Sealing Clay Text Segmentation Based on Radon-Like Features and Adaptive Enhancement Filters

Xia Zheng
Wei Wei
Houbing Song
Wei Li

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications

Digital Commons Citation
Zheng, Xia; Wei, Wei; Song, Houbing; and Li, Wei, "Sealing Clay Text Segmentation Based on Radon-Like Features and Adaptive Enhancement Filters" (2015). Faculty Scholarship. 321.
https://researchrepository.wvu.edu/faculty_publications/321

This Article is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Faculty Scholarship by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
Research Article

Sealing Clay Text Segmentation Based on Radon-Like Features and Adaptive Enhancement Filters

Xia Zheng, 1 Wei Wei, 2 Houbing Song, 3 and Wei Li 4

1 Department of Culture Heritage & Museology, Zhejiang University, Hangzhou, Zhejiang 310028, China
2 School of Computer Science and Engineering, Xi’an University of Technology, Xi’an, Shaanxi 710048, China
3 Department of Electrical and Computer Engineering, West Virginia University Institute of Technology, Montgomery, WV 25136, USA
4 School of Information Technologies, The University of Sydney, Sydney, NSW 2006, Australia

Correspondence should be addressed to Xia Zheng; zhengxia@zju.edu.cn

Received 9 October 2014; Revised 9 January 2015; Accepted 12 January 2015

1. Introduction

In recent years, with the development of computer, network, and multimedia technology, digital preservation of culture relics has become richer in its connotations: from the initial real-time recording and permanent preservation to multiple levels such as preservation studies, dissemination, and utilization. Digital preservation of culture relics aims at effectively protecting the body of culture relics in the process of exploring and taking their historical, artistic, and scientific value. Among culture relics, sealing clay plays an important role in revising and supplementing for ancient documents that record official systems and geographic data. Depending on image processing and analyzing technology, many researches are possible. In these researches, text extraction is the most significant one, which will affect following steps, such as feature quantification and text recognition.

As for the sealing clay text extraction, traditional method is identifying the text by ink rubbing. This method needs to make rubbings on a piece of paper from the sealing clay, in which these sealing clay is directly used and contacted, increasing the risks of damaging. For the seriously incomplete or damaged sealing clay whose preservation condition is not so good, especially, ink rubbings are not advisable from the perspective of culture relic preservation. However, taking sealing clay image as the research objective can directly extract text from the image, without the process of making ink rubbings. Since the sealing clay surface is simple, mainly including the background and the text, thresholding segmentation method is available to the research. Currently, thresholding segmentation method includes histogram-shape-based method, such as peak-valley thresholding [1] and shape modeling thresholding [2]; cluster-based method, such as minimum error thresholding [3] and fuzzy clustering thresholding [4]; entropy-based method, such as cross entropy.
thresholding [5] and fuzzy entropy thresholding [6]; and similar characteristic-based method, such as fuzzy similar thresholding [7] and semantic thresholding [8]. All of the thresholding segmentation methods above can achieve better effects with their own appropriate images. Nevertheless, there are rare systematic studies on the segmentation of sealing clay images. Thus, on the basis of previous studies on image segmentation and insistent demands of sealing clay research and preservation, the present study puts forward a new method for text segmentation based on radon-like features and adaptive enhancement filters compared with 2D Otsu, GA, and FastFCM. Experimental results show that the method proposed by this research cannot only extract text completely and accurately, but also reduce the interference of incomplete or damaged surface and uneven inward of sealing clay, which provides an appropriate solution to the problem of sealing clay text segmentation.

