two break points. In Figure 1e, one diagonal and one anti-diagonal stroke sections intersect at one position to form a bend point. Although there is no abrupt stroke width changes along the spine, the stroke sections at the two sides of the circle have different gradient signs. This circle is a bend point. In this case, the bend point is defined as a middle point which has different gradient sign at its two sides. In Figure 1f-g, four stroke sections intersect at one position to form a joint. The joint consists of four break points. A four joint exists in the center of the break points. A joint is defined as the center of all surrounding break points. Obviously, a 3 joint has 3 surrounding break points, a 5 joint has 5 surrounding break points, and so on. All 3-8 joints are defined in this way. All these break points, end points, bend points and joints are called feature points. The detailed description of the algorithm to detect those feature points is given in the following steps.

(1) Finding average width of character stroke. To extract feature points in Chinese characters, all the Chinese characters are normalized into an 80 x 80 pixel matrix. Then the average stroke width is determined. This is done by scanning the matrix horizontally and vertically and checking each white-to-black and black-to-white transition pairs along the scan line and counting the number of black pixels in between each transition pair as stroke width. After obtaining all stroke width information from the scanned lines, a stroke width histogram is calculated. The minimum stroke width of the character is deduced from the histogram which is the smallest among the stroke width with frequency larger than or equal to 15, and the maximum stroke width of the character is one which is the largest among stroke width with frequency larger than or equal to 15. The average stroke width is obtained by taking the average of the minimum and maximum stroke width. This information is necessary to detect feature points of Chinese characters with different stroke shapes, for example, normal printing or bold printing, even handwritten characters.

(2) Determining the character spine and break points. Upon obtaining the stroke width, the character matrix is scanned horizontally, and then vertically again. For each pair of white-to-black and black-to-white transition along the scan line, its middle point will be marked if the length (stroke width) between the two transition points is lass than the average stroke width, or if its width is larger than average stroke width, but its width change compared to the previous scan line is less than a threshold value Tw. These middle points will form the spine of strokes in the character . All middle points which have stroke width change between the previous and next scan lines larger than threshold T, or the width becomes zero at previous or next scan line are marked as possible break points.

Once all middle points are found and marked, both end points of each continuous middle point section are marked according to the definition. They are checked at next stage to see if they are true end points of a stroke, or just break points of a bend point or a joint.

(3) Detection of end points. In this step, the horizontal middle points and vertical middle points and their surrounding points are used to determine whether each end point of a continuous middle point section is a true end point of a stroke. The criteria for an end point of a continuous middle point section to be classified as a stroke end point is as follows:

A. For horizontal middle points, if the end point is pointing to the right direction from its neighboring middle point, and there are white pixels spanning the width of its stroke width on its next right column, or if the end point is pointing to the left direction from its neighboring middle point, and there are white pixels spanning the width of its stroke width on its next left column.

B. For vertical middle points, if the end point is pointing to the up direction from its neighboring middle point, and there are white pixels spanning the width of its stroke width on its next up row, or if the end point is pointing to down direction from its neighboring middle point, and there are white pixels spanning the width of its stroke width on its next down row.

If the stroke end point test fails, then the end point of the continuous middle point section will be marked as a break point.

(4) Continuity tests to remove false break points. After steps (1-3), all true end points in the character will be detected. However, the remaining break points may contain some false break points. This step will eliminate these false break points. The continuity criterion is used for this purpose: for any break point, if there are adjacent middle points at both sides of it, it is a false break point.

In removing the false break points, some Y type 3 joint is a special case that needs special handling. Figure 2 shows the amplified stroke image grid near a stroke cross where the break point a for the right spine is taken to be continuous with respect to break point b and c on the left, thus being eliminated from break point instead of being identified as a break point of a 3 joint.

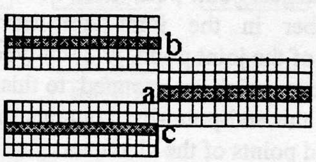


Figure 2 A special Y type 3 joint

To handle this special case, whenever a break point is tested with two other adjacent break points in the opposite direction, this break point is eliminated, the new break point is shifted back 1 pixel along the spine, this new break point will prevent the three break points from being rendered as continuous.

