Research of adaptive threshold model and its application in iris tracking

Qijie Zhao, Dawei Tu, Rensan Wang, Daming Gao
School of Mechatronical Engineering, Shanghai University, Shanghai 200072, P.R.China

ABSTRACT

The relationship between gray value of pixels and macro-information in image has been analyzed with the method in statistical mechanics. After simulating and curve fitting with the experiment data by statistic and regression method, an adaptive threshold model between average gray value and image threshold has been proposed in terms of Boltzmann statistics. On the other hand, the image characteristics around the eye region and the states of eyeball also have been analyzed, and an algorithm to extract the eye feature and locate its position on the image has been proposed, furthermore, another algorithm has been proposed to find the iris characteristic line and then to coordinate the iris center. At last, considering the cases of head gesture, different head position, and the opening state of eyes, some experiments have been respectively done with the function based on the adaptive threshold model and the designed algorithms in eye-gaze input human-computer interaction (HCI) system. The experiment results show that the algorithms can widely be applied in different cases, and real-time iris tracking can be performed with the adaptive threshold model and algorithms.

Key words: image processing, adaptive threshold, Boltzmann statistics, iris tracking, characteristic line

1. INTRODUCTION

One of the most frequently used methods in image processing is thresholding. it distinguishes objects from the background in images [1]. However, what we have in practice is a mixture distribution of objects and background information in an gray-level histogram, but not the two separate distributions, so we may have to make some assumptions about the forms of the two distributions to simplify the problem, under these assumptions, we can then determine the optimal threshold according to the statistical decision theory [2]. Some methods based on gray-level histogram have been proposed in the past to select threshold, such as automatic global threshold method using Renyi’s entropy [1], the two-dimensional entropy method based on the two-dimensional histogram [3], but the first entropy based method was proposed by Pun [1][4], and Pal and Pal have also considered thresholding using some second-order statistics [5]. Iris tracking technique plays an important role in human-computer interaction and biological feature recognition applications, and some methods to track iris have been proposed, a flow-based method of tracking eye corners, irises, and eyelids using simple head, iris and eyelid models has been described in reference [6][7], the deformable templates based on parameters of eyes to track the motion of eyes have been given by reference [8][9]. A kind of iris tracking method for sequence images has been proposed in the light of Kalman Filter and probability evaluation [10], and this method has been improved by using two Kalman Filter in reference [11]. In this paper,
considering the real-time image processing, a kind of thresholding model which adapts to different illumination condition, has been constructed with statistical method, after analyzing its characteristics for eyes gazing objects, a kind of iris tracking method based on characteristic lines has been proposed, and then some experiments have been done with the threshold model and iris tracking algorithm designed in this paper.

2. THRESHOLD MODELING

Global threshold selection methods usually use the gray-level histogram of the image. The optimal threshold is determined by optimizing some criterion function obtained from the gray-level distribution of the image [1]. The gray-level histogram of image describes the relationship between microcosmic pixels and macro-objects by gray values distribution, so it may be a good base to set up a kind of threshold model according to gray-level histogram. Boltzmann statistics has successfully analyzed the relation among the distribution of micro particles and the states of macro objects, and it has been used in many technique fields [12], by this idea, a kind of threshold model is constructed as following:

\[
T(g_A) = T_u - \frac{T_u - T_d}{1 + e^{\frac{g_A - g_0}{k}}}
\]

(1)

Where, \(T_u\) and \(T_d\) are the upper bound of threshold and lower bound of threshold respectively, \(g_A\) is the average gray value, \(g_0\) is a statistical constant, \(k\) is the Boltzmann factor.