2. Radon-Like Features and Adaptive Enhancement Filters

2.1. Radon-Like Features. Radon-like features were originated from Radon transform [9], which is a method of integral transform. The Radon transform formula of 2D continuous function $f(x, y)$ is as follows:

$$ R(m, \tau) = \int_{-\infty}^{\infty} f(x, y = \tau + mx) \, dx. \quad (1) $$

$m$ and $\tau$ are, respectively, the slope and intercept of the straight line and Radon transform is a method of evaluating integrals along a straight line which is determined by $m$ and $\tau$. Radon transform is widely used in tomography, such as MRI (magnetic resonance imaging), and its inverse transformation is often used to reconstruct the original images. The discrete performance of (1) is

$$ R(m, \tau) = \sum_{x} f(x, y = \tau + mx). \quad (2) $$

The Radon-like features just borrowed from the main ideas of Radon transform [9]. Given an image of $I(x, y)$, the parameterized expression of line $L$ is $L(t) = (x(t), y(t))$, which is not to make the information about image $I$ completely map to line $L$ to get an accumulated value, but to make the information distribute to the different line segments of line $L$. The line segments are determined by a collection of knots. Supposing that the collection of crunodes is $(t_1, \ldots, t_n)$, so point $P$ on line $L$ which located between $(x(t_1), y(t_1))$ and $(x(t_{i+1}), y(t_{i+1}))$ is the class Radon-like feature. And its value is determined by

$$ \Psi(p, L, t_i, t_{i+1})[I(x, y)] = T(I, L(t)) \ , \ t \in [t_i, t_{i+1}] . \quad (3) $$

$T$ is an extraction function which determines how to calculate the values of Radon-like features. When fixing the rotation angle $\theta$ of line $L$ and changing its intercept, according to the (3), we can get the Radon-like features image of $I(x, y)$ which contains the values of Radon-like features and is the same size as $I(x, y)$. If rotation angle $\theta$ is also varied, every pixel $(x, y)$ of image $I$ will produce a vector of Radon-like features. As shown in Figure 1(a), two lines from different angles will intersect with the picture to form nodes and line segments. According to the extraction function, $T$, the value of Radon-like features image is highlighted in Figure 1(b).

In Figure 1(a), the line segments are represented by $a, b, c$, and $d$ which are formed by the blue line intersecting with the picture; the line segments are represented by $i, j$, and $k$ which are formed by the red line intersecting with the picture. In Figure 1(b), the value of Radon-like features will be defined by the extraction function $T$.

The usefulness of Radon-like feature depends on knots and the extraction function. For the application at hand, image analysis, knots derived from the edges in the image, provides a useful guide to the structures of the image. When a line scans through the input image, its intersections with the image’s edge map define the knots and the line segments. The choice of the extraction function provides enough flexibility to differentiate among various structures of image.

2.2. The Adaptive Enhancement Filters. The linear combinations of the Gaussian model and its derivatives are suitable for the human visual system, which has been well applied in the areas of image denoising, edge detection, and so on. The first and second derivative of Gaussian function filters are mainly used for strengthening the edges. Based on above theories, this paper adopted Leung-Malik filter bank to strengthen the structural information of images. Leung-Malik filter is the composition of polygon filter, bar filter, and punctate filter.
which are multiscale and multidirection filters, and it can
be divided into 36 first and second derivatives of Gaussian
function filters (6 directions and 3 scales) shown in first three
lines of Figure 2, 8 Gaussian Laplacian filters, and 4 Gaussian
filters, shown in the last line of Figure 2 [10].

As the strokes of Chinese character generally include five
primitive strokes, point, horizontal stroke, vertical stroke, left
slant, and right slant (as shown in Figure 3), we use the 36
second Gaussian derivative filters and 8 Gaussian Laplacian
filters of Leung-Malik filter bank to strengthen the spatial
direction information of image.

The parameters of second derivative of Gaussian filters
are redefined as 12 directions (0°–165°, an interval of 15°) and
3 dimensions ((σx, σy) = (1/2, 2), (2/9, 6), (2/5, 6/9, 2)).
The scales of 8 Gaussian Laplacian filters are 2, 3/2, 2√3/2, 4,
3√3/2, 6, 6√3/2, and 12. After strengthening of image I by these
filters from Leung-Malik filter bank, it will achieve 44 energy
images Eσx,β, as shown in

\[ E_{\sigma,\beta} = I \ast F_{\sigma,\beta}, \]  

where \(*\) is a convolution symbol and Fσ,β is the filter from
36 second derivatives of Gaussian filters and 8 Gaussian
Laplacian filters.

