Chapter 6

HUMAN EAR DETECTION FROM 3D SIDE FACE RANGE IMAGES

H. Chen and B. Bhanu
Center for Research in Intelligent Systems, University of California, Riverside, CA 92521, USA, {hui, bhanu}@cris.ucr.edu

Abstract: Ear is a new class of relatively stable biometrics which is not affected by facial expressions, cosmetics and eye glasses. To use ear biometrics for human identification, ear detection is the first part of an ear recognition system. In this chapter we propose two approaches for locating human ears in side face range images: (a) template matching based ear detection and (b) ear shape model based detection. For the first approach, the model template is represented by an averaged histogram of shape index that can be computed from principal curvatures. The ear detection is a four-step process: step edge detection and thresholding, image dilation, connect-component labeling and template matching. For the second approach, the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. Given a side face range image, step edges are extracted and then the edge segments are dilated, thinned and grouped into different clusters which are the potential regions containing an ear. For each cluster, we register the ear shape model with the edges. The region with the minimum mean registration error is declared as the detected ear region; during this process the ear helix and anti-helix parts are identified. Experiments are performed with a large number of real side face range images to demonstrate the effectiveness of the proposed approaches.

Key words: ear biometrics; ear detection; range images; shape model; shape index.

1. INTRODUCTION

Ear is a viable new class of biometrics since ears have desirable properties such as universality, uniqueness and permanence[1, 2]. The ear has certain

A. Koschan et al. (eds.), 3D Imaging for Safety and Security, 133–155.

advantages over other biometrics. For example, ear is rich in features; it is a stable structure which does not change with the age. It does not change its shape with facial expressions. Furthermore, the ear is larger in size compared to fingerprints and can be easily captured although sometimes it can be hidden with hair and earrings. Although it has certain advantages over other biometrics, it has received little attention compared to other popular biometrics such as face, fingerprint and gait[3,4,5,6,7,8].

The current research has used intensity images and, therefore, the performance of the recognition systems is greatly affected by imaging problems such as lighting and shadows[3,4,5]. Range sensors that are insensitive to above imaging problems can directly provide us 3D geometric information[9,10]. Therefore it is desirable to design a human ear recognition system from 3D side face range images obtained at a distance. Human ear detection is the first task of a human ear recognition system and its performance significantly affects the overall quality of the system.

In this chapter, we propose two techniques for locating human ears in side face range images: template matching based ear detection[11] and ear shape model based detection[12]. The first approach has two stages: off-line model template building and on-line ear detection. Ear can be thought of as a rigid object with much concave and convex areas. As compared to other approaches for object detection in range images[13,14,15,16,17], we use the averaged histogram of shape index to represent the ear model template since shape index is a quantitative measure of the shape of a surface[20,21]. During the on-line detection, we first perform the step edge computation and thresholding since there is a sharp step edge around the ear boundary, and then we do image dilation and connected-component analysis to find the potential regions containing an ear. Next for every potential region, we grow the region and compute the dissimilarity between each region's histogram of shape indexes and the model template. Finally among all of the regions, we choose the one with the minimum dissimilarity as the detected region that contains ear.

For the second approach, the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. Since the two curves formed by ear helix and anti-helix parts are similar for different people, we do not take into account the small deformation of two curves between different persons, which greatly simplifies our ear shape model. Given side face range images, step edges are extracted; then the edge segments are dilated, thinned and grouped into different clusters which are the potential regions containing an ear. For each cluster, we register the ear shape model with the edges. The region with the minimum mean registration error is declared as the detected ear region; the ear helix and anti-helix parts are identified in this process.

The rest of chapter is organized as follows. Section 2 introduces the related work, motivation and contributions. Section 3 describes our first approach to detect ears using template matching based method. Section 4 describes our second approach to build the ear shape model and detect human ears in side face range images. Section 5 gives the experimental results to demonstrate the effectiveness of two approaches. Finally, Section 6 provides the conclusions.

2. RELATED WORK, MOTIVATION AND CONTRIBUTIONS OF THE CHAPTER

2.1 Related work

There are only a few papers dealing with object detection from range images. In the following we give a brief review of object detection techniques from range images.

Keller et al.[13] introduced a fuzzy logic system for automatic target detection from LADAR images. They used two fuzzy logic detection filters and one statistical filter to create pixel-based target confidence values which are fused by the fuzzy integral to generate potential target windows. Features extracted from these windows are fed to a neural network post-processor to make a final decision.