\[\begin{align*}
p(i) & = \frac{n_i}{MN} \\
& \text{where } n_i \text{ represents the number of pixels with gray value } i, \ M \text{ and } N \text{ are respectively the width and height of the image. In Fig.1, } G_m \text{ is the gray value where the probability density function } p(i) \text{ has maximum value, that is,}
\end{align*}\]
\( G_m = \max \{ p(i) \}, i = 0,1,...,255 \), \( g_A \) denotes the average gray value of pixels whose gray-levels are below \( G_m \), so \( g_A \) can be expressed by \( g_A = \frac{\sum_{i=1}^{N_B} \sum_{j=1}^{N_B} g(i,j)}{N_B} \), where \( g(i,j) \) is the gray value of pixel \( (i,j) \), \( N_B \) is the number of pixels whose gray levels are below \( G_m \). \( p_u \) and \( p_d \) are two local minimums of probability density function \( p(i) \), and their corresponding gray-levels are respectively \( \{ T_u = \\min \{ p(i) \}, g_A < i < G_m \} \) and \( \{ T_d = \\min \{ p(i) \}, 0 < i < g_A \} \). If there are not obviously local minimums (e.g. \( p_u \), \( p_d \)) of probability density function \( p(i) \) in gray-level histogram, the upper bound threshold \( T_u \) and the lower bound threshold \( T_d \) can be defined as following:

\[
\begin{align*}
T_u &= g_A + \frac{G_m - g_A}{2} = \frac{G_m + g_A}{2} \\
T_d &= g_A - (T_u - g_A) = \frac{3g_A - G_m}{2}
\end{align*}
\]

Statistical constant \( g_0 \) is given by \( g_0 = \frac{T_u + T_d}{2} \), and Boltzmann factor can be determined by

\[
k = \frac{(p_u - p_d)}{T_u - T_d} = \frac{255(p_u - p_d)}{T_u - T_d}
\]

By this way, a threshold \( T_1 \) can be gotten in this gray range where the gray-levels of pixels are below \( G_m \), and another threshold \( T_2 \) can also be determined with the similar method in the gray range, where the gray-levels of pixels are above \( G_m \), at last, an optimal threshold can be obtained by comparing \( T_1 \) and \( T_2 \) according to the characteristics of the original image.

From the above threshold selection model and method, some experiments have been done with images including face in different illumination condition to extract the iris region, and the average gray values of the sample images in this experiment are between 100 and 230. In order to get a better threshold, three groups of threshold data have been recorded in the light of the information distribution in gray-level histogram and the thresholding results of images, and the three groups of threshold data respectively describe the lower bound threshold, medial threshold and the upper bound threshold for a same image, and then the average value of the three groups data has been calculated to form a group of
average threshold. The simulating and fitting experiments have been done with the four groups of experimental threshold data to analyze the correlation among the average gray values of image and the four groups of data. In this experiment, the sample image only contains the main features in face, and the feature information keeps stable, therefore, the local average gray values have been substituted by global average gray values, the simulating and fitting results are shown in Fig. 2.

![Graphs showing the relation between average gray value and different thresholds](image)

(a) The relation between average gray value and average threshold  
(b) The relation between average gray value and middle threshold  
(c) The relation between average gray value and upper-bound threshold  
(d) The relation between average gray value and lower-bound threshold

Fig.2 the simulating and fitting results

In Fig.2, (a), (b), (c), and (d) respectively illustrate the correlation among average gray values and the average threshold (AVGTHRD symbolized as ▽), the medial threshold (THRD1 symbolized as □), the upper-bound threshold (THRD2 symbolized as ○), and the lower-bound threshold (THRD3 symbolized as △), where the solid line is the fitting result, and the dashed line represents the simulating result. To compare the correlation of the four groups of threshold and the average threshold, the fitting experiment has been done, as shown in Fig.3.
In Fig.3, the experiment data ○, ▽, □, and △ and their corresponding curves ①, ②, ③, and ④ respectively describe the correlation among average gray value and the upper-bound threshold (THRD2), the average threshold (AVGTHRD), medial threshold (THRD1) and the lower-bound threshold (THRD3). From the fitting results, the fitting curve of average threshold and average gray value illustrates the correlation between threshold and average gray values better, and p1, p2, p3 and p4 in the fitting result respectively represent $T_d$, $u_T$, $g_0$, and k in equation (1), so $T_d = 55.87$, $u_T = 117.29$, $g_0 = 189.19$ and $k = 10.94$.