(5) Eliminating noisy strokes. The spine detected from the character often contains some very short noisy spines because of noise or artistic components on the strokes. These short noisy spines can be eliminated due to their short continuity feature. The criterion to eliminate a short noisy spine is that if the length from its end point to its break point is less than 5 pixels. In this case, the whole short spine, including its end point and break point, is eliminated.

(6) Detecting bending points and joints. After all end points and break points are identified in the character, the next task is to identify all bending points and joints in the character from the break points. The feature point detection algorithm can detect only bend points of bend types in Figure 1b-d and their $\pm 90°$ and $180°$ rotations. For the bend type in Figure 1e and its $\pm 90°$ and $180°$ rotations, their detection is left in next stage. Obviously, all break points that belong to a joint must be close each other and linked by black pixels. The test starts by taking any break point in a break point list as a reference point to find out other related break points linked by black pixels. Based on the direction the reference break point is facing from its neighboring middle point, a 3x3 mask centered at the next point along the direction is created. All black points in the mask are stored in a scan list and checked to find whether there is a break point. For each black point in the scan list, the same 3x3 mask is created for it by the same policy, and the black points in the mask but not in the scan list are added into the list and tested. The mask goes as far as the distance from its center to the reference break point is less than or equal to the maximum stroke width, and its center is more than the limit along the reference break point direction that will be described later. In this way, the test spreads from the reference break point to all black points connecting to other break points that belong to the joint, including the break points. If a break point is detected in the scan list, a group break point test is conducted to determine whether this break point really belongs to this joint, thus can be grouped with the reference point. Taking a reference break point facing UP as example, The criteria to qualify break points for the joint is: as long as they are not facing UP and below the reference point, they are qualified for break points of the joint. The limit along the reference break point direction, i.e., the bottom limit of the scan area in this case, is determined as follows: If a break point not satisfying the above qualify condition is found, the bottom limit of the scan area for the reference point is drawn at this line, i.e., the center of all masks must be above the false break point, else the bottom limit is placed at 5 rows up from the reference point. If the test shows that a break point can be grouped with the reference break point, the break point is masked off to prevent it from being grouped to other joint. Based on the number of break points grouped to the reference point, joint type, such as bending point (2 joint), 3 to 8 joint can be detected. The coordinate of the center point of these break points that belong to the joint will be used as the coordinate of the joint. This joint coordinate, together with its joint type, is stored in a joint point list.

Figure 3 shows a normalized 80x80 Chinese characters with grid and all feature points and spines superimposed after feature point detection stage, where

2, 3-8 are placed at the joint coordinates to indicate their joint type respectively.

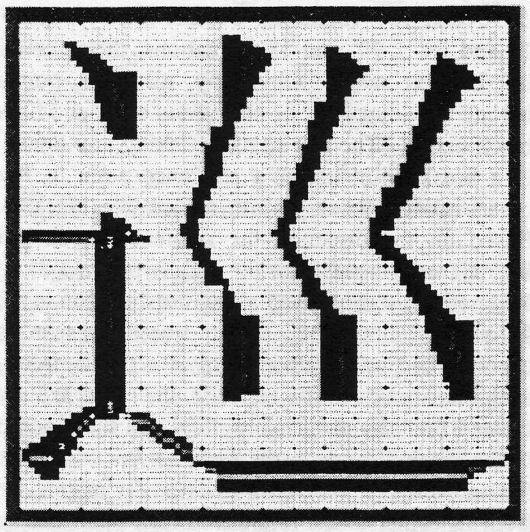


Figure 3 Feature point detection results

From Figure 3, we can see that all feature points are detected. There are some false bend points detected on some strokes. However, these false bend points will not cause any problem on later stroke section detection because checking the two-branch spine direction can easily eliminate them. As stated earlier, the bend points formed by the connection of diagonal and anti-diagonal strokes were not detected by the feature point detection. Notice that a spine sometimes is not continuous, whenever there is a sudden change of stroke width, it will cause discontinuity on the spine.

