Ear Recognition by Major Axis and Complex Vector Manipulation

Ching-Liang Su
Department of Industrial Engineering and Technology Management
Da Yeh University
#168, University Road, Dacun, Chang-Hua, 51505, Taiwan
Tel: 886-4-851-1888 ex 4121, Fax: 886-4-851-1270
cls2@mail.dyu.edu.tw

Received March 15, 2016; revised December 15, 2016; accepted January 16, 2017; published March 31, 2017

Abstract

In this study, each pixel in an ear is used as a centroid to generate a cake. Subsequently the major axis length of this cake is computed and obtained. This obtained major axis length serves as a feature to recognize an ear. Later, the ear hole is used as a centroid and a 16-circle template is generated to extract the major axis lengths of the ear. The 16-circle template extracted signals are used to recognize an ear. In the next step, a ring-to-line mapping technique is used to map these major axis lengths to several straight-line signals. Next, the complex plane vector computing technique is used to determine the similarity of these major axis lengths, whereby a solution to the image-rotating problem is achieved. The aforementioned extracted signals are also compared to the ones that are extracted from its neighboring pixels, whereby solving the image-shifting problem. The algorithm developed in this study can precisely identify an ear image by solving the image rotation and image shifting problems.

Keywords: Major axis length, vector manipulation, and ear recognition

https://doi.org/10.3837/tiis.2017.03.022 **ISSN : 1976-7277**

1. Introduction

This study uses the "moment technique" to determine the major axis length of a cake image. This major axis length is used to recognize an ear. A 16 extracting-circle template is generated to extract the ear features, which are the lengths of major axes. These extracted signals are part of a complex process, using an algorithm to solve the image rotation problem. This 16 extracting-circle template is combined together. Moreover, these extracting-circles center on the ear hole area to extract the ear features. By comparing the extracting signals, which are extracted at different places, one can differentiate and identify different ears. This method is used to solve the image-shifting problem.

Over the past twenty years, researchers have devoted much attention to identify objects, including the Gabor Orientation feature [8, 10, 18], feature segmentation [3], feature detection [1, 2, 4, 7, 11, 14, 16], feature selection [5, 6], feature recognition [15], illumination invariance [9, 11, 14], Gabor filter phase technique [8]. Previous research has establish ear models but with different degrees of variance. This variance is attributed to the fact that human beings have different ears. Subsequently, it is very hard to establish precise ear models. A contributing factor to this variance is that when capturing ear images the ears may be at different positions and poses, thus the image shifting and rotating issue arises. Previous methods have not fully resolved the image shifting and/or rotation problem, the new method proposed in this study solves these two problems.

From past research, the SIFT technique is used to recognize an ear, as shown in **Fig. 1.1**. For employing this SIFT technique to recognize an ear, the feature points of an ear must be located. Subsequently, these feature points are connected. Next, these connection lines are used to identify an ear. Often it is very difficult to locate the feature points of the ear, as these points of an ear are very difficult to obtain. Sometimes, the SIFT technique is unable to establish a complete SIFT pattern.

Figs 1.2 and **1.3** illustrate usage of the ear contour to recognize an ear. For the ear and skin behind the ear, they have similar color. Subsequently it is very difficult to define the contour of the ear due to the cross blending of colors in the ear area. **Fig. 1.3** illustrates how an ear contour can be used to determine the height and width of an ear. As mentioned earlier, it is difficult to fully locate the contour of the outer ear; subsequently the ear contour is not well defined so accurate to decide the length and width of an ear is impossible to obtain.

Fig. 1.4 shows the 3D ear. 3D ear image possesses the depth information of an ear. Subsequently, from this depth information it is easier to locate the key feature points of an ear. With this information it is easier to employ the geometry comparison technique to recognize an ear. Also, as a different reference for comparison methods, the Fisher ear [11] does not give a good solution to image shifting and rotating problems.

Fig. 1.5 shows a pair of pliers and a staple case with its major axis respectively. An object has its own unique major axis, which can serve as a critical feature to identify the object. As shown in this figure, the pliers have its own unique major axis length and the staple case has its own unique major axis length as well. These unique major axis lengths provide information to identify the pliers and the staple case with greater consistency and accuracy. **Fig. 1.6** shows three pliers with three different poses. However, they are from the same source. The major axis lengths of these three pliers are identical. Although the orientations of them in the images are different, they have the same major axis lengths. Similarly, **Fig. 1.7** shows three different staple cases, which are from the same source. For them, since they are from the same source,

the lengths of these staple cases are identical. This shows that although the objects are rotated, the major axis length is preserved maintaining consistency across images. This property is used to solve the image-rotating problem.

When performing the ear recognition, especially when performing the pixel-by-pixel comparison, one individual pixel should posses its own unique salient feature; therefore, an ear can be recognized. As is common, the ear color in an ear image is always plain white. When comparing these plain white pixels, an ear cannot be recognized. **Fig. 1.8** shows three different era images. The left ear image (a) is in the normal position, the center ear image (b) is obtained by rotating the left ear image (a) clockwise a little bit, and the right ear image (c) is obtained by rotating the center ear image (b) a little bit further. These three ears belong to one person and they need to be identified. **Fig. 1.9** shows the same ears of **Fig. 1.8** again. However, in this figure images (a), (b), and (c) contain one large white ring, respectively. Moreover, near the ear lobe each ear image has a black slant line cake. **Fig. 1.10** shows three large rings. Each ring contains four black spots. These black spots represent the symbols of the black slant line cakes in **Fig. 1.9**. The major axis length of the black slant line cake of **Fig. 1.9** is computed and the number is presented in **Fig. 1.10**, as 22, 8, 3… 10…. As the images (a), (b), and (c) of **Fig. 1.9** are rotated, the major axis lengths, 22, 8, 3… 10 in **Fig. 1.10** are rotated. The summing result of the major axis lengths, 22, 8, 3… 10 in **Fig. 1.10** is the same. This implies that the summing result of the major axis lengths of figures (a), (b), and (c) in **Fig. 1.9** is the same. This method is used to recognize an ear. **Fig. 1.9** shows only one white circle resides in one image. Actually, in this study, for every pixel in an ear image, each pixel is used as a centroid to generate the aforementioned white circle. The small cakes with the black slant lines are part of this process of computing all contributing axes for an aggregate total to recognize an ear.

The flow chart for this study is illustrated in **Fig. 1.11**. The remainder of this report consists of four sections. In Section 2, major axis length computing is presented. In Section 3, vector computing is presented. In Section 4, an ear signal is extracted from various locations and a signal comparison is performed. In Sections 5 and 6 conclusions for this study are presented.

Fig. 1.1. SIFT.

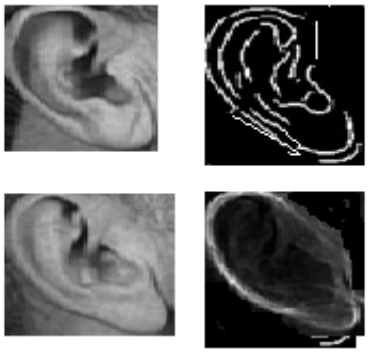

Fig. 1.2. Ear edgeSIFT.

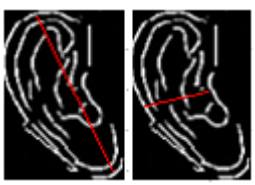

Fig. 1.3. Ear height and Width.

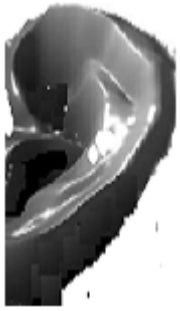

Fig. 1.4. Dears.

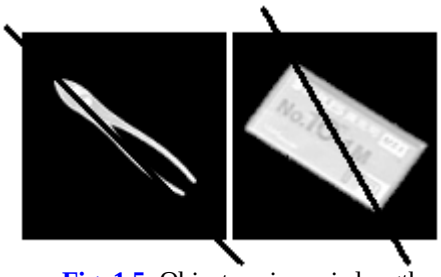

Fig. 1.5. Object major axis length.

Fig. 1.6. Pilers with different orientations.