Meier and Ade[14] proposed an approach to separate image features into ground and road obstacles by assuming the road was flat. They distinguished obstacles and road pixels using the separating plane. The plane model is updated by fitting a plane through all pixels marked as ground. Connected component analysis is used to partition detected obstacles into different objects.

Sparbert et. al.[15] presented a method to detect lanes and classify street types from range images. First they calculated the lane's width, curvature and relative position to the car, then compared them with a prior knowledge on construction rules of different street types, and finally achieved street type based on the mean value of lane's width.

Garcia et. al.[16] generated a unique signature of a 3D object by the Fourier transform of the phase-encoded range image at each specific rotation. The signature defined in a unit sphere permitted the detection of 3D objects by correlation techniques.

Heisele and Ritter[17] proposed a method for segmenting temporal sequences of range and intensity images. The fusion of range and intensity data for segmentation is solved by clustering 4D intensity/position features.

Kalman filters are then used to stabilize tracking by predicting dynamic changes in cluster positions.

Boehnen and Russ[18] proposed an algorithm to utilize the combination of the range and registered color images for automatically identifying facial features. The foreground is first segmented using the range data and then the skin color detection is used to identify potential skin pixels, which are further refined using z-based range erosion to compute the eye and mouth maps. Finally, the geometric-based confidence of candidates is computed to aid in the selection of best feature set.

Tsalakanidou et. al.[19] introduced a face localization procedure combining both depth and color data. First the human body is easily separated from the background by using the depth information and its geometric structure; and then the head and torso can be identified by modeling the 3D point distribution as a mixture model of two Gaussians; and finally the position of the face is further refined using the color images by exploiting the symmetric structure of the face.

2.2 Motivation

The anatomical structure of the ear is shown in Figure *6-1*. The ear is made up of standard features like the face. These include the outer rim (helix) and ridges (anti-helix) parallel to the helix, the lobe and the concha that is a hollow part of ear. From Figure *6-1*, we clearly see that two curves formed by ear helix and anti-helix parts are easily identified. We can use these two curves to develop the procedures to locate the ear in side face range images.

2.3 Contributions of the Chapter

The contributions of the chapter are: (a) Based on the geometric characteristics, we develop a template matching based approach for ear detection in side face range images. (b) We also propose a ear shape model based approach for locating 3D ears more accurately in side face range images. (c) We present many examples to illustrate the performance of these two approaches.

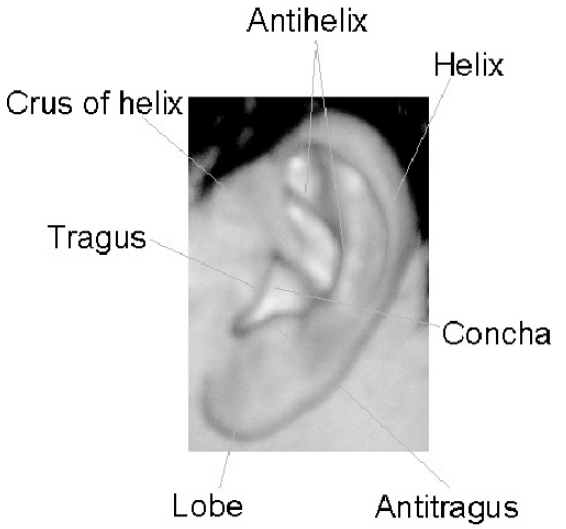

Figure 6-1. The external ear and its anatomical parts.

3. TEMPLATE MATCHING BASED EAR DETECTION

3.1 Off-line model template building

3.1.1 Shape Index

Shape index S_i, a quantitative measure of the shape of a surface at a point p, is defined by Eq. (1).

$$S_i = \frac{1}{2} - \frac{1}{\pi} \tan^{-1} \frac{k_1(p) + k_2(p)}{k_1(p) - k_2(p)} \tag{1}$$

where $k_1(p)$ and $k_2(p)$ are maximum and minimum principal curvatures respectively[20]. With this definition, all shapes can be mapped into the interval $S_i \in [0,1]$. The shape categories and corresponding shape index range are listed in Table *6-1*[20,21]. From Table *6-1*, we can see that larger shape index values represent convex surfaces and smaller shape index values represent concave surfaces.

Table 6-1. Surface shape categories and the range of shape index values.