3 Stroke Section Extraction

Traditionally, a natural stroke is defined to be a continuous draw between the pen fall and the pen rise, which means that a natural stroke may contain two or more stroke sections in different directions. At this stage, we concentrate only on the extraction of stroke sections. The information obtained at this stage can be used to merge the related stroke sections to form a natural stroke or any stroke types. Four types of primitive stroke sections are defined in this paper. They are Horizontal stroke H (—), Vertical stroke V (|), Diagonal stroke D (\ ) and anti-diagonal stroke d (/), as illustrated in Figure 4.

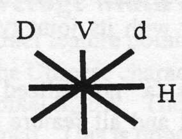


Figure 4 Four types of stroke sections

Assume that the horizontal stroke pointing to the right has 0°, and the angle is incremented in anti-clockwise manner from 0° to 360°. Each stroke type will have angle range tolerance. For H type, any stroke sections with angle or gradient between 340° and 25.5° or 160° and 205.5° will be classified as H type stroke; for V type, any stroke sections with angle falling between 65° and 110° or 245° and 290° will be classified as V type; similarly, any stroke section with angle from 110° to 160° or from 290° to 340° is a D type, and from 25.5° to 65° or from 205.5° to 245° is a d type. There are point strokes in some Chinese characters that have very short length or have stroke width almost equal to the stroke length. However, at this stage, we temporarily classify the point strokes as one of the four basic stroke types. Later at the stroke merge stage, these point strokes can be identified by their length.

For any Chinese character, each single stroke breaks at the intersections into several stroke sections. These stroke sections terminate at the break points of joints or end points. After the feature point detection stage, all feature points, joint points and their types are detected and marked. The maximum, minimum and average stroke widths are also obtained. Two linked lists of feature points are passed out from the feature point detection stage: joint-point-list and end-point-list. Each member in the joint-point-list contains the coordinates of the joint point, its joint type and pointers to break points which are grouped to this joint. On the other hand, the end-point-list contains the coordinates of all the end points of the character.

To extract the information for stroke sections, each stroke section is traced pixel by pixel along its spine. The tracing of a stroke section starts from the break point of each joint point in the joint-point-list, and the tracing will follow the path of the spine pixel by pixel. Spine tracing is different from binary object tracing in that the spine will not be continuous at the point where the width of the stroke suddenly changes. Whenever the tracing process meets a discontinuous point on the spine, a special scanning is invoked to search for next discontinuous spine point. The tracing process for each branch of the joint will stop when a valid break point belonging to a joint or an end point is met.

As mentioned in section 2, feature point detection cannot detect bend type in Figure 1e and its ±90° and 180° rotations, i.e. bend points formed by two connected diagonal and anti-diagonal strokes. This type of bend point is detected in the stroke section tracing stage. For the detection of this type of bend point along each spine, the tracing process will record the gradient between the current spine pixel and previous spine pixel as a reference gradient during tracing each branch. If gradient sign change is detected, the coordinates of the middle point where the sign change starts are recorded and the change value of subsequent traced middle points that have different sign from the reference gradient is accumulated. If the total change value is greater than a threshold which is defined as the half of the difference between maximum and average stroke width, a new bend point is set at the middle point where the sign starts change and its two neighboring middle points are used as two break points. This will prevent two merged anti-diagonal and diagonal strokes from being traced as a single stroke. The total change value will be reset to zero if the sign is changed back to the reference type and the total change value is less than or equal to the threshold value. This will tolerate small gradient sign changes caused by noise. After stroke section tracing, The required information of each stroke section is obtained. The results obtained are equivalent to traced thinning results.

The final output of the tracing process is a linked list of joint-branches as shown in Figure 5.

