Effective Binarization of Document Images with Uneven Shading

XIAOHUA ZHANG† Member, HEMING HUANG† Non-member, NING XIE‡ Non-member, YUELAN XIN§ Non-member

(Received November 08, 2015, revised December 24, 2015)

Abstract: With rapid popularization of mobile cameras, capturing a document and storing it become easy. However, when the document is illuminated under poor conditions, the document image may appear uneven shading. In that case, it is difficult to restore texts for character recognition and document analysis etc. In this paper, we propose an effective and simple approach to remove uneven shading from a document image and binarize it. First, a local uneven shading is estimated using a constant-time weighted median filter. Then, reflectance is computed by removing the uneven shading based on the retinex theory, followed by a bandpass thresholding to partially binarize the clean document image. Finally, the unbinarized pixels are classified into text and background by employing a graphcut approach in which the graph is constructed by considering a contrast map. A host of experimental results demonstrate that the proposed method runs effectively and has plausible performance.

Keywords: Uneven shading, Weighted median filter, Retinex Theory, Graphcut, Binarization

1. Introduction

With rapid popularization of mobile cameras, capturing a document and storing it become easy in our daily life. However, when a document is illuminated under poor conditions, the document image may appear unevenly shaded and is hard for binarization to read the text characters. It is known that the document image binarization is a very important preprocessing for character recognition and document analysis etc. Over the past decades a lot of ink has been spent on binarizing document images. These methods fall into two distinct types: global thresholding[1] and local adaptive thresholding[2]~[5]. Otsu’s method is a well-known global thresholding, however, it requires that the histogram of a document image should be bimodal. Unfortunately, this is not true for the document image in question when uneven shading ingredients are contained in it. Local adaptive thresholding, such as the famous Sauvola’s method, seems good when there exists uneven shading in image. However, even the parameter is well tuned, the Sauvola’s method cannot perform satisfactorily when the background in the document image is unevenly shaded.

Because of the uneven shading in the background, both global and local adaptive thresholding methods fail to binarize images. If a document image holds an uniform background, it is often easier to be binarized. This observation drives us to remove or at least reduce the complexity of the uneven shading by making it almost uniform in gray level.

This paper concentrates on binarizing document images taken under poor illumination and thus having very uneven shading in the background, as shown in Fig. 1(a). If a document image contains very complicated uneven background, the gray levels are very different at each pixel location. It is assumed that the document image is in gray scale in this paper. If a color image is given, it is easy to convert it to gray scale. Suppose that the media (such as paper) is in white, while the “inked” texts are in black. The texts may be interfered with dark shading.

There have existed several attempts so far trying to remove the uneven shading before binarization of a document image. Gatos et al. estimated the shading surface directly[6], using the binarized result obtained by Sauvola’s technique. A transform approach was reported that transformed an image into curvelet field and then removed the uneven shading[7]. Lu et al.[8] proposed a surface modeling approach for thresholding poorly illuminated document images by employing a single polynomial model to fit variations in the shading surface. The problem is that a single polynomial cannot properly fit a complicated surface. An improvement of Lu’s method was given by using piecewise polynomials to estimate the shading surface[9] more accurately. However, the polynomials are based on each image block, deciding block size is not a convenient task. Recently, a new approach was reported, focusing on shading estimation, which used Gaussian kernels to estimate the uneven shading, followed by a removal of shading from the document image[10]. This method has superior performance to others, however, it computed a Gaussian model at each pixel, thus it is time-consuming.

The approaches mentioned above perform well in most
cases, however, they cannot handle the complicated uneven shading or take too much time to acquire the final binarized results. To solve these problems, this paper presents an effective and simple method to remove uneven shading from a document image. The proposed method does not need a mathematical model to fit the shading surface. It estimates local lightness for every pixel by using a constant-time contrast weighted median filter, followed by a calculation of reflectance based on the retinex theory. If the image contains darker shading, the recovered foreground may become noisy. A guided image filter is employed to reduce noises. After shading removal, the background in
the image becomes almost uniform. Then, it is partially binarized by employing a bandpass thresholding based on Sauvola’s method, by using threshold interval instead of a single threshold value. Finally, the remained unbinarized pixels are classified into texts and background by using contrast weighted graphcut approach. Intensive experiments show that the proposed approach is superior to others.

This paper is organized as follows: Section 2 gives an overview of the proposed method. Section 3 describes how to remove the uneven shading from document images. A bandpass thresholding is introduced in Section 4 and the graphcut for binarization is provided in Section 5. Section 6 shows experimental results and Section 7 draws a conclusion.