### 3. IRIS TRACKING METHOD

When eyes gaze at object, the upper and the lower part of iris can be occluded by eyelids, but the left and right boundary of iris can not be occluded, as shown in Fig.4 (a), the dashed curves represent the eyelid edge. $l_i$ is a chord length of the circle formed by iris boundary, in the set, $L = \{l_1, l_2, \ldots, l_i\}$, there must be a maximum value $l_{max} = \max\{l_i\}$, and $l_{max}$ is also the characteristic line of iris in image, let $P$ as the center of $l_{max}$, if $l_{max}$ represents the diameter(d) of the circle formed by iris boundary, then $P$ is nearly the center of iris(c), so as long as $P$ can be located, the iris tracking can be performed.
In Fig.4(b), \( p_1(x_1, y_1) \) and \( p_2(x_2, y_2) \) are the end points of \( l_{\text{max}} \), and \( P(\bar{x}, \bar{y}) \) is the center of \( l_{\text{max}} \), so \( \bar{x} \) and \( \bar{y} \) can be defined as the following:

\[
\begin{align*}
\bar{y} &= y_1 = y_2 \\
\bar{x} &= x_1 + (x_2 - x_1) / 2
\end{align*}
\]

(3)

Where \( x_2 - x_1 \) is the length of \( l_{\text{max}} \).

Let \( I(x, y) \) be the binarized image of the eye region while iris is completely naked, as shown in Fig.5, and \( \varphi(x, y) \) be the feature-depicting image of \( I(x, y) \), pixels set \( R = \{(x, y)|\varphi(x, y) = 1\} \) denotes the iris and pupil in image, \( R^c \) is the supplementary set of \( R \), and \( P(x, y) \) express a pixel of image, that is,

\[
\varphi(x, y) = \begin{cases} 
0, & P \in R^c \\
1, & P \in R
\end{cases}
\]

(4)

The integration for the image \( I(x, y) \) along x-axis and y-axis is as following:

\[
f_j(x, y_j) = \int_{x_i}^{x_2} \varphi(x, y_j)dx \quad \text{and} \quad f_i(x, y) = \int_{y_i}^{y_2} \varphi(x_j, y)dy
\]

(5)

For the discrete pixels in image, the integration can be substituted by addition, so equation (5) can be expressed by:

\[
f_k = \sum_{i=x_1}^{x_2} \varphi(i, j_k) \quad \text{and} \quad f_z = \sum_{j=y_1}^{y_2} \varphi(i_z, j)
\]

(6)

where \( j_k \in [y_1, y_2] \), \( i_z \in [x_1, x_2] \). In Fig.5, \( H \) and \( V \) respectively represent the distribution of the integration results along x and y axis, and the maximum integration value along x-axis is \( f_z \), its corresponding position is \( i_z \), by
the same way, there must be \( f_k \) along y-axis, and its corresponding position is \( j_k \). According to the characteristic of 
\[ \varphi(x, y), \quad (i_z, j_k) \] is the intersection of the characteristic lines, and it is also the tracking target \( P(\bar{x}, \bar{y}) \), that is,
\[
\bar{x} = i_z \left| \varphi_z = \max \left\{ \sum_{j=j_k}^{n} \varphi(i_z, j) \right\} \right.
\]
\[
\bar{y} = j_k \left| \varphi_y = \max \left\{ \sum_{i=i_k}^{m} \varphi(i, j_k) \right\} \right.
\]
(7)

In order to get a much more precise location of characteristic lines, the influences caused by eyebrow, eyelash and eye corner must be eliminated or decreased (shown in Fig.6), and then extracting the feature region of eyes (in Fig.7).

The algorithm to eliminate eyebrow and extract the characteristic area of eyes is designed as following:

1) In the binarized image, make the integration of \( \varphi(x, y) \) along x-axis, 
\[ f_{k} = \sum_{i=0}^{n-1} \varphi(i, j_k), \]
and determine the global maximum 
\[ GV_{\max} = \max \left\{ \sum_{i=0}^{n-1} \varphi(i, j_k) \right\}, \quad y_0^' = \hat{f}_k \bigg|_{f_k=GV_{\max}} \]
where \( j_k = 0, 1, \cdots, n-1 \), \( m \) and \( n \) are the width and height of the image.
2) For \( j_k > y_0 \) and \( j_k < y_0 \), two local minimum values which is beside and near to the global maximum \( GV_{\text{max}} \), can be determined, that is, \( LV_{\text{min}_1} = \min \left\{ \sum_{i=0}^{m-1} \phi(i, j_k) | j_k < y_0 \right\} \) and \( LV_{\text{min}_2} = \min \left\{ \sum_{i=0}^{m-1} \phi(i, j_k) | j_k > y_0 \right\} \), and their coordination are \( y_1 = j_{k1} \big|_{f_i = LV_{\text{min}_1}}, y_2 = j_{k2} \big|_{f_i = LV_{\text{min}_2}} \).