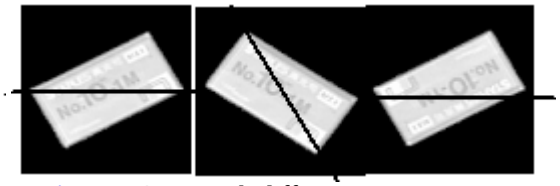

Fig. 1.7. Cases with different orientations.

Fig. 1.8. Ear with different orientations.

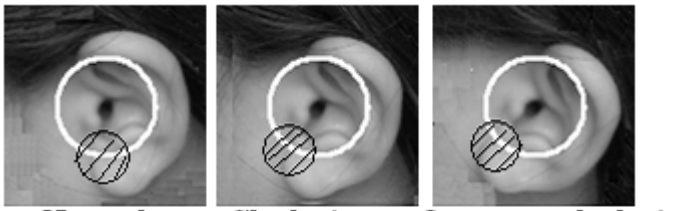

Fig. 1.9. Black slash cake areas to compute major axis length

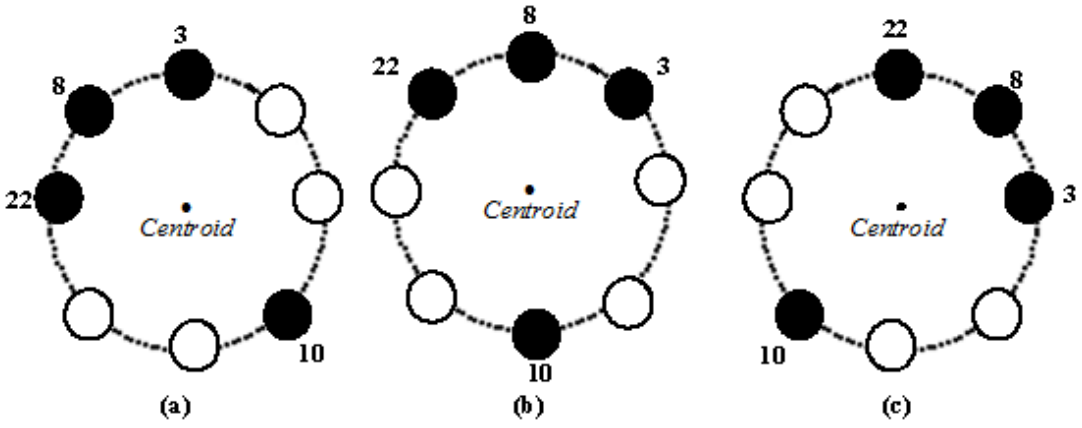

Fig. 1.10. The extracted major axis length, (a), (b), (c) are coming from the identical source.

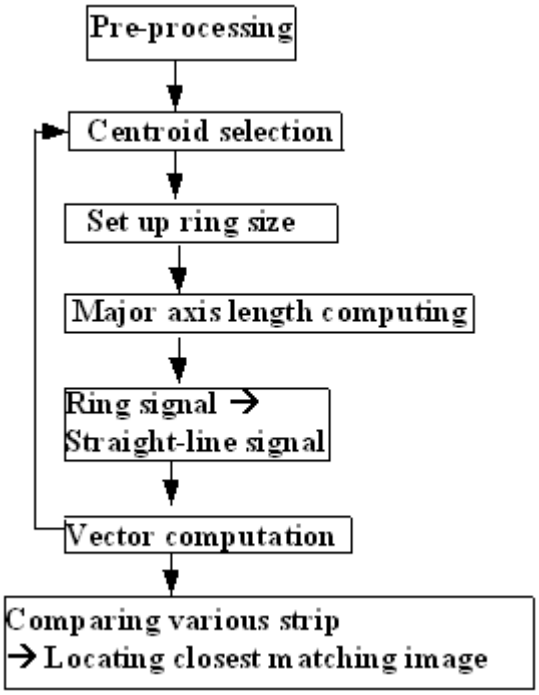

Fig. 1.11. Processing steps.

2. Major axis length computing

Fig. 2.1 shows three ears again and they belong to the same person. If the left ear is in the normal position (image a), one can rotate this ear (image a) clockwise to obtain the middle ear (image b). Furthermore, rotating the left ear (image a) a little bit more, the right ear (image c) can be obtained. As checking these three ear-images, one can see there is a white ring inside the ear-image. Moreover, near the ear lobe, there is a small black slash cake for computing major axis length. The length of the major axis of this small black slash cake is calculated and this length represents the feature of this small black slash cake. This length serves as a feature to recognize an ear. As mentioned earlier, all these three cakes are located near the ear lobes and the major axes of these three cakes are calculated. One can move the cakes to the top sections of the ears and the other major axis lengths can then be obtained. Similarly, these additional lengths obtained can be used to identify an ear. In this study, each point that is located on the circumference of the white circle serves as a centroid. Furthermore, this centroid is used to generate a cake. Following the previous discussion, the major axis length of each cake can be determined. As mentioned before, these major axis lengths are used to recognize an ear. After obtaining the major axis length, the length is saved in the position of the centroid of each cake itself. Therefore, the white circumference of the white circle contains the information of the lengths of the major axes. Subsequently, these signals are used to recognize an ear. The technique used to obtain the major axis of a black slash cake is described as follows.

Equation 2.1 is used to compute the parameters m_{pq}, wherein variables p and q may vary different case, as 0, 1, or 2. Location (x, y) represents the location of a pixel inside the aforementioned black slash cake, to which $b(x, y)$ represents the gray level of this pixel.

Location (x_c, y_c) is the centroid of this cake, which is computed by equation 2.2, in which m_{10} and m_{00} are used to compute x_c and y_c.

The centroid of the black slash cake is defined as (x_c, y_c). Equation 2.3 is used to compute variables a, b, and c, wherein x' represents the horizontal distance of a pixel from its current position to its centroid (x_c, y_c) and y', vertical distance. By further analysis, equation 2.4 can be derived. In equation 2.4, θ represents the orientation of an aforementioned cake. By determining the values of θ, x', and y', then variables $x_{\min}$ and $y_{\max}$ can be determined, as shown in equation 2.5. Subsequently, by equation 2.6, the major axis length of a cake can be determined. This length is used as a feature to recognize an ear. **Fig. 2.2** shows the white circle again. This white circle contains the major axis lengths. For easy processing, this white circle can be mapped to a straight line, as shown in the image on the right in **Fig. 2.2**. The line has a length of 128 pixels. The left figure of **Fig. 2.3** uses the symbols $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$ to represent the lengths of the major axes of different cakes. As the same situation of **Fig. 2.2**, the signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$ are mapped to a straight line. The result is shown on the right in **Fig. 2.3**.

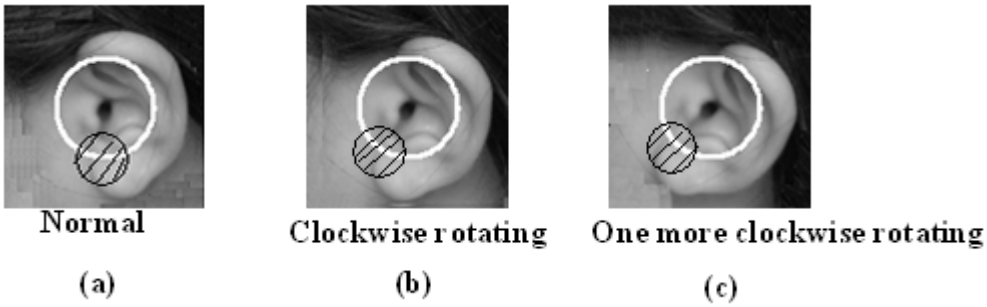

Fig. 2.1. Black slash cake areas to compute major axis length

$$m_{pq} = \iint (x - x_c)^p (y - y_c)^q b(x, y) dxdy \quad \textbf{(2.1)}$$

$$\left.\begin{aligned} x_c &= \frac{m_{10}}{m_{00}} = \frac{\iint x b(x, y) dxdy}{\iint b(x, y) dxdy} \\ y_c &= \frac{m_{10}}{m_{00}} = \frac{\iint y b(x, y) dxdy}{\iint b(x, y) dxdy} \end{aligned}\right\} \quad \textbf{(2.2)}$$

$$\left.\begin{aligned} a &= \iint (x')^2 b(x, y) dxdy \\ b &= 2\iint (x'y')^2 b(x, y) dxdy \\ c &= \iint (y')^2 b(x, y) dxdy \end{aligned}\right\} \quad \textbf{(2.3)}$$

$$\cos 2\theta = \pm \frac{a-c}{(b^2+(a-c)^2)^{1/2}} \quad \text{(2.4)}$$

$$x_{\min} = \iint (x'\cos\theta - y'\sin\theta)^2 \, dxdy$$
$$x_{\max} = \iint (x'\sin\theta - y'\cos\theta)^2 \, dxdy \quad \text{(2.5)}$$

$$Major \quad axis \quad length = (\frac{4}{\pi})^{\frac{1}{4}} [\frac{x_{\min}^3}{x_{\max}}]^{\frac{1}{8}} \quad \text{(2.6)}$$