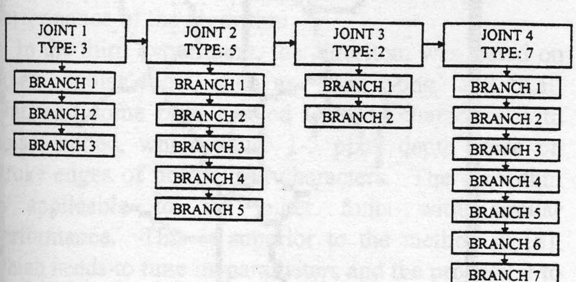


Figure 5 Joint-branch list of tracing output

The information for all the joint points and stroke sections (branches) is stored in the list. The information stored in every joint point includes the coordinates of the joint, the type of the joint, a pointer to a stroke section (branch) which is linked to this joint, and a pointer to next joint, as indicated by the arrows in Figure 3. The information stored at each stroke section(branch) which intersects other branches at the joint includes the coordinates of the starting and ending points of the stroke section, its length, its type (V, H, D, d), a pointer to a joint where this stroke section (branch) terminates, and a pointer to the next stroke section which belongs to the same joint. Those end points in the end-point-list that are not traced are considered to belong to isolated strokes. An isolated stroke is traced from one end point to another end point in a similar way and all information of isolated strokes is stored separately.

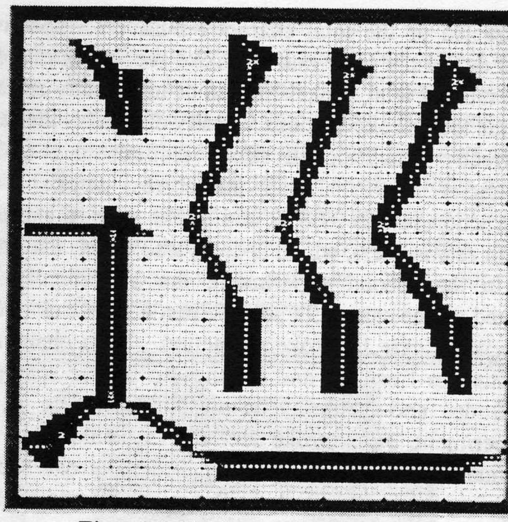


Figure 6 Stroke section tracing results

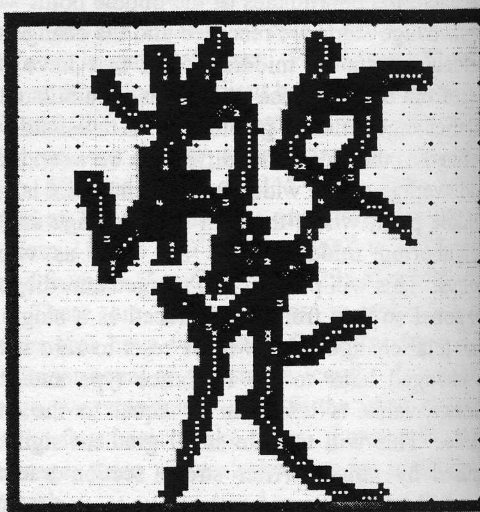
Figure 6 shows the tracing results superimposed on the Chinese character of Figure 3 based on the feature points detected. Notice that the bend points formed by the connection of diagonal and anti-diagonal strokes in Figure 3 which were not detected in feature point detection stage have been detected in the tracing stage. Every stroke branch has been traced