Shape Category	S_i Range
Spherical cup	[0, 1/16)
Trough	[1/16, 3/16)
Rut	[3/16, 5/16)
Saddle rut	[5/16, 7/16)
Saddle	[7/16, 9/16)
Saddle ridge	[9/16, 11/16)
Ridge	[11/16, 13/16)
Dome	[13/16, 15/16)
Spherical cap	[15/16, 1]

Ear has significant convex and concave areas, which gives us a hint to use the shape index for ear detection. The original ear range image and its shape index image are shown in Figure *6-2*. In Figure *6-2*(b), the brighter points denote large shape index values that correspond to ridge and dome surfaces. We believe the ridge and valley areas form a pattern for ear detection. We use the distribution of shape index as a robust and compact descriptor since 2D shape index image is much too detailed. The histogram h can be calculated by $h(k) = \#$ of points with shape index $\in bin(k)$. The histogram is normalized during the implementation.

3.1.2 Curvature Estimation

In order to estimate curvatures, we fit a quadratic surface $f(x, y) = ax^2 + by^2 + cxy + dx + ey + f$ to a 5×5 window centered at the surface point of interest and use the least squa re method to estimate the parameters of the quadratic surface. After we get the parameters, we use differential geometry to calculate the surface normal, Gaussian and mean curvatures and principal curvatures[6,22].

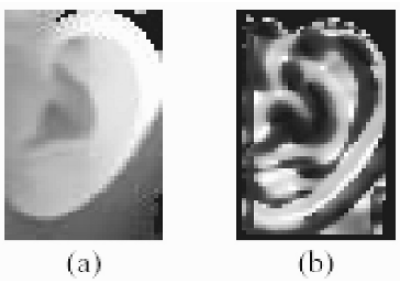

Figure 6-2. (a) Ear range image. Darker pixels are away from the camera and the lighter ones are closer. (b) Its shape index image. Darker pixels correspond to concave surfaces and lighter ones correspond to convex surfaces.

3.1.3 Building Model Template

Given a set of training side face range images, first we extract ears manually, and then calculate its shape index image and histogram the shape index image. After we get the histograms for each training image, we average the histograms and use the averaged histogram as our model template. Figure *6-3* shows the model template, obtained using 20 training images, in which the two peaks correspond to the convex and concave regions of the ear.

3.2 Step Edge Detection, Thresholding and Dilation

The step edge magnitude, denoted by M_{step}, is calculated as the maximum distance in depth between the center point and its neighbors in a small window[23]. M_{step} can be written as Eq. (2):

$$M_{step}(i,j) = \max |z(i,j) - z(i+k, j+l)|, \quad where \quad -(w-1)/2 \le k,l \le (w-1)/2 \tag{2}$$

In Eq. (2) w is the width of the window and $z(i,j)$ is the z coordinate of the point (i,j). To get the step edge magnitude image, a $w \times w$ window is translated over the original side face range image and the maximum distance calculated from Eq. (2) replaces the pixel value of the pixel covered by the center of the window. The original side face range image and its step edge magnitude image are shown in Figure *6-4* (a) and (b).

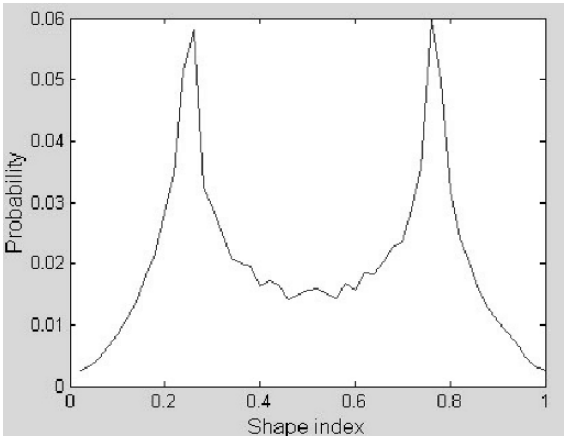

Figure 6-3. Model template (discretized into 50 bins).

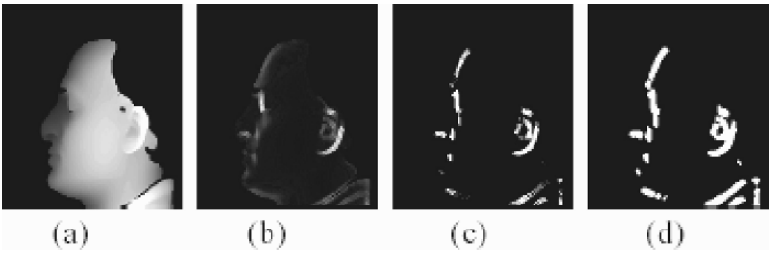

Figure 6-4. (a) Original side face range image. (b) Its step edge magnitude image. (c) Thresholded binary image. (d) Dilated image.