2. Overview

Figure 2 illustrates a flowchart of the proposed method. A rectangle shows image data at different stages, while a rectangle with rounded corners means a process. A dotted arrow indicates where a process is required; A solid arrow illustrates data flow. A small circle with a cross inside means multiple data are used.

Given a gray scale document image \( F(x, y) \) with uneven shading such as those shown in Figure 1(a). A constant-time median filter is firstly applied to the original image to estimate shading at each pixel. The median filter is weighted by a contrast map \( F_{cont}(x, y) \) computed from the original image \( F(x, y) \). The resultant is a lightness map \( L(x, y) \) constructed from the whole illumination of all pixels. By combining the original image \( F(x, y) \) with the lightness map \( L(x, y) \), most of uneven shading components are removed or weakened based on the retinex theory, and an image \( R(x, y) \) with almost uniform background is obtained. At the same time, the strokes of text are darkened by controlling a gain of darkness. A bandpass thresholding based on Sauvola’s method is applied to separate text strokes partially from the background, by considering threshold interval instead of a single threshold value. Then a temporary binarized result \( I_{bin}(x, y) \) is acquired. Finally, after computing again a contrast map \( R_{cont}(x, y) \) from the “clean” image \( R(x, y) \), a graphcut approach is employed to classify the remained unbinarized pixels into texts and background and obtain the final result \( I_{b} \). The contrast map \( R_{cont}(x, y) \) is used in weight setting in the graph of graphcut.

In the following sections, Fig. 1 is used as a running example to explain the processes of the proposed method.

3. Uneven Shading Removal

It is believed that in most cases the uneven shading is a big obstacle in thresholding document images. For example, when using Otsu’s method to binarize the image in Fig. 1(a), the result is in Fig. 1(b). It is obvious that the texts in this resultant image cannot be well recognized. The local adaptive Sauvola’s method works well on this document image, but a lot of noises remained, as shown in Fig. 1(c). It is known that removing previously the uneven shading from document images can greatly improve thresholding results. Suppose the original document image is notated as \( F(x, y) \), and the dynamic range is \([0, 255]\). A contrast map[11] is constructed from the original image \( F(x, y) \) and the contrast at each pixel is computed in a window centered at the current pixel:

\[
F_{cont}(x, y) = \frac{F_{max}(x, y) - F(x, y)}{F_{max}(x, y) + \epsilon},
\]  

where \( F_{max} \) is the maximum inside the window and \( \epsilon \) is a tiny positive value for avoiding the denominator being zero. The window size depends on the width of strokes. A size from 3 to 10 is used in all our experiments. The contrast defined in Equation (1) has a property that it assigns a larger value to the pixel on strokes and a smaller value to the pixel on background. The computed contrast map from Fig. 1(a) is shown in Fig. 1(g). Note that \( F_{max}(x, y) \) can be found by using an order filter, which is in the same framework as median filter introduced later.

The retinex theory says that the observed image intensity is expressed as a multiplication of illumination \( L(x, y) \) (incident light) and reflectance \( R(x, y) \) of scene[12]. A median filter is used for estimating illumination as a lightness map.

Moreover, in order to improve robustness for estimating the lightness accurately, a weighted median filter is adopted. The histogram in a kernel window $W$ is computed as follows:

$$h(i) = \sum_{(x,y) \in W} w(x,y) \delta(F(x,y) - i),$$

where $w(x,y) = 1.0 - F_{cont}(x,y)$; $\delta(\cdot)$ is a Kronecker function which is one if $F_{cont}(x,y) = i$, and zero otherwise. $i$ represents a pixel level. It is known that the median filter is time consuming when the kernel window size becomes large. This problem is solved by employing a constant-time algorithm[13], which is independent of the window size and thus has time complexity in $O(1)$. The weighted median filter is embedded in the same framework as constant-time median filter. It is important that pixel values are weighted by their contrasts, thus pixels on the background have high possibility than pixels on strokes to be as output of median filter. The size of filter kernel is set to triple the average stroke width which is computed using stroke width spectrum(SWS)[14].