4 Experimental Results

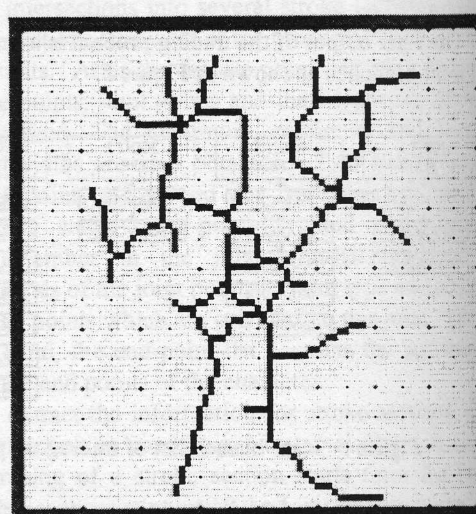
In order to show the effectiveness of the feature point based stroke section extraction method, three experiments have been conducted on a PC pentium-120Mhz based computer system with a scanner. The algorithm was tested on about 5000 most frequently used Chinese characters over the popular 'Song', 'Fang Song', 'Kai' and other bold printed fonts to extract strokes. All the characters are printed by a laser printer and scanned in by a scanner both at 300 dpi. The average size of those Chinese characters is 76x71 pixels. Each character is picked and normalized into 80x80 points in size. The algorithm is found to be applicable to all characters. To evaluate the accuracy of the algorithm, 197 characters are selected to examine the stroke section extraction results. The extracted number of stroke sections will be compared to that of human visual extraction. In these 197 characters, 4197 stroke sections extracted agree with human visual stroke extraction, and 7 stroke sections are found to be missed in 7 characters, the total stroke sections visually extracted is 4204. Define the correct stroke section or character extraction rate as the ratio of total correctly extracted stroke sections or total characters which has all stroke sections extracted correctly to the total stroke section numbers visually extracted or the total character numbers respectively, and define the error rate of stroke section extraction as the ratio of total strokes missed or incorrectly extracted to the total stroke sections visually extracted, the rate of correctly extracting all stroke sections in a Chinese character is 96.4% and the correct stroke section extraction rate is 99.8%, and stroke section extraction error rate is 0.2%, which is comparable with that in [8].

In the second experiment, the algorithm is also tested on the same 197 handwritten Chinese characters, the worst case in stroke extraction. Each character is written in free size and scanned in by the scanner at 300 dpi and normalized to 80x80 size. In this experiment, 3642 stroke sections extracted agree with human visual extraction result, and 82 stroke sections are found to be missed or incorrectly extracted in 61 characters, and the total stroke sections extracted visually is 3727. The rate of correctly extracting all stroke sections in a Chinese character is 69%, the correct stroke section extraction rate is 97.7%, the stroke section extraction error rate is 2.2%. Since different character database is used to test different methods, it is difficult to compare the performance of this method to other methods. So it is compared with a thinning method for the same handwritten characters. In this case, the same 197 handwritten Chinese characters are normalized to 80x80 size and thinned by using the pseudo 1-subcycle thinning algorithm[3]. 3671 stroke sections extracted agree with human visual stroke extraction result, 444 stroke sections are found to be missed or incorrectly extracted in 136 characters. Compared to the 3727 visually extracted stroke sections and total 197 characters, the rate of correctly extracting all stroke sections in a Chinese character is 31%, the correct stroke section extraction rate is 98.5%, and stroke section extraction error rate is 11.9%. Comparing the experiment results of the proposed algorithm with the thinning results, we can find that although the correct stroke section extraction rate for the thinning method is slightly higher than that of feature point based method, they are comparable. The stroke section extraction error rate of this method is about one fifth of that of the thinning algorithm, and the rate of correctly extracting all stroke sections in a Chinese character of this method is more than doubled compared to that of the thinning method. An example of stroke section extraction and thinning results of a handwritten Chinese character is shown in Figure 7.

From Figure 7, we can find that the algorithm is immune to small holes on or between strokes which often cause distorted strokes or extra strokes being generated by thinning algorithms, as can been see in Figure 7b. The algorithm also extracts much fewer noisy strokes due to inherent structural problems as thinning algorithms do. For example, the 5 joint in Figure 7a is formed by five stroke sections intersect at



7a



7b

Figure 7 Stroke section extraction and thinning results

it, the algorithm extracted all five stroke sections on it. However, the thinning algorithm produced a 3 joint and a 4 joint close each other, which are linked by an extra stroke section instead of generating a true five joint

From the experiment results, we can see that the proposed algorithm is much better than thinning algorithm in that it has comparable accuracy in extracting desired stroke sections and much lower error rate in producing undesired stroke sections. The reason that the algorithm is superior to thinning algorithm is that it is not sensitive to various types of stroke shapes, noise or artistic effects on strokes, thus avoiding most undesirable side effects, such as generating many noisy or distorted stroke sections, which cannot be avoided by most thinning methods. In fact, most stroke section extraction errors of this method in the experiment are due to missed stroke sections. On the other hand, although thinning methods rarely miss desired stroke sections, they usually produce many incorrect stroke sections. As compared with the method in [8], the algorithm need not adjust any parameter or procedure to suit the handwritten character, and the stroke extraction is stable for all characters.