From Figure *6-4*(b), we clearly see that there is a sharp step edge around the ear boundary since brighter points denote large step edge magnitude. The step edge image is thresholded to get a binary image that is shown in Figure *6-4*(c). The threshold is set based on the maximum of M_{step}. Therefore, we can get a binary image by using Eq. (3). There are some holes in the thresholded binary image and we want to get the potential regions containing ears. We dilate the binary image to fill the holes using a 3×3 structure element. The dilated image is shown in Figure *6-4*(d).

$$F_T(i,j) = \begin{cases} 1 & \text{if} \quad M_{step}(i,j) \geq \eta \, Max\{M_{step}\} \\ & \quad 0 \leq \eta \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

3.3 Connected Component Labeling

Using the above result, we proceed to determine which regions can possibly contain human ears. To do so, we need to determine the number of potential regions in the image. By running connected component labeling algorithm, we can determine the number of regions. We used an 8-connected neighborhood to label a pixel. We remove smaller components whose areas are less than since the ear region is not small. The labeling result is shown in Figure *6-5*(a) and the result after removing smaller components is shown in Figure *6-5*(b).

After we get regions, we need to know the geometric properties such as the position and orientation. The position of a region may be defined using the center of the region. The center of area in binary images is the same as the center of the mass and it is computed by Eq. (4)

Figure 6-5. (a) Labeled image (12 components). (b) Labeled image after removing smaller components (6 components). (See also Plate 22 in the Colour Plate Section)

$$\bar{x} = \frac{1}{A}\sum_{i=1}^{n}\sum_{j=1}^{m} jB[i,j] \qquad \bar{y} = \frac{1}{A}\sum_{i=1}^{n}\sum_{j=1}^{m} iB[i,j] \tag{4}$$

where B is $n \times m$ matrix representation of the region and A is the size of the region.

For the orientation, we find the axis of elongation of the region. Along this axis the moment of the inertia will be the minimum. The axis is computed by finding the line for which the sum of the squared distances between region points and the line is minimum. The angle of θ is given by Eq. (5).

$$\theta = \frac{1}{2}\tan^{-1}\frac{b}{a-c} \tag{5}$$

The parameters a, b and c are given by Eq. (6), Eq. (7) and Eq. (8) respectively.

$$a = \sum_{i=1}^{n}\sum_{j=1}^{m} (x_{ij}^{'})^2 B[i,j] \tag{6}$$

$$b = 2\sum_{i=1}^{n}\sum_{j=1}^{m} x_{ij}^{'} y_{ij}^{'} B[i,j] \tag{7}$$

$$c = \sum_{i=1}^{n}\sum_{j=1}^{m} (y_{ij}^{'})^2 B[i,j] \tag{8}$$

where $x' = x - \bar{x}$ and $y' = y - \bar{y}$. θ gives us the hint about the direction along which region growing must take place.

3.4 Template Matching

As mentioned in Section 3.1.1, the model template is represented by an averaged histogram of shape index. Since histogram can be thought of as an approximation of probability density function, it is natural to use the $\chi^2 - divergence$ function Eq. (9)[24].

$$\chi^2(Q,V) = \sum_1 \frac{(q_i - v_i)^2}{q_i + v_i} \tag{9}$$

where Q and V are normalized histograms. From Eq. (9), we know the dissimilarity is between 0 and 2. If the two histograms are exactly the same, the dissimilarity will be zero. If the two histograms do not overlap with each other, it will achieve the maximum value 2.

From Section 3.3, we get the potential regions which may contain the ears. For each region, we find a minimum rectangular box to include the region, then we grow this region based on the angle θ. If $0 \le \theta \le \pi/2$, we grow the rectangle by moving the top-right vertex right, up and anti-diagonal and moving the bottom-left vertex left, down and anti-diagonal. If $\pi/2 \le \theta \le \pi$, we grow the rectangle by moving the top-left vertex left, up and diagonal and moving the bottom-right vertex right, down and diagonal. For every region, we choose the grown rectangular box with the minimum dissimilarity as the candidate ear region. Finally over all of the candidate regions, we select the one with the minimum dissimilarity as the detected region. We set a threshold γ for region growing, which controls the size of the region.