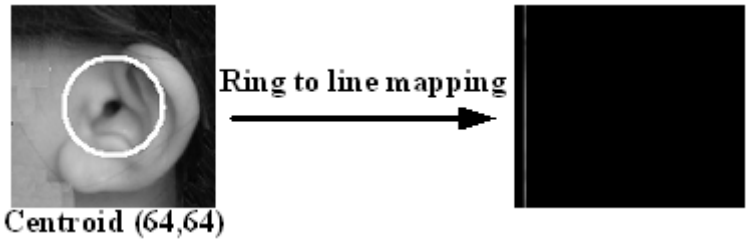

Fig. 2.2. Ring-signal to line signal: radius 28.

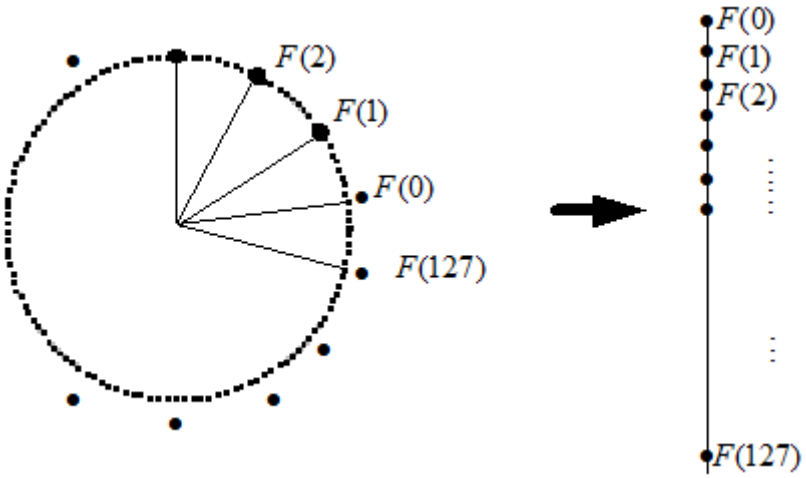

Fig. 2.3. Ring to line signals.

3. Vector computation

Fig. 3.1 shows the image extracting process. Image (a) shows the original 1224-by-1632 image, taken with the Nikon 8400 camera. The techniques used to extract the essential part of the image are as follows: Sobel operator, three by three cross medium filter, and edge searching and thinning operator. Image (b) shows the 1224-pixels-by-1224-pixels image,

which is obtained by using the aforementioned techniques. Furthermore, the ear hole of an ear is determined. Moreover, in this study, the whole ear following the ear hole is moved to the center of an ear image, as shown image (b). Image (b) is further processed to become a smaller image with 128-by-128 pixels, as show in figure (c).

Fig. 3.2 shows again the three different ears, belonging to the same person. The left ear image (a) is in the normal position; the middle ear image (b) is obtained by rotating the left ear image (a) a little bit clockwise. The right ear image (c) is obtained by rotating the middle ear image (b) a little bit further in a clockwise direction. These three ears belong to the same person and the rotation invariant transform technique is used to classify them. This rotation invariant transform technique is described in the follows.

Fig. 3.3 shows one ring in counter-clock-wise positions contain signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$. For this ring, there are 128 signals residing in it. Each two adjacent signals mapped to the circle centroid from an angle. For example, signals $F(0)$, $F(127)$ and the circle $Centroid$ can generate an angle $\angle F(0)\ Centroid\ F(127)$. There are 128 signals in this ring, thus the angle of $\angle F(0)\ Centroid\ F(127)$ is $(360/128)^{\circ}$. In **Fig. 3.3**, if signal $F(0)$ resides at $((360/128)\cdot 0)^{\circ}$ position, $F(1)$ should be at $((360/128)\cdot 1)^{\circ}$ position. Similarly, $F(2)$ should be at $((360/128)\cdot 2)^{\circ}$ position and $F(127)$ should be at $((360/128)\cdot 127)^{\circ}$ position. For efficient processing result, the circle signal in the left is mapped to the line signal on the right hand side. After these circle signals are placed in one straight line, one can determine signal $F(0)$ places on the top of the line and $F(127)$ at the bottom. These signals are the major axis lengths, as computed and obtained in the previous section. **Fig. 3.4** shows another ring, which also contain the same signals of **Fig. 3.3**, as signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$. However, these signals are shifted when rotating clockwise to different positions, with signals $F(44)$, $F(43)$, $F(42)$, and $F(41)$ on the right side of the circle. Furthermore, $F(42)$ is shifted at $((360/128)\cdot 0)^{\circ}$ angle position. These signals are also placed on a straight line, as shown in the right picture in **Fig. 3.4**, with signal $F(42)$ on the top and $F(41)$ at the bottom**.** **Fig. 3.5** shows another ring, which contains the same signals as **Figs. 3.3** and **3.4**. In this case, the signals are also rotated clockwise a little bit further, with signals $F(86)$, $F(85)$, $F(84)$, and $F(83)$ on the right hand side. In this figure $F(84)$ is on the $((360/128)\cdot 0)^{\circ}$ angle position. These signals can also be placed on one straight line, with signal $F(84)$ on the top and $F(83)$, bottom. Equation 3.1 shows a vector with signal magnitude $F(n)$ emitting from the origin of the complex plane with angle $\angle\theta^{\circ}$, which can be denoted as $F(n)e^{j\theta}$, which can be also denoted as follows.

$$F(n)e^{j\theta} = F(n)\cdot\cos\theta^{\circ} + F(n)\cdot\sin\theta^{\circ}.$$

It represents the horizontal vector adding the vertical vector. Equation 3.1 serves as the rule to generate various vectors. Equation 3.2 shows $\| Vector_0 \|$ equal to absolute value of vector $F(0)\angle 0^{\circ} + F(1)\angle 0^{\circ} + F(2)\angle 0^{\circ} + \cdots + F(127)\angle 0^{\circ}$. For short denotation, $\| Vector_0 \|$ is denoted as $\| V_0 \|$. As in equation 3.1, when $\theta = 0$, $\cos\theta^{\circ} = \cos 0^{\circ}$, $\cos 0^{\circ} = 1$ and

$\sin\theta^\circ = \sin 0^\circ$, $\sin 0^\circ = 0$; thus $F(n)e^{j0} = F(n)$. Therefore, in this case the following result can be obtained.

$$\| V_0 \| = F(0) + F(1) + F(2) + \cdots + F(127)$$

Equation 3.3 shows $\| Vector_1 \|$ equal to absolute value of vector $F(0)\angle(\phi \cdot 0)^\circ + F(1)\angle(\phi \cdot 1)^\circ + F(2)\angle(\phi \cdot 2)^\circ + \cdots + F(127)\angle(\phi \cdot 127)^\circ$, where ϕ is equal $360/128$. Equation 3.4 shows $\| Vector_2 \|$ equal to absolute value of vector $F(0)\angle(\phi \cdot 0 \cdot 2)^\circ + F(1)\angle(\phi \cdot 1 \cdot 2)^\circ + \cdots + F(127)\angle(\phi \cdot 127 \cdot 2)^\circ$. In equation 3.5, it shows again that $\| Vector_0 \|$ equal to absolute value of vector $F(0)\angle 0^\circ + F(1)\angle 0^\circ + F(2)\angle 0^\circ + \cdots + F(127)\angle 0^\circ$.