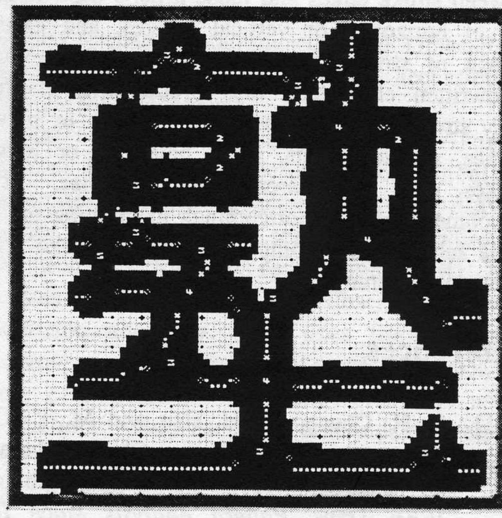
The major shortcoming for the algorithm is its speed. The algorithm is a few magnitudes slower compared to the thinning algorithm. However, as computer speed becomes faster and faster, this barrier will be eliminated in the future. The algorithm sometimes generates incorrect joint types due to imperfection in feature point grouping stage. The algorithm will also miss some stroke sections if its width is much larger than average stroke width. For example, if two parallel stroke sections merge into one stroke section due to noise, the spine for this section cannot be detected because its stroke width is much larger than average stroke width and width change threshold. There is still room to further improve the performance of the algorithm.

In the third experiment, the algorithm was tested on other popular fonts, such as 'Fang Song' and 'Kai' font, and some bold printed Chinese characters with noises added, which cause 1-3 pixel depth noise on stroke edges of normalised characters. The algorithm is applicable to the other fonts with similar performance. This is superior to the method in [8], which needs to tune its parameters and the procedure to suit each font. The algorithm is also applicable to bold printed characters with various noises. The noises generally do not affect the performance of the algorithm. Figure 8a-b show the results of the algorithm and the thinning method on a bold printed Chinese character with very bad noise on or between strokes, some of which are impossible to be eliminated by smoothing algorithms. The algorithm shows its robustness to stroke width changes and noise.

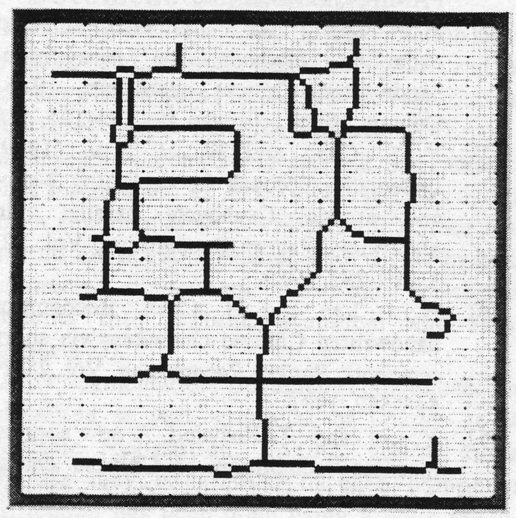
From Figure 8a, we can find that one stroke is missed, and only five very short noisy strokes are produced by the algorithm, which can be eliminated easily in the next stage to merge stroke sections into complete strokes. Comparing to the thinning results of Figure 8b, we can find that many noisy and distorted stroke sections are produced by the thinning method which are impossible to eliminate.