The construction of the actual LM filter bank needs to
determine the maximum window’s size of the filter bank:
K \times K. The value of K decides the fine level of enhance-
ment. Therefore, this paper adopts an adaptive enhancement
mechanism to decide the window size by estimating noise of
the original image. The detailed information of a sealing clay
image is not complicated, which is only interfered by damage,
stains, and so on. Thus, in order to boost the system speed,
this paper adopts the method of fast noise variance estimation
[11]. For an image measuring W \times H as seen in I, the formula
of variance for noise is presented as

\[ \sigma_n = \sqrt{\frac{1}{2} \frac{\pi}{6(W-2)(H-2)} \frac{1}{\sum_{x=1}^{W} \sum_{y=1}^{H} |I(x,y) \ast N|}}, \]  

where \( \sigma_n \) is the noise variance; \(*\) is the convolution symbol; N is a 3 \times 3 template. Consider

\[ N = \begin{pmatrix} 1 & 1 & -2 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{pmatrix}. \]  

On the basis of distribution of noise variance, we determine
threshold parameters T1 and T2. According to (7), we get the
value of K:

\[ K = \begin{cases} 5, & \sigma_n > T_1 \\ 9, & T_2 < \sigma_n < T_1 \\ 13, & \sigma_n < T_2. \end{cases} \]  

3. The Process and Description of Algorithm

When calculating the Radon-like feature values of an image,
the function is as follows:

\[ T(I, L(t)) = \int_t^{t+1} \frac{E(L(t)) \, dt}{\| L(t_{i+1}) - L(t_i) \|_2}, \quad t \in [t_i, t_{i+1}]. \]  

This function calculates the mean values of image E’s pixel
values locating along line L and situating between knots t_i and
t_{i+1}. E is the maximum energy image:

\[ E(x,y) = \max_{\sigma,\beta} E_{\sigma,\beta}(x,y). \]  

Besides, this paper extracts the edge of E by using Canny
operator to generate knots. 12 lines with different direction
(0°–360°, with an interval of 30°) are used to scan images
to locate knots and determine Radon-like features. The basic
process of algorithm is as follows.

Step 1. For an input image I, the method of fast noise variance
estimation is used to calculate \( \sigma_n \). Based on the values of \( \sigma_n \)
and T1 and T2, windows parameter K is determined.

Step 2. Enhancement is achieved by using the 36 second
Gaussian derivative filters and 8 Gaussian Laplacian filters in
the LM filter bank. 44 energy images Eσ,β are obtained and
the maximum energy image E is got.

Step 3. Canny operator is used to extract the edge of E, and
an edge graph is got.

Step 4. Based on the edge graph, 12 lines with different
direction are used to scan image E. The knots are determined
by the intersections of the scan line and the edge graph.
The pixels between two knots in $E$ are the found Radon-like features. Their values are calculated according to the extraction function which is defined in formula (8). 12 scan lines will make 12 different Radon-like feature pictures.

**Step 5.** Thresholding is conducted on the mean value of the 12 Radon-like feature pictures, and the segmentation result is got.

In Step 2, the adaptive LM filter bank is used to enhance the texture features and it also means to characterize a texture by its responses to a set of filters. In order to reduce the dimensionality of the filter response space, the maximum energy image $E$ over both scales and orientations is selected from $E_{\sigma,\theta}$. Figure 4 shows some results in Step 2. In Figure 4, the top of (a) is a sealing image and its some energy images from 12 directions at scale $(\sqrt{2}, 3\sqrt{2})$ are listed in (b). The maximum energy image is shown in lower (a). We can see that maximum energy image contains enhanced main parts and local details.

In Step 3, canny operator is used to get edge of $E$. Though the operator can be directly applied to input image $I$, its result is less complete. It can be seen in Figure 5 that the result of $E$ has more complete edges.

In Step 4, based on the edge graph and $E$, Radon-like features are found. After that, these features’ values are calculated according to the extraction function and make 12 different pictures. Figure 6 shows the 12 Radon-like feature pictures of sealing clay image shown in Figure 4. It can be seen that the features of text in 12 directions are selected and enhanced.