As the previous discussion, the following result is obtained.

$$\| Vector_0 \| = F(0) + F(1) + F(2) + \cdots + F(127)$$

For equations 3.5, 3.6 and 3.7, if signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$ are derived from the identical source, the computed results of $\| Vector_0 \|$ have the same value. As the case of the ears shown in left, center, and right images in **Fig. 3.2**, the ears are recognized to belong to the same person. Therefore, the signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$ are the same. Thus, the computed results of $\| Vector_0 \|$ would be the same. This methodology can be used to identify the ears. In a similar case, equations 3.8-3.10 can compute the same result of $\| Vector_1 \|$, if signals $F(0)$, $F(1)$, $F(2)$, $\cdots$, $F(127)$ come from the same source. Equations 3.11-3.13 also compute the same result. Equation 3.14 is the summarized result of equations 3.2-3.13, which is used to identify different ears. In this equation, the variable $factor$ varies from 0 to 127 and for each $factor$ case the variable m also varies from 0 to 127. The reason for variables $factor$ and m having a variance range from 0 to 127 is because the signals residing in the circles in the left images in **Figs. (3.3-3.5)** contain 128 points.

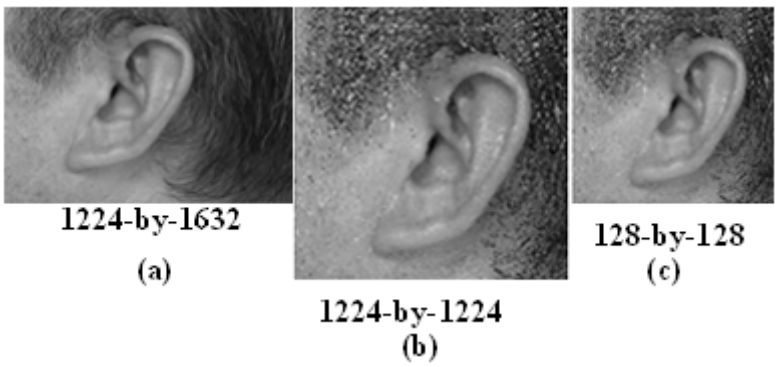

Fig. 3.1. The extracted imges.

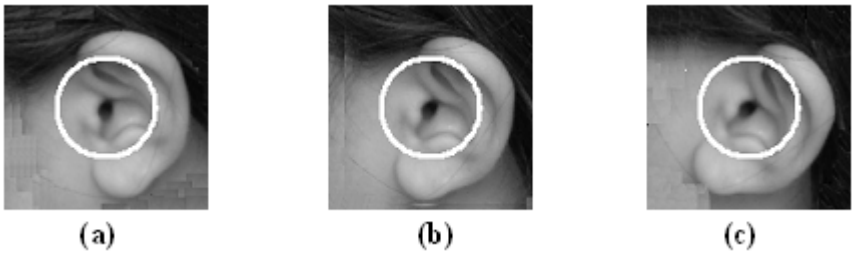

Fig. 3.2. Ear rotated to different positions.

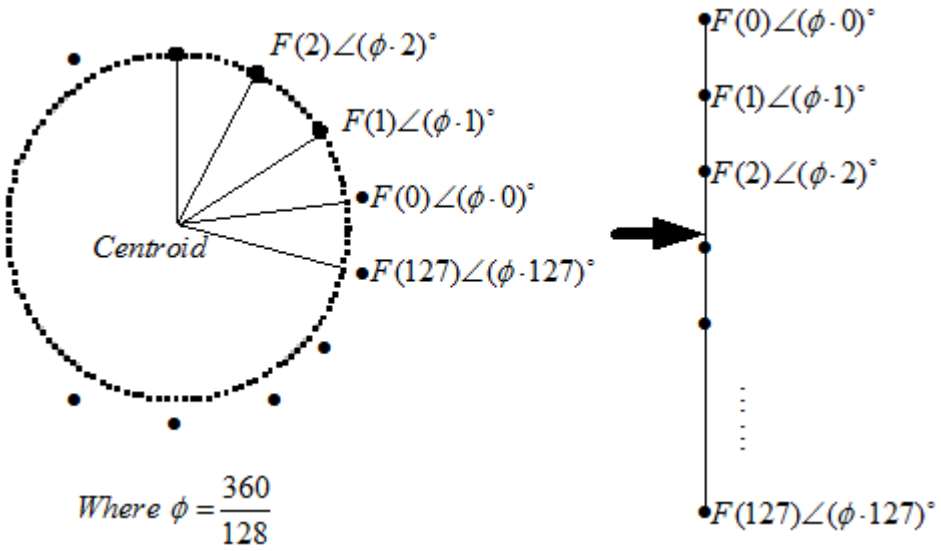

Fig. 3.3. Ring to line signals, normal position.

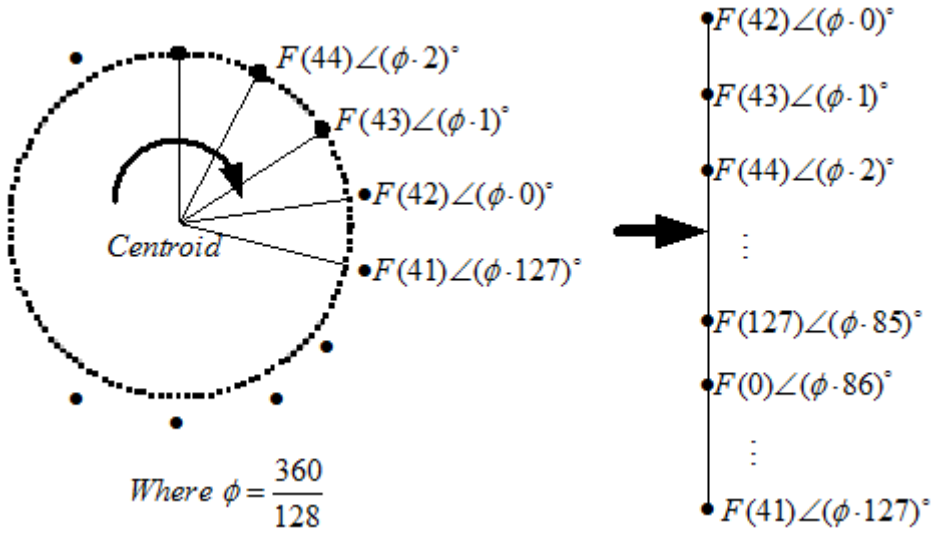

Fig. 3.4. Ring to line signals, normal position, clockwise rotated.

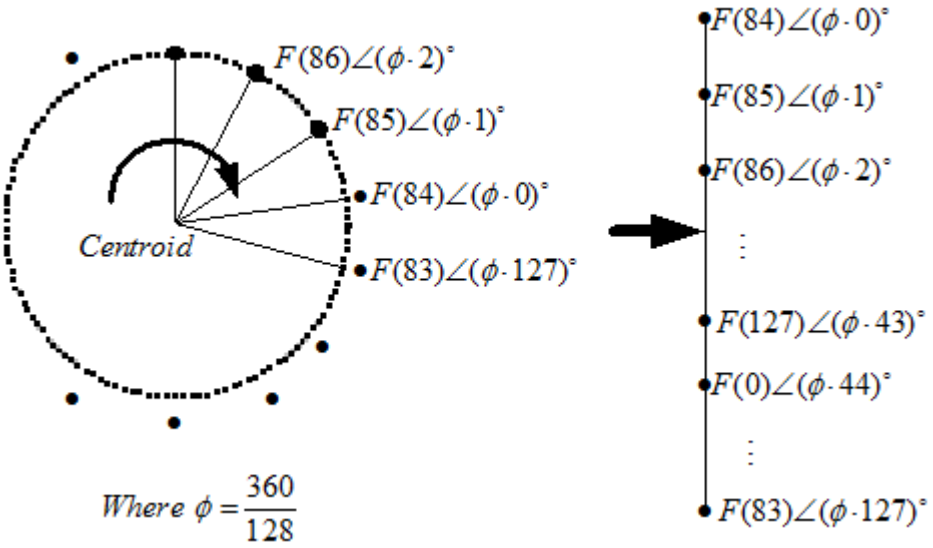

Fig. 3.5. Ring to line signals, normal position, one more clockwise rotated.

$$F(n)\angle\theta^\circ = F(n)\cdot\cos\theta^\circ + F(n)\cdot\sin\theta^\circ = F(n)_{HorizontalVector} + F(n)_{VerticalVector} \quad \textbf{(3.1)}$$

$$\|V_0\| = \|F(0)\angle 0^\circ + F(1)\angle 0^\circ + \cdots + F(127)\angle 0^\circ\| \quad \textbf{(3.2)}$$

$$\|V_1\| = \|F(0)\angle(\phi\cdot 0)^\circ + F(1)\angle(\phi\cdot 1)^\circ + \cdots + F(127)\angle(\phi\cdot 127)^\circ\| \quad \textbf{(3.3)}$$

$$\|V_2\| = \|F(0)\angle(\phi\cdot 0\cdot 2)^\circ + F(1)\angle(\phi\cdot 1\cdot 2)^\circ + \cdots + F(127)\angle(\phi\cdot 127\cdot 2)^\circ\| \quad \textbf{(3.4)}$$

Where $\|Vector_0\| = \|V_0\|$, $\|Vector_1\| = \|V_1\|$, $\|Vector_2\| = \|V_2\|$, and $\frac{360}{128} = \phi$.