Figure 4 compares the effect of the general median filter and the contrast weighted median filter. The red curve data is taken from one row of the original image shown in Fig. 1(a). The blue and green curves are the estimated background by using the non-weighted and weighted median filters, respectively. It is obvious that the green curve fits well the background, while the blue curve contains values close to the values of pixels on the strokes. The estimated uneven shading by applying the above method to the whole image in Fig. 1(a) is shown in Fig. 1(h). It is observed visually that the uneven shading is well estimated. The estimated uneven shading is used as a lightness map $L(x,y)$. The reflectance $R(x,y)$ is computed as $R'(x,y) = F(x,y)/L(x,y)$ according to the retinex theory. This operation removes or reduces the uneven shading $L(x,y)$ from the original image $F(x,y)$. Since the resultant image $R'(x,y)$ may appear in low contrast, the darkness of text strokes is enhanced using a gain control. The final reflectance is modified as

$$R(x,y) = 255(\alpha (R'(x,y) - 1.0) + 1.0),$$

where $\alpha$ controls the gain of darkness. In our experiments, $\alpha \in [1.0, 3.0]$. Since the median filter is essentially a smooth filter, the reflectance image may appear noisy after shading removal using division operation. To solve this problem, a fast guided image filter[15][16] is applied to reduce noises. Although it cannot remove noises completely, the pseudo edges on the background are weakened. The kernel size of the guided filter is $3 \times 3$, the regularization coefficient is $\epsilon = 0.001$. After these computations, the final result is an image with almost uniform background, as shown in Fig. 1(i). Compared with the original image, this image is easier to binarize.

### 4. Bandpass thresholding

Sauvola’s technique[3] calculates a single threshold value for a pixel $(x,y)$ in a local window with size $w \times w$ as

$$T(x,y) = \mu(x,y) \left( 1 + k \frac{\sigma(x,y)}{R_{\text{max}}} \right),$$

where $\mu(x,y)$ and $\sigma(x,y)$ are the mean and standard deviation, $R_{\text{max}}$ is the maximum standard deviation and is set to 128 for gray scale image, while $k$ is a parameter and is always positive. Generally its value is less than 0.5. Since the standard deviation $\sigma(x,y)$ is less than $R_{\text{max}}$, a smaller parameter $k$ produces a larger threshold value $T(x,y)$, thus captures pixels on weak strokes. The Sauvola’s original method uses threshold value defined in Equation (4) to determine whether a pixel $(x,y)$ belongs to the foreground or background. However, a pixel value close to the threshold may contain uncertainty and the pixel is probably misclassified due to the uncertainty and noises. Figure 5 shows an example of the histogram of a small $11 \times 11$ region. The pixel values are distributed in the range $[156, 255]$. The computed threshold value is 204 when using Sauvola’s technique, shown in this graph as a green vertical line. It may not be a proper threshold value, because it incorrectly classifies pixels with value close to this threshold. It is very difficult to observe visually where the threshold is. A solution is to binarize pixels into white or black as much as possible, and leave pixels with value close to the threshold unclassified. To this end, a bandpass thresholding is adopted that uses a threshold interval instead of a single threshold value.

$$I(x,y) = \begin{cases} 0 & R(x,y) < T(x,y) - t \\ 255 & R(x,y) \geq T(x,y) + t \\ R(x,y) & \text{otherwise} \end{cases},$$

where $t$ is a parameter to control the width of the band, that is, the bandwidth is $2t$. The word “band” means a
Effective Binarization of Document Images with Uneven Shading

5

range in the intensity space. It is obvious that the image is partially binarized. While the pixels with value in the range \([T(x, y) - t, T(x, y) + t]\) are not classified, pixels with value out of this range are classified into white or black. If \(t = 0\), the above thresholding becomes the original Sauvola’s method. In all our experiments, \(t\) is set to 15.

From this point of view, the above bandpass thresholding is a generalization of the common thresholding. The computation of the threshold value \(T(x, y)\) is not limited to the Sauvola’s method. It certainly can be computed by other techniques such as Otsu’s method. Figure 1(j) illustrates the bandpass thresholding result \(I(x, y)\) from the image in Fig. 1(i). Note that the unclassified pixels are marked as red for visualization. One part of image is enlarged and shown in Fig. 3 for confirmation.

5. Graphcut for Binarization

After bandpass thresholding, there may remain many pixels unclassified yet, since their values fall in the range \([T(x, y) - t, T(x, y) + t]\) are not classified, pixels with value out of this range are classified into white or black. If \(t = 0\), the above thresholding becomes the original Sauvola’s method. In all our experiments, \(t\) is set to 15. From this point of view, the above bandpass thresholding is a generalization of the common thresholding. The computation of the threshold value \(T(x, y)\) is not limited to the Sauvola’s method. It certainly can be computed by other techniques such as Otsu’s method. Figure 1(j) illustrates the bandpass thresholding result \(I(x, y)\) from the image in Fig. 1(i). Note that the unclassified pixels are marked as red for visualization. One part of image is enlarged and shown in Fig. 3 for confirmation.