5 Conclusion

In this paper, a feature point based stroke section extraction method is proposed which can extract stroke sections from the binary image of a Chinese character without using thinning and smoothing algorithms. The stroke sections extracted are only four primitive stroke types (H, V, D, d) which can be used to form natural strokes or any stroke types defined in a Chinese character recognition system. Hence, this stroke section extraction method can be used for any Chinese character recognition system which use strokes as the feature to recognize Chinese characters.



8a



8b

Figure 8 Stroke section extraction and thinning results of bold character

In the experiment, the algorithm achieved 99.8% correct stroke section extraction accuracy in extracting stroke sections from printed Chinese characters of

Song font, and 97.7% correct stroke section extraction accuracy in extracting stroke sections in handwritten Chinese characters in the other worst extreme, only a little bit deterioration, and achieved comparable results in extracting stroke sections from other popular fonts of printed Chinese characters. Although the correct stroke section extraction rate is comparable to that of the thinning method for handwritten characters, the algorithm has much lower stroke section extraction error rate. The algorithm resists most noise on or between stroke edges and those artistic parts near the end of horizontal or vertical strokes, and dramatic change of stroke shapes in different fonts and in handwritten characters, thus produces much fewer noisy stroke sections caused by noise on or between strokes, artistic effects, small holes on or between strokes, even by inherent structural noise. It has overcome the most shortcomings in thinning methods while achieving the comparable correct stroke section extraction accuracy.

References

[1]. E. H. Cheng, W. H. Hsu and M. Y. Chen, Recognition of handwritten Chinese characters by modified Hough transform techniques, IEEE Trans. *Pattern Analysis Mach Intell.* PAMI-11, 429-439, 1989.

[2]. F. H. Cheng and W. H. Hsu, Three stroke extraction methods for recognition of handwritten Chinese characters, 1986 Int. Comput. Symp., Taiwan, pp. 191-195, 1986.

[3]. Y. S. Chen and W. H. Hsu, (1989), A systematic approach for designing 2-subcycle and pseudo 1-subcycle parallel thinning algorithms, Pattern Recognition, Vol. 22, No. 3, 267-282, 1989.

[4]. H. J. Lee and B. Chen, Recognition of handwritten Chinese characters via short line segments, *Pattern Recognition*, Vol. 25, No. 5, 543-552, 1992.

[5]. Y. Liu and T. Kasvand, A new approach to machine recognition of Chinese characters, *Proc. 7th Int. Joint Conf. on Pattern Recognition*, 381-384, 1984.

[6]. Y. L. Ma and J. F. Kiang, Chinese character recognition by discrete Fourier transformation, *Proc. Natn. Comput. Symp.*, Taiwan, 165-171, 1981.

[7]. S. Mori and T. Sakakura, Line filtering and its application to stroke segmentation of handprinted Chinese characters, Proc. 7th Int. Joint Conf. on Pattern Recognition, 366-369, 1984.

[8]. L. Y. Tseng, and C. T. Chuang, An efficient knowledge-based stroke extraction method for multi-font Chinese characters, *Pattern Recognition*, Vol. 25, No. 12, 1445-1458, 1992.

[9]. J. F. Wang and K. F. Wu, A new approach for handwritten Chinese character recognition, 1988 *Int. Comput. Symp.*, Taiwan, 731-736, 1988.

[10]. P. S. P. Wang and Y. Y. Zhang, Decomposition and recognition of handprinted Chinese characters, 1986 *Int. Comput. Symp.*, Taiwan 131-135, 1986.

[11]. H. Yamama, Contour DP matching method and its application to handprinted Chinese character recognition, *Proc. 7th Int. Joint Conf. on Pattern recognition*, 389-392, 1984.

[12]. Z. W. Zang, M. S. He and C. H. Ge, A research on printed Chinese character recognition based on stroke features with optical/digital hybrid realization, *Proc. 9th Int. Joint Conf. on Pattern recognition*, 573-575, 1988.

[13]. H. D. Chang and J. F. Wang, A robust stroke extraction method for handwritten Chinese characters, Intl. Journal on Pattern Recognition and Artificial Intelligence, Vol. 8, No. 5, 1223-1239, 1994.