General terms for figure 3.3

$$\|V_0\| = \|\sum_{n=0}^{127} F(n)\angle 0^\circ\| \quad \textbf{(3.5)}$$

(*For figure* 3.3)

$$\|V_0\| = \|\sum_{n=42}^{127} F(n)\angle 0^\circ\| + \|\sum_{n=0}^{41} F(n)\angle 0^\circ\| \quad \textbf{(3.6)}$$

(*For figure* 3.4)

$$\|V_0\| = \|\sum_{n=84}^{127} F(n)\angle 0^\circ\| + \|\sum_{n=0}^{83} F(n)\angle 0^\circ\| \quad \textbf{(3.7)}$$

(*For figure* 3.5)

For vector $\|V_0\|$

$$\| V_1 \| = \| \sum_{n=0}^{127} F(n) \angle (\phi \cdot n)^{\circ} \| \quad \text{(For figure 3.3)} \quad \textbf{(3.8)}$$

$$\| V_1 \| = \| \sum_{n=42}^{127} F(n) \angle (\phi \cdot (n-42))^{\circ} \| + \| \sum_{n=0}^{41} F(n) \angle (\phi \cdot (n+86))^{\circ} \| \quad \text{(For figure 3.4)} \quad \textbf{(3.9)}$$

$$\| V_1 \| = \| \sum_{n=84}^{127} F(n) \angle (\phi \cdot (n-84))^{\circ} \| + \| \sum_{n=0}^{83} F(n) \angle (\phi \cdot (n+44))^{\circ} \| \quad \text{(For figure 3.5)} \quad \textbf{(3.10)}$$

Where $\| \| Vector_1 \| = \| V_1 \|$ and $\frac{360}{128} = \phi.$

For vector $\| V_1 \|$

$$\| V_2 \| = \| \sum_{n=0}^{127} F(n) \angle (\phi \cdot n \cdot 2)^{\circ} \| \quad \text{(For figure 3.3)} \quad \textbf{(3.11)}$$

$$\| V_2 \| = \| \sum_{n=42}^{127} F(n) \angle (\phi \cdot (n-42) \cdot 2)^{\circ} \| + \| \sum_{n=0}^{41} F(n) \angle (\phi \cdot (n+86) \cdot 2)^{\circ} \| \quad \text{(For figure 3.4)} \quad \textbf{(3.12)}$$

$$\| V_2 \| = \| \sum_{n=84}^{127} F(n) \angle (\phi \cdot (n-84) \cdot 2)^{\circ} \| + \| \sum_{n=0}^{83} F(n) \angle (\phi \cdot (n+44) \cdot 2)^{\circ} \| \quad \text{(For figure 3.5)} \quad \textbf{(3.13)}$$

Where $\| Vector_2 \| = \| V_2 \|$ and $\frac{360}{128} = \phi.$

For vector $\| V_2 \|$

$$\| V_{factor} \| = \| \sum_{n=m}^{127} F(n) \angle (\phi \cdot (n-m) \cdot factor)^{\circ} \| + \| \sum_{n=0}^{m-1} F(n) \angle (\phi \cdot (n+(128-\mathrm{m})) \cdot factor)^{\circ} \| \quad \textbf{(3.14)}$$

where factor varies from 0 to 127 and m also varies from 0 to 127.

4. Managing image-shifting problem and comparing signals

Fig. 4.1 shows the ring signals are mapping to straight-line signals. In this white circle it contains the lengths of major axes, which are obtained by using the aforementioned technology. In this figure, the ring radius of the white ring is 28 pixels in length. The image frame of **Fig. 4.1** is a 128-pixels-by-128-pixels frame. The centroid of this 128-pixels-by-128-pixels frame is located at (64,64). The circle centroid of the white circle of **Fig. 4.1** is located at (64,64). As mentioned earlier, the major axis lengths contained in this white circle are mapped and a straight line is generated, as shown in the left side of **Fig. 4.1**. The length of the straight line is 128 pixels. This mapped line signals are used to identify different ears. If more major axis lengths are obtained, more identifiers can be generated for comparison to identify an ear. Subsequently, one can obtain more accurate identification results. To obtain more major axis lengths extra extracting circles are added to extract the major axis length. For this process, in total there are 16 circles used to extract major axis lengths, as shown in **Figs. (4.2-4.5)**. For each circle, the radius is 13, 16, 19, 22, ⋯, 58, respectively, which represent the pixel lengths.

To achieve ear recognition, the image-shifting problem must be solved. To accomplish this, all 16 circles are combined to generate a template. Furthermore, this 16-circle template is positioned at multiple locations around the ear to extract the major axis lengths. The extract signals are shown in the right hand side, represented as strips in **Figs. (4.2-4.5)**. These strips are called bands. Each ear generates its own unique bands. Each ear has a 16-circle template at different locations; this 16-circle template generates different bands. For two ears, band-to-band comparisons are made to locate the maximum match of two bands and to identify each ear. **Fig. 4.2** depicts the use of location (54,54) as the centroid to extract the major axis length. **Fig. 4.3** shows by centroid (54,74) to carry out this procedure; **Fig. 4.4**, centroid (74,74); **Fig. 4.5**, centroid (74, 54).

For the right line in **Fig. 4.1**, this straight line is processed by the rotation invariant transform or vector manipulation technique to obtain the recognition data, as discussed in the previous section. Each band is composed of 16 individual lines as shown in **Figs. (4.2-4.5)**. Each individual line is used and processed by the same procedure shown in **Fig. 4.1** to obtain recognition data. The 128-pixels-by-128-pixels frames are used for this study to obtain ear recognition-data. The centroid of this 128-pixels-by-128-pixels frame is located at (64,64) position. The ear is moving to the centroid position of this 128-pixels-by-128-pixels frame. The ear hole is located exactly at the (64,64) position. As mentioned earlier, the 16-circle template is used to extract the major axis lengths. For managing this shifting-problem, the locations (64,64), (63,63), (63,65), (65,65), (65,63), (62,64), (64,66), (66,64), and (64,62) are used as the centroids to extract the major axis lengths, as shown in **Fig. 4.6**. In total, there are nine different locations to serve as the centroids to extract the major axis lengths. The extracted results are similar to other results shown on the right images in **Figs. (4.2-4.5)**. For more efficient processing result, these nine bands are put together to generate a larger band, as shown in **Fig. 4.7**. The corresponding image results are shown in the bottom of **Fig. 4.8**. In **Fig. 4.6** for each centroid, the extract strips are identified as strip0, strip1… strip8, respectively. For **Fig. 4.7**, the band extracted from centroid at location (64,64) is saved at positions (0) 0-15. The number 0 represents the 0th band and 0-15 represents the 0th column to 15th columns for the combined larger band. The previous nine-centroid extracted signals are saved at positions (0) 0-15, (1) 16-31, (2) 32-47, ⋯, (8) 128-143, respectively, as shown in **Fig. 4.7**. After this step, the rotation invariant transform technique is applied to extracted signals to obtain the comparison data, as shown at the bottom of **Fig. 4.8**. The left figure of **Fig. 4.9** shows the symbol of strip 0 of ear one; the right figure shows all of the strips. This

figure shows strip 0 of ear one comparing with all of the strips in ear two. By this comparison, the maximum match can be obtained. By this method, the algorithm can judge whether ear one and ear two is coming from the same source.