Figure 6: Original document images and binarized results.

In the following graphcut segmentation, the costs between nodes of graph are computed by considering pixel contrast. A contrast map is calculated before graphcut segmentation. Note that the contrast map \(R_{cont}(x, y)\) is constructed from the image \(R(x, y)\), the result after shading removal, using the same formula in Eq. (1). This is a different contrast map with that produced at shading estimation stage. Figure 1(k) illustrates a contrast map constructed from Fig. 1(i). Compared with the map in Fig. 1(g), this map is cleaner due to shading removal. It is natural that pixels with similar contrast should be classified into the same category. This will be considered in the graph construction in graphcut segmentation.

Considering the above temporary binarized image \(I(x, y)\), denote the set of pixels on text strokes as \(O\), the set of pixels on background as \(B\), and the set of all pixels as \(P\). It is obvious that \(O \in P\), \(B \in P\), and \(O \cap B = \emptyset\). To show assignments of pixels \(p \in P\), a binary vector is denoted as \(A = (A_1, A_2, \ldots, A_P)\). Each component \(A_p\) of vector \(A\) belongs to either object “obj” \(O\) or background “bkg” \(B\). The vector \(A\) defines a segmentation of image \(I(x, y)\) in which each pixel belongs to the text stroke or background. This is
accomplished by minimizing the cost function

$$E(A) = \lambda \cdot R(A) + B(A),$$

where $R(A)$ is a regional term and $B(A)$ is a boundary term. The parameter $\lambda$ controls the relative importance of regional term $R(A)$ versus boundary term $B(A)$. This optimization belongs to the graph segmentation. To binarize the given image $I(x, y)$, a graph $G < V, E >$ is constructed. The node corresponds to the pixel $p \in P$ of the image. Two additional nodes are added to represent source $S$ (text stroke object) and sink $T$ (background). The edges $E$ consist of two kinds of undirected links. The first kind is links between neighborhoods called $n$-links and the second kind is links between pixel nodes of $P$ and terminal node $S$ or $T$ called $t$-links. Generally, the weights of edges in $E$ are given in a manner similar to the technique in reference [17]. However, when $[p, q] \in P, [p, q] \notin O \cup B$, the weight between pixel nodes is computed as follows:

$$w_{pq} = e^{-\frac{(I_p - I_q)^2}{2\sigma_1^2}} \cdot e^{-\frac{(R_{cont}(p) - R_{cont}(q))^2}{2\sigma_2^2}} \cdot \frac{1}{dist(p, q)},$$

where $I_p$ and $I_q$ are pixel values at $p$ and $q$ in the image $I(x, y)$, $R_{cont}$ and $R_{cont}$ are contrasts at $p$ and $q$ in the contrast map $R_{cont}(x, y)$, respectively. $dist(p, q)$ is a geometric distance between pixels $p$ and $q$. The standard deviations $\sigma_1$ and $\sigma_2$ are computed from the temporary binarized image $I(x, y)$ and the contrast map $R_{cont}(x, y)$, respectively.

On the other hand, the weight between $p$ and terminal $S$ or $T$ is defined as

$$w_{p,S} = (1 - c(p)) \cdot \lambda \cdot R_p(\text{"bkg"}),$$

and

$$w_{p,T} = c(p) \cdot \lambda \cdot R_p(\text{"obj"}),$$

Figure 7: Experimental results
Figure 8: Original document image, intermediate and final results.

Table 1: OCR performances of different methods applied to the document images

<table>
<thead>
<tr>
<th>Methods</th>
<th>Fig. 1(a)</th>
<th>Fig. 6(a)</th>
<th>Fig. 6(e)</th>
<th>Fig. 6(i)</th>
<th>Fig. 7(a)</th>
<th>Fig. 8(a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Otsu</td>
<td>7.69</td>
<td>9.03</td>
<td>4.39</td>
<td>3.83</td>
<td>3.33</td>
<td>4.63</td>
</tr>
<tr>
<td>Sauvola</td>
<td>79.12</td>
<td>88.36</td>
<td>91.20</td>
<td>95.92</td>
<td>95.21</td>
<td>76.85</td>
</tr>
<tr>
<td>S. Poly.</td>
<td>20.07</td>
<td>33.73</td>
<td>63.73</td>
<td>77.93</td>
<td>10.98</td>
<td>10.42</td>
</tr>
<tr>
<td>P. Poly.</td>
<td>95.60</td>
<td>98.57</td>
<td>90.10</td>
<td>98.74</td>
<td>91.36</td>
<td>95.86</td>
</tr>
<tr>
<td>G. kern.</td>
<td>95.60</td>
<td>98.57</td>
<td>91.20</td>
<td>98.08</td>
<td>98.84</td>
<td>92.22</td>
</tr>
<tr>
<td>Proposed</td>
<td>94.51</td>
<td>97.39</td>
<td>95.70</td>
<td>98.32</td>
<td>94.62</td>
<td>97.92</td>
</tr>
</tbody>
</table>