4. SHAPE MODEL BASED EAR DETECTION

4.1 Shape Model Building

Considering the fact that the curves formed by ear helix and anti-helix parts are similar for different people, we construct the ear shape model from one person only. We plan to work on building a generic ear model from multi persons. We extract ear helix and anti-helix parts by running a step edge detector with different thresholds and choose the best edge extraction result which detects ear helix and anti-helix parts. We define the ear shape

model s as 3D coordinates $\{x, y, z\}$ of n vertices that lie on the ear helix and anti-helix parts. The shape mode s is represented by a $3n \times 1$ vector $\{x_1, y_1, z_1, x_2, y_2, z_2, \ldots, x_n, y_n, z_n\}$. Figure *6-6* shows the 3D side face range image with textured appearance, in which the ear shape model s marked by yellow vertices is overlaid.

4.2 Step Edge Detection and Thresholding

Given the step face range image, the step edge magnitude, denoted by M_{step}, can be calculated as described in Section 3.2. One example of step edge magnitude image is shown in Figure *6-7*(b). In Figure *6-7*(b), larger magnitudes are displayed as brighter pixels. We can clearly see that most of the step edge magnitudes are small values. To get edges, the step edge magnitude image can be segmented using a threshold operator. The selection of threshold value is based on the cumulative histogram of the step edge magnitude image which is different from Section 3.2. Since we are interested in larger magnitudes, in our approach the top α ($\alpha = 3.5\%$) pixels with the largest magnitudes are selected as edge points. We can easily determine the threshold by investigating the cumulative histogram. The thresholded binary image is shown in Figure *6-7*(c), while the original side face range image is shown in Figure *6-7*(a).

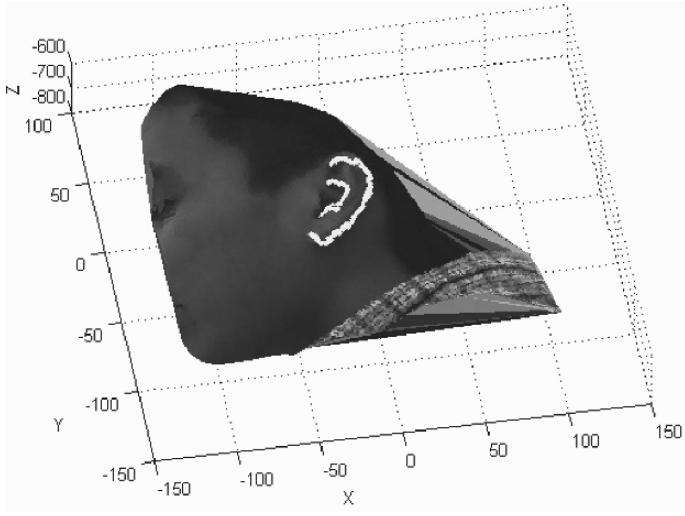

Figure 6-6. The textured 3D face and overlaid ear shape model (The units of X, Y and Z are mm).

Figure 6-7. (a) Original side face range image. (b) Its step edge magnitude image. (c) Its step edge image.

4.3 Edge Thinning and Connected Component Labeling

Since some step edge segments are broken, we dilate the binary image to fill the gaps using a 3×3 structure element. The dilated image is shown in Figure *6-8*(a). We proceed to do edge thinning and the resulting image is shown in Figure *6-8*(b). The edge segments are labeled by running connected component labeling algorithm and some small edge segments (less than 15 pixels) are removed. The edge segments that are left are shown in Figure *6-8*(c).

4.4 Clustering Edge Segments

Since the ear region contains several edge segments, we group edge segments which are close to each other into different clusters. The clustering procedure works as follows:

Figure 6-8. (a) Dilated edge image. (b) Thinned edge image. (c) Edge segments left after removal of small segments.

1. For each edge segment e_i
2. Initialize e_i into a cluster C_i, calculate its centroid $\{\mu_{xi}, \mu_{yi}\}$
3. For each edge segment e_j while $i \neq j$
   - (a) calculate its centroid $\{\mu_{xj}, \mu_{yj}\}$
   - (b) if $\max\{|\mu_{xi} - \mu_{xj}|, |\mu_{yi} - \mu_{yj}|\} \leq \varepsilon_1$, put e_j into cluster C_i, remove e_j and update the centroid of C_i

In the implementation, ε_1 =36 pixels since we want to put ear helix and anti-helix parts into a cluster. Three examples of clustering results are shown in the second row of Figure *6-9*, in which each cluster is bounded by a red rectangular box. The first row of Figure *6-9* shows side face range images.