To determine the maximum match for recognition for the two images, each strip of one ear is compared with the entire nine strip of another ear. For example, strip 0 of ear one is compared to strips 0, 1, 2, ⋯, and 8 of ear two; strip 1 of ear one is compared with strips 0, 1, 2, ⋯, and 8 of ear two; strip 8 of object one is compared to strips 0, 1, 2, ⋯, and 8 of ear two, as shown in **Fig. 4.10**. After completing these steps, the minimum-subtracted-result, which representing the maximum matching, can be obtained. Finally, they are recorded in a table. The actual images of subtracted-results are shown in **Fig. 4.11**.

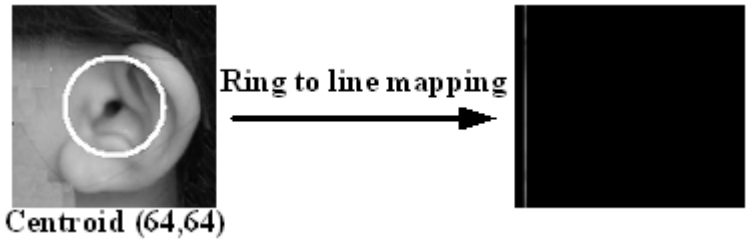

Fig. 4.1. Ring-signal to line signal: radius 28.

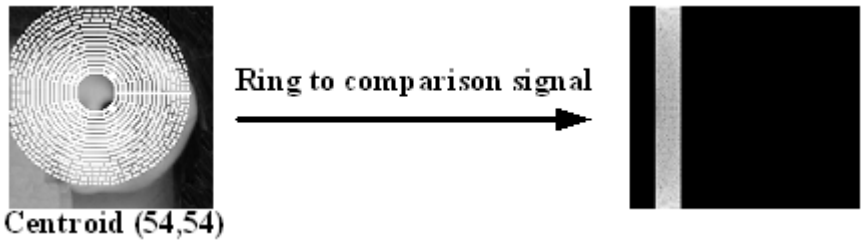

Fig. 4.2. Location (54,54) as centroid to generate comparison signals.

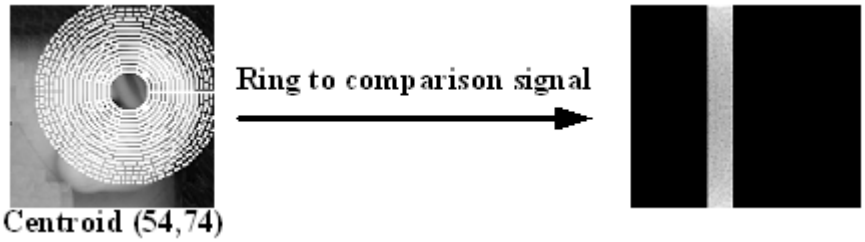

Fig. 4.3. Location (54,74) as centroid to generate comparison signals.

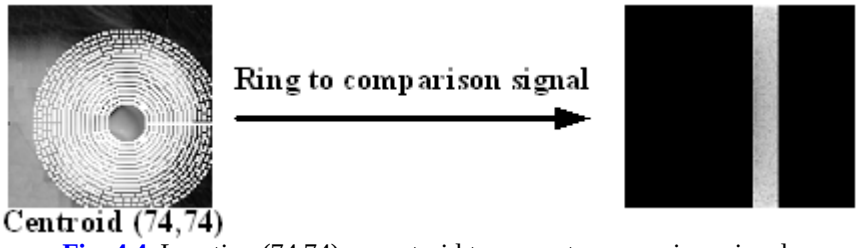

Fig. 4.4. Location (74,74) as centroid to generate comparison signals.

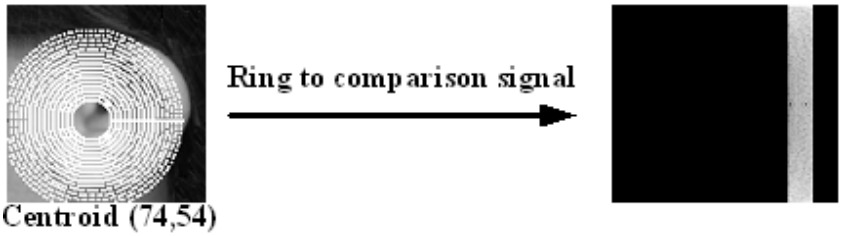

Fig. 4.5. Location (74,54) as centroid to generate comparison signals.

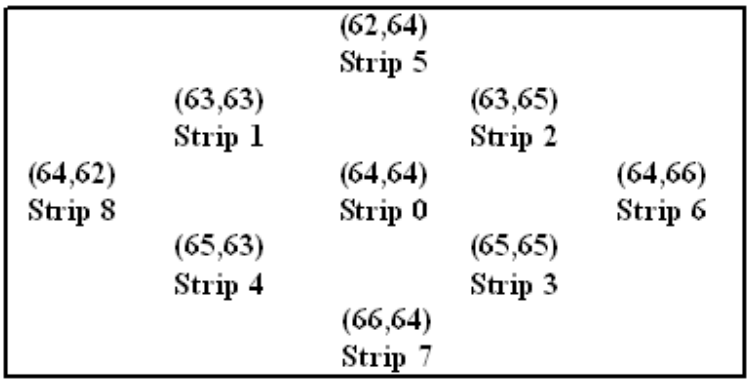

Fig. 4.6. Centroid locations to generate various strips.

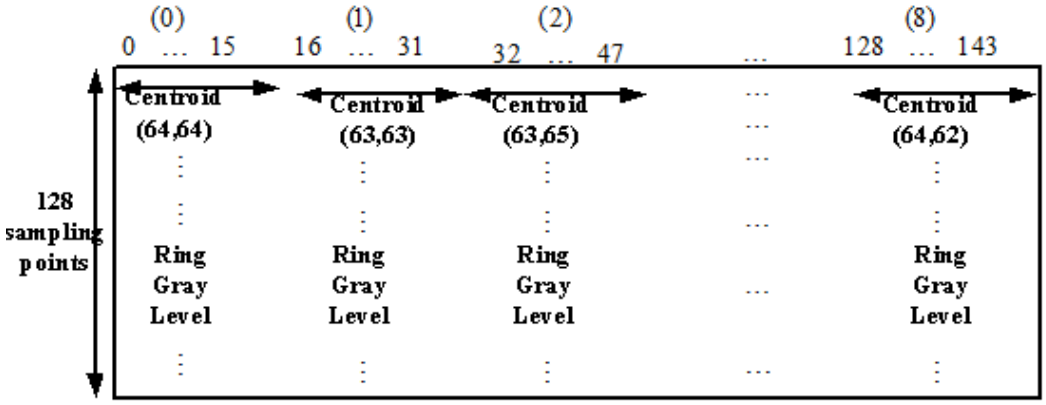

Fig. 4.7. Various centroids generate various vertical bands.

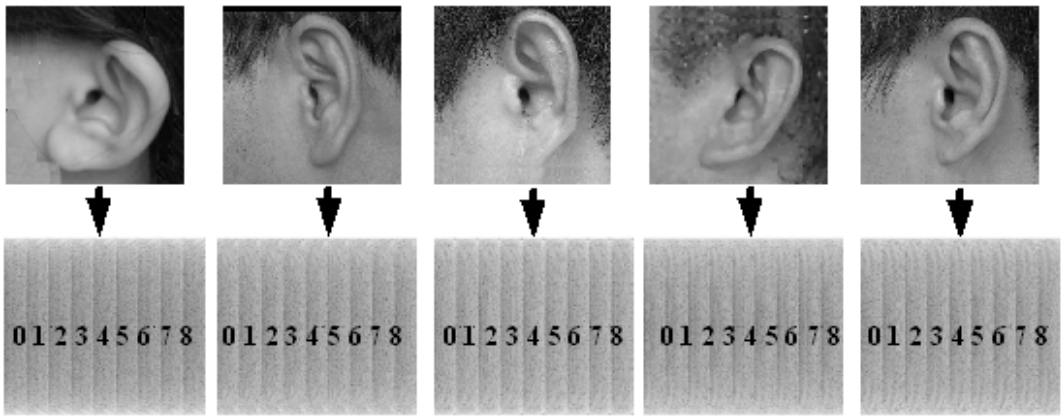

Fig. 4.8. Various centroids generate various vertical comparison strips.