respectively. $c(p) \in [0, 1]$ is the contrast at pixel $p$, taken from the contrast map $R_{cont}(x, y)$. A larger contrast indicates the corresponding pixel has a higher possibility to be assigned to a stroke object. $R_{p}(\text{obj})$ and $R_{p}(\text{bkg})$ are penalties for classifying pixel $p$ into the text stroke and background, respectively. The previously classified pixels at the stage of bandpass thresholding are used as a prior information to calculate the intensity distributions: $P_r(I|O)$ and $P_r(I|B)$. Then the penalties are set to be negative log-likelihoods:

$$R_{p}(\text{obj}) = -\ln P_r(I|O)$$  \hspace{1cm} (10)

$$R_{p}(\text{bkg}) = -\ln P_r(I|B).$$  \hspace{1cm} (11)

The distributions are computed using Gaussian models in our experiments. After graph construction, the efficient graphcut algorithm in reference [18] is employed to obtain the optimal binarization result. Fig. 1(1) shows the final binarized result of Fig. 1(a). It is observed that the original image is well binarized, even though there exists very heavy uneven shading.

6. Experimental Results

To verify the effectiveness of the proposed method, a host of experiments have been conducted on document images with uneven shading in different patterns. One experimental result is shown in Fig. 1, which has been used as a running example in the explanation of the proposed method. Figure 1(b) is a threshold result using Otsu’s method; Fig. 1(c) is a binarized result using the classical Sauvola’s technique. In both of them, the text strokes and background are not well separated. Fig. 1(d) is the binarized result by using a single polynomial to estimate the uneven shading. It is obvious that a single polynomial cannot well fit complicated shading. Fig. 1(e) and Fig. 1(f) are the binarized results by using piecewise polynomials and Gaussian kernels to esti-
mate the uneven shading, respectively. Both of them acquired pleasurable results by visual observation. However, piecewise polynomial models need to set a proper block size. If the block size is small, the binarization may take long time. On the other hand, the Gaussian kernel models compute coefficients at each pixel using a row or a column, thus take long time. Figure 1(f) is the final binarized result of the proposed method.

Experiments on images with different uneven shading patterns are further conducted. Figure 6 shows three of them. In this figure, the first column shows three original document images with different uneven shading patterns. The second column illustrates the estimated uneven shadings using the algorithm described in Section 3. The third column demonstrates images with uneven shading removed. By observing these images, most of uneven shadings have been removed or reduced, although there still exists weak shading. Compared with the original images, these images become easier to binarize because the backgrounds are almost uniform gray level. The rightmost column in Fig. 6 demonstrates the final resultants after bandpass thresholding and graphcut segmentation. The parameter settings are the same as those for Fig. 1. It is observed that the text strokes are well restored from the unevenly shaded document images in our experiments.

Figure 7 shows experimental results using different methods on a document image in Fig. 7(a) with sharp shading edge. It is certain that Otsu global thresholding failed to binarize this image due to the non-bimodality as shown in Fig. 7(b). Figure 7(c), Fig. 7(d), Fig. 7(e) and Fig. 7(f) are binarized results obtained by using Sauvola’s technique[3], single polynomial model[8], piecewise polynomial models[9], and Gaussian kernel models[10], respectively. By carefully observing these results, it is known that some background pixels are classified into foreground in Fig.7(c); some pixels on text strokes near the shading edges are classified into background in Fig.7(d) and Fig.7(e). The result by using the Gaussian kernel models is pleasurable, but this method takes long runtime. The proposed method estimates the shading very well as shown in Fig.7(g). After removing the shading from the original image, the clean image with almost uniform background is in Fig.7(h). The final binarized result is illustrated in Fig.7(i). Visual observation indicates that the result produced by the proposed method is superior to those by other methods.