3) For \( j_k < y_1 \), if there is a local maximum of the integration result, e.g. \( LV_{\text{max}} = \max \left\{ \sum_{i=0}^{m-1} \phi(i, j_k) | j_k < y_1 \right\} \), then \( y_0 \) is position of the eyebrow, and \( y_1 \) is the lower boundary of eyebrow, so \( y_1 \) is the separate line to eliminate the eyebrow, \( sptline = y_1 \). For \( j_k > y_2 \), if there is a local maximum of the integration result, e.g. \( LV_{\text{max}} = \max \left\{ \sum_{i=0}^{m-1} \phi(i, j_k) | j_k > y_2 \right\} \), then \( y_0 \) is the location of iris, and \( y_2 \) is the upper boundary of iris region, so \( y_2 \) is the separate line to eliminate the eyebrow, \( sptline = y_2 \).

4) For pixels \( P(x, y) \), if \( y > sptline \), set its gray value as background gray-level to eliminate the eyebrow, and make integration of \( \phi(x, y) \) along \( x \) and \( y \). The horizontal and vertical maximum of the integration, e.g. \( HT_{\text{max}} \) and \( VT_{\text{max}} \), can be determined by the method in step 1), and their corresponding coordination, e.g. \( h_0 \) and \( w_0 \) can also be gotten (in Fig.7).

5) Make use of the method in step 2), these coordination \( h_1, h_2, w_1, \) and \( w_2 \) which are beside and nearest to the local minimum \( HT_{\text{max}} \) and \( VT_{\text{max}} \).

6) This area which is formed by \( h_1, h_2, w_1, \) and \( w_2 \), is the feature region \( S \), in Fig.7) of eyes to be extracted.

To decrease the influences caused by eyelash, eyelid and the spots formed by reflex light on cornea, the medial filter with template \( 3 \times 1 \) has been used to eliminate noise, and the fill algorithm is designed to fill the cavity caused by reflection spot on cornea(shown in Fig.8(a)), make the horizontal and vertical integration of the processed image(in Fig.8(b)), at last, the iris center can be determined by the method in section 3.1 and equation(3), and the tracking result in original image can be gotten (shown in Fig.8(c)).
The procedure of filling algorithm is described as the following:

1) The upper and lower boundary (e.g. $y_i$, $y_b$) of iris in binarized image can be determined with the feature extracting algorithm step 5.

2) Let $\{(i, j) | G(i, j) = 1\}$ be the pixels set of iris, and $\{(i, j) | G(i, j) = 0\}$ be background or cavity pixels set, for $y \in (y_b, y_i)$ and $y = j_k$, scan the binarized image from left to right, $(i = 0, 1, 2, \ldots, m)$, and if $(G(i, j_k) = 1)$ then break, the left boundary $x_l$ has been found, and then scan the image from right to left, $(i = m, m - 1, \ldots, 0)$, and if $(G(i, j_k) = 1)$ then break, the right boundary $x_r$ has been found, where $m$ is the width of the image, and $j_k$ is the y-coordinate for a row in the image.

3) $i \in (x_l, x_r)$, if $(G(i, j_k) = 0)$, seeds[stackpoint $s$] $\leftarrow$, and add the seeds into stack.

4) if (not (stack = NULL)) then seed = seeds[$k$], where $(k = 0, 1, \ldots, stackpoint s)$. For the pixels whose x-coordinates meet this condition, $x_l < i < seed.x$ and if $(G(i, j_k) = 0)$, so fill this pixel, that is, let $G(i, j_k) = 1$, by the same way, scan and fill the image, while $seed.x < i < x_r$, where $seed.x$ denotes the x-coordinate of seed.