4.5 Locating Ears by Use of the Ear Shape Model

For each cluster obtained in the previous step, we extract step edges around ear helix and anti-helix parts. We use a threshold ε_2 = 1.9 mm since the step edge magnitudes of vertices in ear anti-helix are at least 1.9 mm and the magnitude of vertices in anti-helix part is smaller than that of vertices in the helix part for the collected data. The problem of locating ears is to minimize the mean square error between ear shape model vertices and their corresponding edge vertices in the bounded rectangular box.

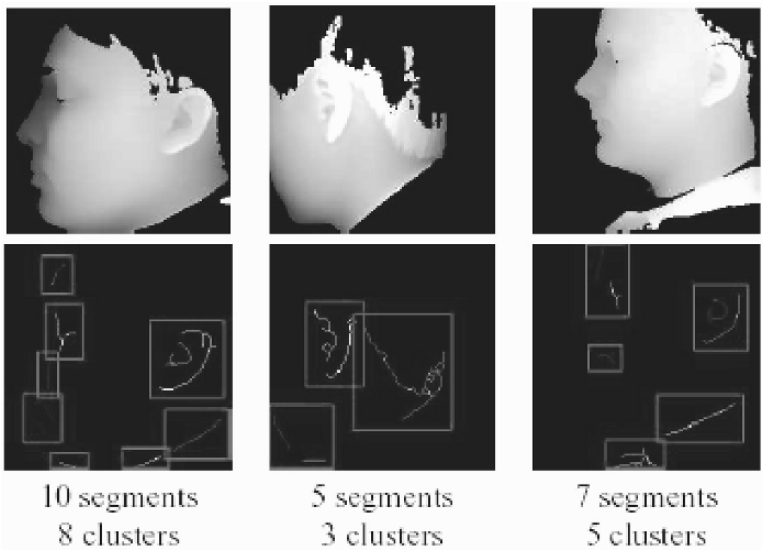

Figure 6-9. Examples of edge clustering results.

$$E = \frac{1}{n}\sum_{l=1}^{n} | T_r(s_i) - I(s_i) |^2 \quad (10)$$

where T_r is the rigid transformation and $I(s_i)$ is vertex in the 3D side face image closest to the $T_r(s_i)$. Iterative Closest Point (ICP) algorithm developed by Besl and Mckay[25] is well-known method to align 3D shapes. However, ICP requires that every point in one set have a corresponding point on the other set, we can't guarantee that edge vertices in the potential regions satisfy this requirement. Therefore, we use a modified ICP algorithm presented by Turk[26] to register the ear shape model with the edge vertices. The steps of modified ICP algorithm to register a test shape Y to a model shape X are:

- 1) Initialize the rotation matrix R_0 and translation vector T_0.
- 2) Given each point in Y, find the closest point in X.
- 3) Discard pairs of points that are too far apart.
- 4) Find the rigid transformation (R, T) such that E is minimized.
- 5) Apply the transformation (R, T) to Y.
- 6) Goto step 2) until the difference $| E_k - E_{k-1} |$ in two successive steps falls below a threshold or the maximum number of iterations is reached.

By initializing the rotation matrix R_0 and translation vector T_0 to the identity matrix and difference of centroids of two vertex sets respectively, we run ICP iteratively and finally get the rotation matrix R and translation vector T, which brings the ear shape model vertices and edge vertices into alignment. The cluster with minimum mean square error is declared as the detected ear region; the ear helix and anti-helix parts are identified in this process.

5. EXPERIMENTAL RESULTS

5.1 Data Acquisition

We use real range data acquired by Minolta Vivid 300. During the acquisition, 52 subjects sit on the chair about 0.55 ~ 0.75m from the camera. The first shot is taken when subject's left side face was approximately parallel to the image plane; two shots are taken when the subject was asked to rotate his/her head to left and right side within 35 degrees with respect to his/her torso. The same acquisition procedure was repeated once for another three shots. Six images per subject are recorded. Therefore, we have 312

images in total. Each range image contains 200×200 grid points and each grid point has a 3D coordinate $\{x, y, z\}$. The ear shape model is built from a side face range image described in Section 4.1. Examples of side face range images are shown in Figure *6.10*.