### 4. Experimental Results and Analysis

Based on the above algorithm, this paper selects 30 sealing clay images to conduct segmentation experiment. These images are obtained from three types of sealing clay with different materials and color. Because the algorithm proposed by this paper is to divide the text of sealing clay image, before the experiment, preprocessing of the original sealing clay image is conducted. The central text area is extracted from the sealing clay images and different areas are unified into the same size measuring 128 in length and 128 in width. To demonstrate the advantages of the proposed method we implemented it with Matlab R2010b and compared it with 2D Otsu [12], Genetic Algorithm [13], and FastFCM [14]. Figure 7 lists five sealing clay images using different segmentation methods, among which the last line is the result yielded by manual division under the guidance of experts [15].

We can see that 2D Otsu and GA yield similar results which do not completely separate the text information from the noise, with many areas integrating with each other especially in Figures 7(a) and 7(b). In comparison, FastFCM not only extracts the major form of characters but also suppresses parts of the noise, yielding more detailed information comparing with the former two ways, which is shown in Figures 7(a), 7(b), 7(c), and 7(d). However, looking at every single character, the specific strokes and the intactness of structures are not satisfactory and Figures 7(a), 7(b), 7(c), and 7(d) can illustrate this problem. It should be noticed that, comparing the methods of 2D Otsu, GA, and FastFCM, only GA can obtain the approximate form of the Chinese character “司” in Figure 7(b). Therefore, GA has a better recognizability. For Figure 7(e), GA and FastFCM yield similar results which contain much noise. Though 2D Otsu can obtain better result, the form of characters is incomplete. The method proposed by this paper can solve the problems mentioned above, which not only clearly separates the major body of a character but also has a better performance in terms of richness and intactness. What is more, this method can greatly filter the noise so as to yield clearer and more definite results [16–20].

As is shown in Table 1, by comparing the operation time, we found that 2D Otsu method requires the minimum time. The time required by FastFCM method is longer than that of 2D Otsu method, but is shorter than that of the method used in this paper. GA method required the longest operation time. The method we used in this paper was not advantageous in operating the time, which needed twice the time of FastFCM method and three times the time of 2D Otsu method. Time
Figure 7: The segmentation results using different methods including 2D Otsu, GA, FastFCM, and the proposed method.
Table 1: Operation time of different method.

<table>
<thead>
<tr>
<th>Method</th>
<th>(a) (ms)</th>
<th>(b) (ms)</th>
<th>(c) (ms)</th>
<th>(d) (ms)</th>
<th>(e) (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D Otsu</td>
<td>86</td>
<td>93</td>
<td>95</td>
<td>86</td>
<td>89</td>
</tr>
<tr>
<td>GA</td>
<td>541</td>
<td>527</td>
<td>520</td>
<td>513</td>
<td>518</td>
</tr>
<tr>
<td>FastFCM</td>
<td>248</td>
<td>249</td>
<td>243</td>
<td>232</td>
<td>241</td>
</tr>
<tr>
<td>The method in this paper (with 12 scanning directions)</td>
<td>476</td>
<td>481</td>
<td>472</td>
<td>468</td>
<td>473</td>
</tr>
</tbody>
</table>

Table 2: Average values of ME, TPR, FPR, and EP of 4 methods.

<table>
<thead>
<tr>
<th></th>
<th>2D Otsu</th>
<th>GA</th>
<th>FastFCM</th>
<th>The proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average value of ME</td>
<td>0.2163</td>
<td>0.2545</td>
<td>0.1536</td>
<td>0.0982</td>
</tr>
<tr>
<td>Average value of TPR</td>
<td>87.3%</td>
<td>84.7%</td>
<td>91.2%</td>
<td>96.5%</td>
</tr>
<tr>
<td>Average value of FPR</td>
<td>10.1%</td>
<td>12.7%</td>
<td>7.3%</td>
<td>4.7%</td>
</tr>
<tr>
<td>Average value of EP</td>
<td>12.3%</td>
<td>15.2%</td>
<td>8.6%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

Figure 8: The relationship between average operation time and the number of directions in the proposed method.

was mainly spent in scanning and searching the Radon-like features in different directions. The more directions we have scanned, the more detailed Radon-like features we would find and the longer time it would take. Figure 8 shows the relationship between average operation time and the number of directions we scanned. According to several experiments in this paper and with the principle to keep the operation time as short as possible, we found that it could achieve preferable result when the number of directions was 12.