5) For the pixels whose y-coordinates meet the condition, $y \in (y_b, y_i)$, and $y = j_k$, $(j_k = y_b + 1, y_b + 2, \ldots, y_i)$, repeat step 2), 3), and 4).
4. EXPERIMENTS

The experiment system consists of a PC with a 866 MHz CPU, 256M memory, an image capture board and a CCD camera, some experiment results have been obtained with the proposed model and method in this paper, as shown in Fig.9 and table 1.

(a) Illumination is dark, and thresholded at T=77  

(b) Illumination is light, and thresholded at T=105

(c) Illumination is lighter, and thresholded at T=120

Fig.9 the results in different illumination condition

In Fig.9,(a), (b), and (c) respectively illustrate the experiment results in different illumination condition, and the images from left to right in each group are original image, gray-level histogram, binarized and eliminated eyebrow, the feature region of eyes, and the tracking result image, moreover, the intersection point of the white lines in the tracking result image represents the determined iris center. The parameters and image information in (a), (b), and (c) have been recorded in table 1, and these information includes global average gray value (GAVG), local average gray value (\( g_\delta \)), and upper-bound threshold (\( T_u \)), lower-bound threshold (\( T_l \)) and optimal threshold (\( T^* \)).
Table 1 the parameters of thresholding

<table>
<thead>
<tr>
<th>Image №</th>
<th>GAVG</th>
<th>$g_A$</th>
<th>$T_u$</th>
<th>$T_d$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>92</td>
<td>73</td>
<td>83</td>
<td>63</td>
<td>77</td>
</tr>
<tr>
<td>(b)</td>
<td>128</td>
<td>102</td>
<td>111</td>
<td>93</td>
<td>105</td>
</tr>
<tr>
<td>(c)</td>
<td>163</td>
<td>125</td>
<td>115</td>
<td>135</td>
<td>120</td>
</tr>
</tbody>
</table>

Considering the variation of the distance from eyes to camera, head gesture and eyes opening state in practice, the experiments to naked eyes, wearing glasses, head rotation, and eyes half open have been done with the proposed model and method in this paper, the images of size 300×300 have been used, and the processing speed is 12 fps, consequently, the results have been shown in Fig.10. The intersection point of the white lines represents the iris center determined with the proposed iris-tracking algorithm. In the figure, images in above and below row respectively express the experiment results when the subject is naked of eyes and wearing glasses, and the first three images in each row are the results that eyes move near to camera and eyes position in camera vision field changes from left-top corner to right-bottom corner, moreover, the last two images in each row illustrate the situations that eyes are half opening and head rotation.

![Fig.10 the experiment results of different eye states](image)

From the results in Fig.9, Fig.10 and table 1, the threshold model based on Boltzmann statistics can adapt to different illumination condition, the main feature of eyes in original image has been kept in binarized image, and the iris center determined with the proposed method is basically right. Furthermore, the iris tracking can be performed in these situations, such as naked eyes, wearing glasses, head rotation and so on, as long as the iris has not been occluded completely, the iris tracking can be performed with this method.

5. CONCLUSIONS

In this paper, a threshold model based on Boltzmann statistics has been proposed, and some simulation and fitting experiments have been done to determine the parameters in the threshold model, after comparing the experiment results,
a threshold model applied for iris tracking has been set up, and which can adapt to different illumination condition. In addition, a kind of iris tracking method based on characteristic lines has been proposed, to verify the effect of threshold model and iris tracking method in practice, some experiments have been done, and the results show that the threshold model based on Boltzmann statistics can adapt to different illumination condition, moreover, the main feature of eyes in original image has been kept in binarized image, and the iris center determined with the proposed method is basically right. Furthermore, the iris tracking can be performed in these situations, such as naked eyes, wearing glasses, head rotation and so on, and this method can be applied to both single-frame image processing and sequence image processing, as long as the iris has not been occluded completely, the iris tracking can be performed with this method.

REFERENCES

2. Hong Yan, Unified formulation of a class of image thresholding techniques, Pattern Recognition, vol.29, No.12, pp.2025-2032, 1996