5.2 Results for the Template Matching Approach

We test the template matching based detection method on 312 side face range images. The parameters are $w = 5$ pixels, $\eta = 0.35$, $\beta = 99$ pixels and $\gamma = 35$ pixels. Figure *6-11* shows examples of positive detection in which the detected ears are bounded by rectangular boxes. If the detected region contains part of ear, we think it is a positive detection; otherwise it is a false detection. From Figure *6-11*, we observe that the ear region is correctly detected.

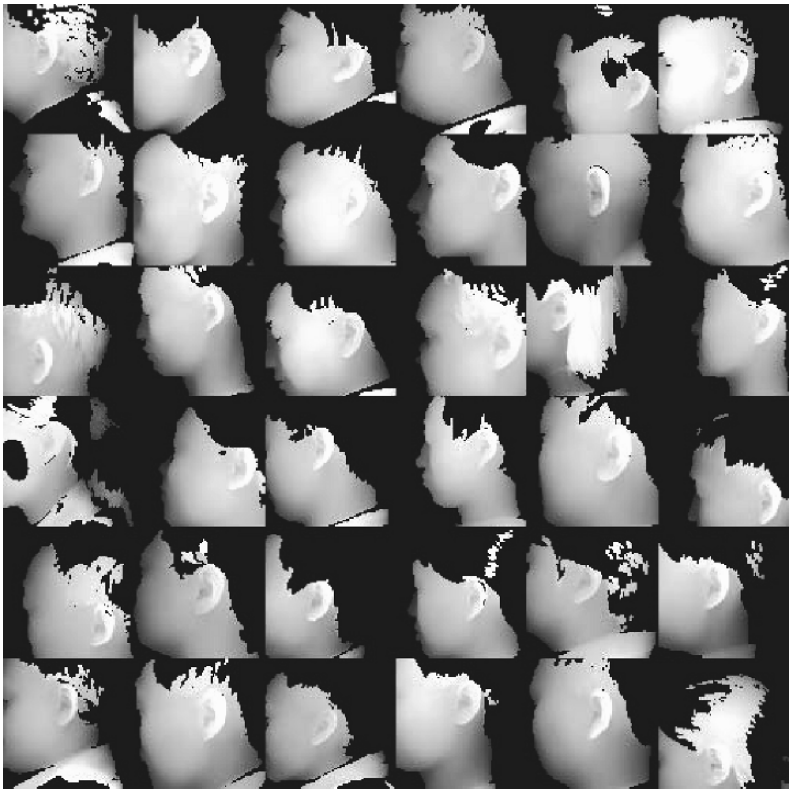

Figure 6-10. Examples of side face range images.

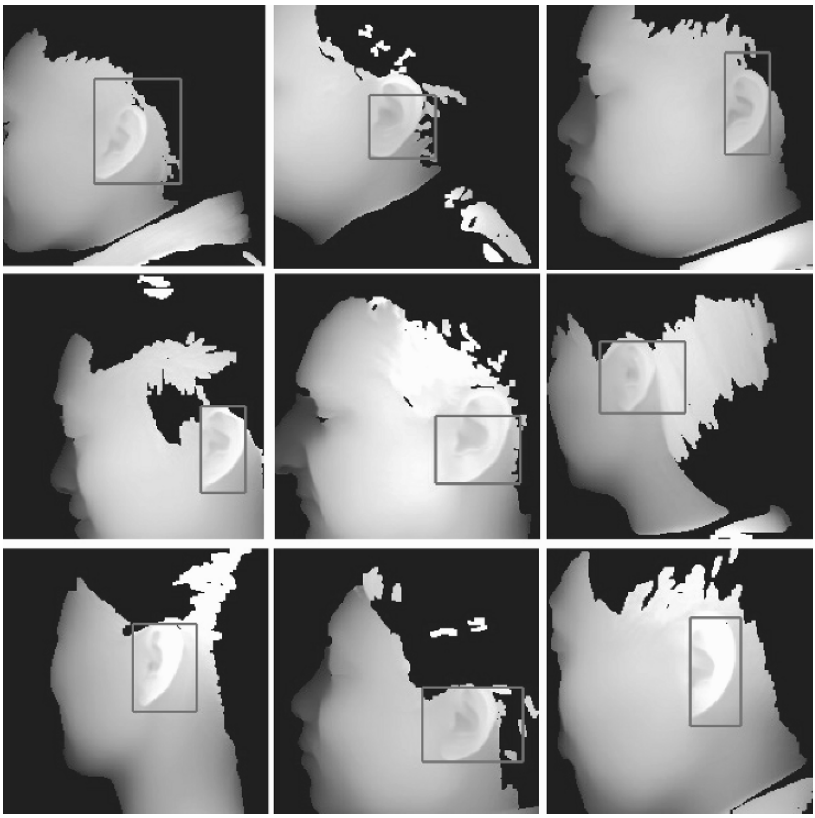

Figure 6-11. Examples of the positive detection using the template matching approach.