To measure the result of segmentation and further validate effectiveness of the method, the ground-truth image, which was manually segmented under the guidance of experts, was used as a standard reference pattern to compare with the results made by other 4 methods. Misclassification error (ME) was adopted in the measurement of specific differences, which was defined as follows [21]:

\[
ME = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o| + |F_o|},
\]

where \(F_o\) and \(F_T\) refer to pixels of the text area in standard reference figure and test pattern, respectively, and \(B_o\) and \(B_T\) stand for pixels of the background region in standard reference image and test pattern, respectively. \(|F_o \cap F_T|\) represents the pixel of statistical texts which are segmented correctly, while \(|B_o \cap B_T|\) means the pixel of statistical backgrounds which are segmented correctly. The value of ME ranges from 0 to 1, with 0 standing for completely correct segmentation while 1 means completely incorrect segmentation. Therefore, the smaller the ME value is, the more precise the segmentation is. Moreover, the average values of true positive rate (TPR), false positive rate (FPR), and error probability (EP) are also used to measure the performance, which are calculated the following, respectively:

\[
TPR = \left( \frac{TP}{TP + FN} \right) \times 100\% ,
\]

\[
FPR = \left( \frac{TN}{FP + TN} \right) \times 100\% ,
\]

\[
EP = \left( \frac{FP + FN}{TP + FN + FP + TN} \right) \times 100\% ,
\]

where TP, TN, FP, and FN are true positives, true negatives, false positive, and false negatives, respectively. In this paper, 30 images are tested, and the average values of ME, TPR, FPR, and EP obtained from 4 methods are shown in Table 2.

It is apparent from Table 2 that, in all experiments, the proposed method is superior compared to other methods, with the least ME (0.0982), FPR (4.7%), and EP (4.2%) and the maximum value of TPR (96.5%), while FastFCM method has a slightly lower accuracy on segmentation, with its values of ME (0.1536), TPR (91.2%), FPR (7.3%), and EP (8.6%). By comparison, 2D Otsu method and GA method have the lowest accuracy on segmentation with higher values of ME (0.2163, 0.2545), FPR (10.1%, 12.7%), and EP (12.3%, 15.2%). All the results indicate that the proposed method is the most efficient to segment the text of sealing clay in terms of accuracy and stability.
5. Conclusion

Traditional text extraction method of sealing clay needs to contact the sealing clay object directly, which increases the risk of damage. Due to the gradually mature digital technology of cultural relics, this paper proposes a text extraction method based on digital images to facilitate sealing clay research. This method combines Radon-like features and adaptive enhancement LM filter bank and is implemented by two major steps of enhancing the filtering and calculating Radon-like feature images. In experimental testing phase, the result obtained from the method used in this paper is compared with those of 2D Otsu method, GA method, and FastFCM method. According to the result, the method proposed in this paper could not only effectively reduce noise like breakage and muddiness, but also extract text with higher definition, completeness, and accuracy. The disadvantage of this method is its long execution time. The major reason is that the directional scanning process which identifies Radon-like features is relatively time consuming. The more directions it scans, the longer execution time it takes. Therefore, we should promote the scanning strategy to improve the execution speed. Besides, to identify the Radon-like feature, more abundant partial statistical information can be further utilized to extract function so as to acquire more detailed information to enhance the representativeness of Radon-like features. As for applying the method, we can extend the method to solve the problem of segmentation in other types of images, such as blood vessels image segmentation, cells image segmentation, and road image segmentation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partially supported by the National Natural Science Foundation of China (Grant no. 61202198) and Social Science Foundation of Zhejiang Province (Grant no. 11JCWH13YB). This program is supported by Scientific Research Program Funded by Shaanxi Provincial Education Department (Program no. 2013Jk139) and supported by China Postdoctoral Science Foundation (no. 2013M542370) and supported by the Specialized Research Fund for the Doctoral Program of Higher Education of China (Grant no. 20136118120010).

References