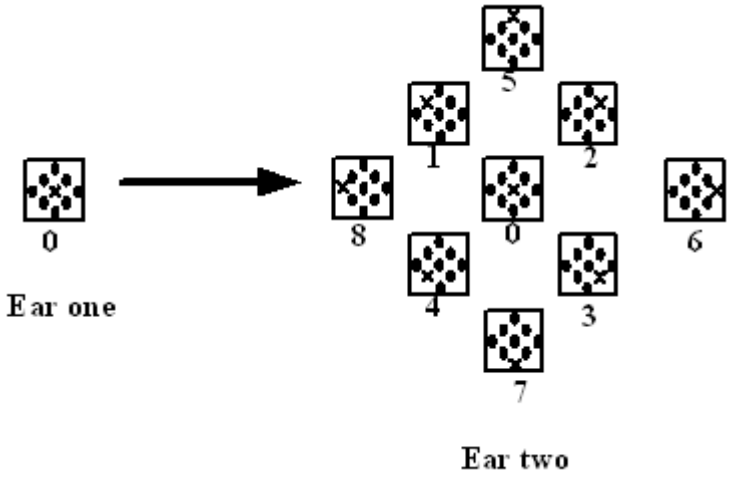

Fig. 4.9. Strop 0 of ear one comparing with strips 0, 1, 2,… , 8 of ear two.

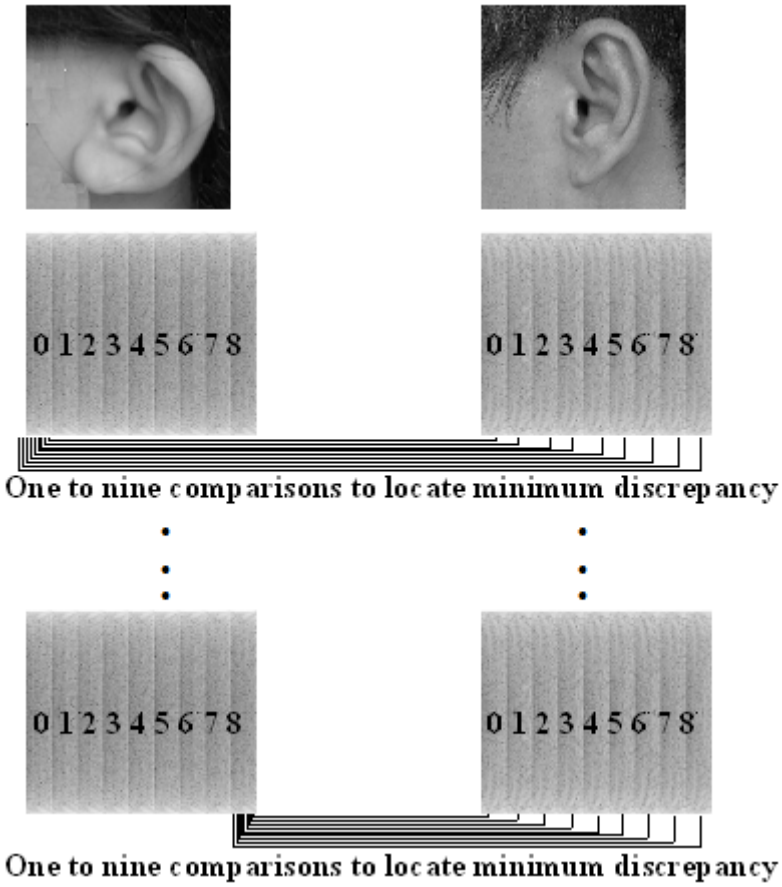

Fig. 4.10. Vertical strip subtractions.

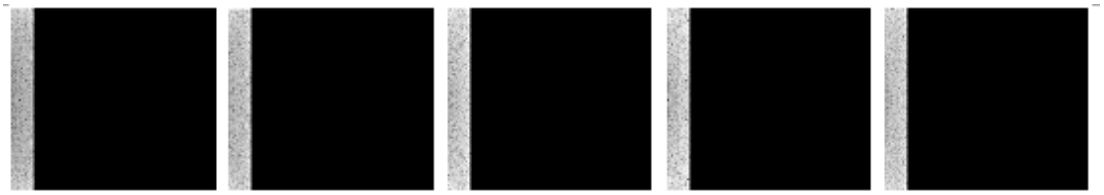

Fig. 4.11. Vertical comparison results.

5. Results

Fig. 5.1 shows the original ear images. In this study, three ear-images were taken for each person and 40 people participated in this test. Therefore, there are a total of 120 ear-images available for comparison. In this study, the 15 ear-images which belonging to five different persons were designated as a group and these 15 ear-images are tested to verify the accuracy of the system. Five persons are assigned together as a group and there are 40 persons joining in this study; therefore, totally there are eight different groups. As previously mentioned, three different ear-images are taken for one person. Five persons are assigned in one group; therefore, there are 15 ear-images for test per group. One ear-image is compared with the other 14 ear-images. In total, there are 105 comparisons in one group, wherein one ear was compared with 14 others. Of the 105 comparisons conducted altogether, 15 were run for genuine comparisons, and the other 90 being imposters. In table 5.1, the data inside the bold rectangular boxes represent comparisons of identical sources; the others, represent different sources. Table 5.2 and **Fig. 5.2** summarize the results listed in table 5.1 for eight different groups encompassing 120 genuine and 720 imposter tests. In table 5.2, the subtracted values of imposter comparisons are 21198155~10977696; whereas, identical, 8709516~4338999; the tolerance gap between the two is 2268180; the error rates for both tests are zeros.

Fig. 5.1. Original 128-by-128 images.

Table 5.1. Comparison data of original images.

4761459	5992151	16813061	15763595	17615733	18038174	18405252	18338316	15315822	15017388	15389453	17108676	17670210	18489724
	5007977	17007179	15715624	17600403	20119960	20211563	20229597	14925616	14581296	15072131	16860539	17943265	18933808
		16893492	16064645	17637636	19313545	19370533	19430543	15196398	15072823	15485935	17164847	17879179	18976910
			4338999	5849972	20956475	21141297	21917007	12932600	12538691	12364234	10977696	11164622	12319893
				6194770	20914581	21198155	21891745	12717075	12339666	12199495	10812079	11204967	12289427
					20333801	20849351	21310188	13898467	13577968	13089961	11144800	11427573	12225167
						8709516	8703878	21685888	21385373	21564802	19912972	19658002	20045606
							7938753	21798751	21541180	21433354	20478014	20614826	20181336
								21854089	21567335	21695652	20627208	20633657	20585874
									5723653	5525073	12149953	13265395	13951084
										5445762	12161588	13086700	13930077
											11914435	12955484	13621990
												6192490	4941719
													7118037

Comparisons of identical sources

Table 5.2. Range of subtraction values.

	Range of subtraction values	Gap tolerance between genuine and imposter hands	Errors (Times)	Error rate
Imposter comparisons	21198155 ~ 10977696	2268180	0	840:0
Identical comparisons	8709516 ~ 4338999	2268180	0	840:0

120 different ears; 840 comparisons; 9838556 as threshold.