Figure 8 is another experiment which further demonstrates the effectiveness of the proposed method. Figure 8(a) is an original image containing a different shading pattern. The contrast map constructed from the original image is in Fig. 8(b) and the intermediate result after shading removal is shown in Fig. 8(c) in which both strong and weak strokes exist. The bandpass thresholding result is demonstrated in Fig.8(d), while the clean contrast map constructed from Fig.8(c) is shown in Fig.8(e). The final binarized result is illustrated in Fig.8(f).

The proposed algorithm was implemented on a notebook PC with Windows OS 8.1, CPU 1.8GHz, and memory 4.00GB. The C++ language is used for coding. Since pixel level ground truth is not available, OCR performances are compared by using the ABBYY FineReader 11 software[19] to recognize text characters. The comparisons are shown in Table 1. The proposed method is compared with the other five methods: Otsu, Sauvola, Single polynomial (S. Poly.), Piecewise polynomial (P. Poly), and Gaussian kernel (G. Kern.). The recognition rates are in words (“W”) and letters (“L”). Figures in parentheses after “W” and “L” are the number of words and letters, respectively, in the original images. The largest rates are marked in bold font in each column. It is observed that the proposed method achieved superior recognition performance for most of images. Note that since Japanese characters and alphabetical letters are mixed in Fig.7(a), the words mean Chinese characters, Hiragana, and Katakana, the letters only mean alphabetical letters. All methods have lower recognition rates when applied to Fig. 8, due to that the original image is blurred with low quality.

The runtime comparisons are demonstrated in Table 2. It is certain that Otsu’s method is the fastest, but with the worst recognition rates. Note that Sauvola’s original technique is very time-consuming. A quick algorithm is implemented by using integral image in this paper, so it takes short time to run on images. Both single polynomial and piecewise polynomial models cost moderate runtime, while Gaussian kernel model takes the longest one. As for the proposed method, since a constant-time weighted median filter was adopted, the time complexity is independent of filter size; the graphcuts process employed a library from [18], which is also very fast. As shown in the table, the proposed method has lower time cost than Gaussian kernel model, while having superior recognition rate.

7. Conclusion

An efficient approach for restoring text strokes from document images with uneven shading was proposed in this paper. First, for removing the uneven shading, a weighted median filter was adopted whose time complexity is independent of window size of the filter and thus the filter...
Effective Binarization of Document Images with Uneven Shading

has constant-time complexity. Then, the classical Sauvola’s technique was employed to compute threshold value, but a bandpass threshold interval was used instead of a single threshold value. Finally, the unclassified pixels were further assigned using graphcut segmentation. The experiments demonstrate that the plausible results are acquired despite the document images with complicated uneven shading.

Acknowledgment
The authors would like to thank Jiangtao Wen from Yan-shan University, China, who kindly permitted us to use his original images in Fig. 1(a), and in Fig. 6(th the first column). This work is part of the research supported by the Chunhui project of the Education Ministry of China under Grant Nos. Z2012100 and Z2014020, and the National Nature Science Foundation of China under Grant No. 61462072.

References


Xiaohua Zhang (Member) received a B.S. degree in computational mathematics from JiangXi University, China, in 1984, an M.S. degree in computer science and engineering from Jilin University of Technology, China, in 1990, and a Dr. Eng. degree in information science and engineering from the Tokyo Institute of Technology, Japan, in 2000. After working at NHK Engineering System Inc. from 2000 to 2003, he joined the faculty of the Hiroshima Institute of Technology as an associate professor. His research interests include computer graphics, image processing, computer vision, pattern recognition, and machine learning. He is a member of the IEEE, the IEICE, the ITE, and the IIEE.

Ning Xie (Non-member) received the M. E. from Department of Computer Science, Tokyo Institute of Technology, Tokyo, Japan, in 2009, a Dr. Eng. degree from the same department in 2011. He is currently a lecture at Tongji University, China. His research interests include machine learning, computer graphics and computer vision.

Heming Huang (Non-member) received a B.S. degree in mathematics from Shanxi Normal University, China, in 1992, an M.S. degree in computer applications from LanZhou University, P. R. China, in 2004, and a Dr. Eng. degree in pattern recognition and artificial intelligence from Southeast University, China, in 2014. He is currently a professor at Qinghai Normal University. His research interests include Tibetan information processing and pattern recognition.
Yuelan Xin (Non-member) received a B.S. degree in physics from Shaanxi Normal University, China, in 1997, an M.S. degree in computer science and engineering from Qinghai Normal University, China, in 2009. She is currently an associate professor at Qinghai Normal University. Her research interests include image processing, pattern recognition.