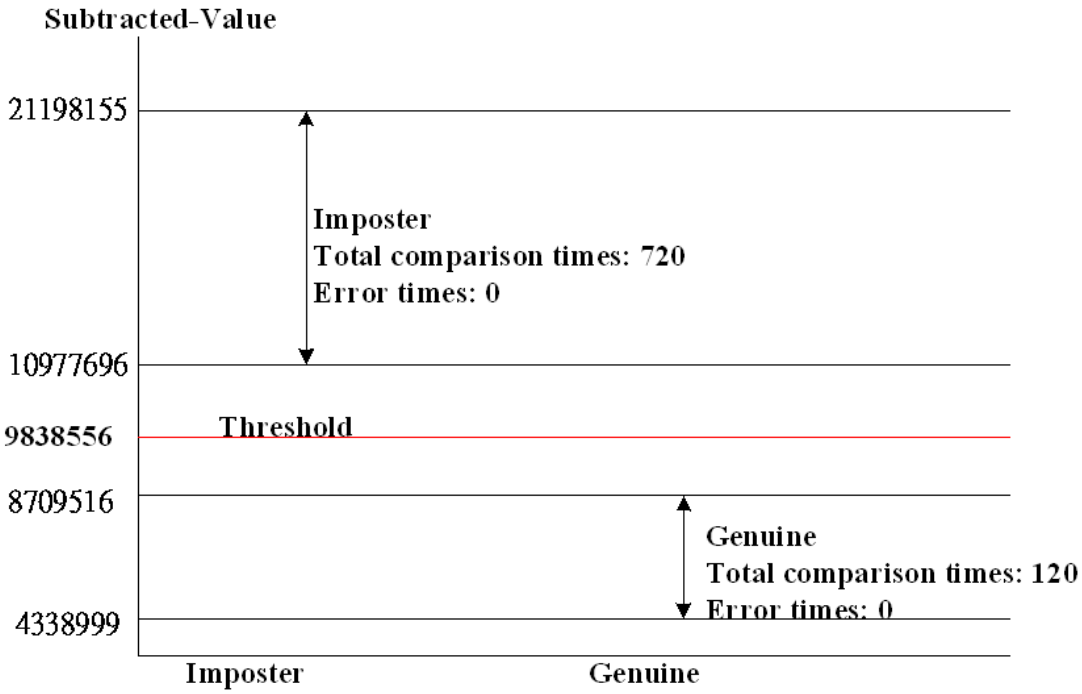

Fig. 5.2. Original Response characteristics of genuine and imposter comparisons: 120 different ears; 40 comparisons; 9838556 as threshold.

6. Conclusions

In the previous technique, outlines of ears are located and these are used to determine the heights and widths of the ears, as shown in **Fig. 6.1**. However, the ears may be hidden behind the hair. Additionally, the ear and the skin behind the ear have the same skin color. Therefore, there is difficulty to distinguish between the ear-line and the adjacent skin behind the ear location. Thus, these types of interferences may cause difficulty to obtain an image of the full ear outline. This study does not extract the outer ear outline. Instead, the ear hole is used as a centroid and a 16-circle template are used to extract the major axis lengths of the ear. The 16-circle template extracted signals are used for ear recognition. Alternatively, as a

comparison, as shown in **Fig. 6.2**, the 3D ear extraction technique is a good method to recognize an ear. The 3D-ear contains ample object depth information. Furthermore, it is easier to obtain this depth information. **Fig. 6.3** shows one can obtain the position of highest tip of the ear edge. The vectors from ear holes to these highest tips can be established. Furthermore, the positions of the inner lowest valleys of the ears and the left isolated ear balls can be located as well. The vectors from the inner lowest valleys of the ears to the left isolated ear balls can be established. One can use these vector features to recognize an ear.

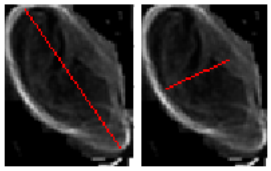

Fig. 6.1. Ear height and width.

Fig. 6.2. 3 D ears.

Fig. 6.3. Image depth to establish SIFT vectors.

Acknowledgement

The Ministry of Science and Technology, Taiwan, supported this work under grant MOST 105-2622-E-212-007 -CC3.

References

[1] Banafshe Arbab-Zavar and Mark S. Nixon, "On guided model-based analysis for ear biometrics," *Computer Vision and Image Understanding*, 115, 487–502, 2011. Article (CrossRef Link).

[2] Septimiu Crisan, Ioan Gavril Tarnovan, and Titus Eduard Crisan, "Radiation optimization and image processing algorithms in the identification of hand vein patterns," *Computer Standards & Interfaces*, 32, 130–140, 2010. Article (CrossRef Link).

[3] R. M. Farouk, "Iris recognition based on elastic graph matching and Gabor wavelets," *Computer Vision and Image Understanding*, 115, 1239–1244, 2011. Article (CrossRef Link).

[4] Zhen-Hua Feng, Guosheng Hu; Kittler, J., Christmas, W., Xiao-Jun Wu, "Cascaded Collaborative Regression for Robust Facial Landmark Detection," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3425 – 3440, 2015. Article (CrossRef Link).

[5] Miguel A. Ferrer and Aythami Morales, "Hand-Shape Biometrics Combining the Visible and Short-Wave Infrared Bands," *IEEE transactions on information forensics and security*, vol. 6, no. 4, December. 2011. Article (CrossRef Link).

[6] Cong Geng and Xudong Jiang, "Face Recognition Based on the Multi-Scale Local Image Structures," *Pattern Recognition*, 44, 2565–2575, 2011. Article (CrossRef Link).

[7] S.G. Kong, R.O. Mbouna, "Head Pose Estimation From a 2D Face Image Using 3D Face Morphing With Depth Parameters," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1801 – 1808, 2015. Article (CrossRef Link).
[8] Ajay Kumar and Chenye Wu, "Automated Human Identification Using Ear Imaging," *Pattern Recognition*, 45, 956–968, 2012. Article (CrossRef Link).
[9] Haifeng Hu, "Multiscale illumination normalization for face recognition using dual-tree complex wavelet transform in logarithm domain," *Computer vision and image understanding*, 115, 1384–1394, 2011. Article (CrossRef Link).
[10] Deng-Yuan Huang, Wu-Chih Hu, and Sung-Hsiang Chang, "Gabor filter-based hand-pose angle estimation for hand gesture recognition under varying illumination," *Expert Systems with Applications*, 38, 6031–6042, 2011. Article (CrossRef Link).
[11] Hong Huang, JiaminLiu, Hailiang Feng, and Tongdi He, "Ear Recognition Based on Uncorrelated Local Fisher Discriminant Analysis," *Neurocomputing*, 74, 3103–3113, 2011. Article (CrossRef Link).
[12] Vivek Kanhangad, Ajay Kumar, and David Zhang, "A Unified Framework for Contactless Hand Verification," IEEE transactions on information forensics and security, vol. 6, no. 3, September 2011. Article (CrossRef Link).
[13] S.G. Kong, R.O. Mbouna, "Head Pose Estimation From a 2D Face Image Using 3D Face Morphing With Depth Parameters," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1801 – 1808, 2015. Article (CrossRef Link).
[14] Antonio Iula and Michele De Santis, "Experimental evaluation of an ultrasound technique for the biometric recognition of human hand anatomic elements," *Ultrasonics*, 51, 683–688, 2011. Article (CrossRef Link).
[15] Leila Mansourian, Muhamad Taufik Abdullah, Lilli Nurliyana Abdullah, Azreen Azman, and Mas Rina Mustaffa, "A Salient Based Bag of Visual Word Model (SBBoVW): Improvements toward Difficult Object Recognition and Object Location in Image Retrieval," *KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS*, VOL. 10, NO. 2, pp. 769 - 786, Feb. 2016. Article (CrossRef Link).
[16] D. Muramatsu, A. Shiraishi, Y. Makihara, M.Z. Uddin, Y. Yagi, "Gait-Based Person Recognition Using Arbitrary View Transformation Model," *IEEE Transactions on Image Processing*, vol. 24, no. 1, pp. 140 – 154, 2015. Article (CrossRef Link).
[17] Lin Zhang, Ying Shen, Hongyu Li, Jianwei Lu, "3D Palmprint Identification Using Block-Wise - Features and Collaborative Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, 1730–1736, 2015. Article (CrossRef Link).
[18] Yan Zhou, Qingwu Li, and Guanying Huo, "Human Visual System based Automatic Underwater Image Enhancement in NSCT domain," *KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS*, VOL. 10, NO. 2, pp. 837 - 856, Feb. 2016. Article (CrossRef Link).

Dr. Ching-Liang Su was a visiting professor at Electrical Engineering Department, Stanford University in 2010 and 2014 respectively. He was also a visiting scholar at Electrical Engineering Department, University of Washington in 2011. Currently he is a professor at Department of Industrial Engineering and Management, Da Yeh University, in Taiwan. In 1995 and 1993 he received his Ph.D. and M.S. degrees in computer engineering from the University of Louisiana at Lafayette, respectively. In 1989 he also received another M.S. degree in electrical computer engineering from the University of Louisville at Kentucky. His research interest is image processing and pattern recognition.