However, we may obtain a part of an ear; we may obtain parts that do not belong to an ear. Figure *6-12* shows examples of false detection. Each column in Figure *6-12* shows the step edge magnitude image, the dilated binary edge map and the detection result respectively. Since the ear helix part cannot be extracted, we made false detections. The average time to detect an ear from a side face range image is 5.2 seconds with Matlab implementation on a 2.4G Celeron CPU. We achieve 92.4% detection rate.

5.3 Results for the Shape Model-Based Approach

We test the ear shape model based detection method on 312 side face range images. If the ear shape model is aligned with the ear helix and anti-helix parts, we classify it as a positive detection; otherwise false detection.

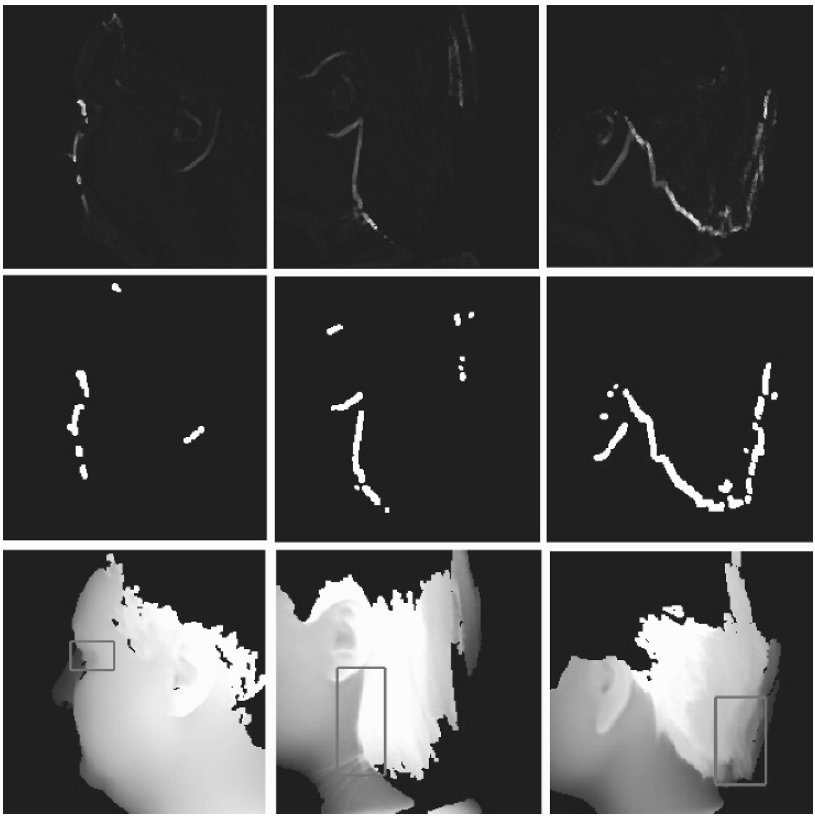

Figure 6-12. Examples of false detection using the template matching approach. Each column shows the step edge magnitude image, the dilated binary edge map and the detection result respectively.

In our experiments, the number of vertices of the ear shape model is 113; the average number of edge segments is 6 and the average number of clusters is 4. The average time to detect an ear from a side face range image is 6.5 seconds with Matlab implementation on a 2.4G Celeron CPU. Examples of positive detection results are shown in Figure *6-13*. In Figure *6-13*, the transformed ear shape model marked by yellow points is superimposed on the corresponding textured 3D face. From Figure *6-13*, we can observe that the ear is correctly detected and the ear helix and anti-helix parts are identified from side face range images. The distribution of mean square error defined in Eq. (10) for positive detection is shown in Figure *6-14*. The mean of mean square error is 1.79 mm. We achieve 92.6% detection rate. After we locate the ear helix and anti-helix parts in a side face range image, we put a minimum rectangular bounding box that contains the detected ear.

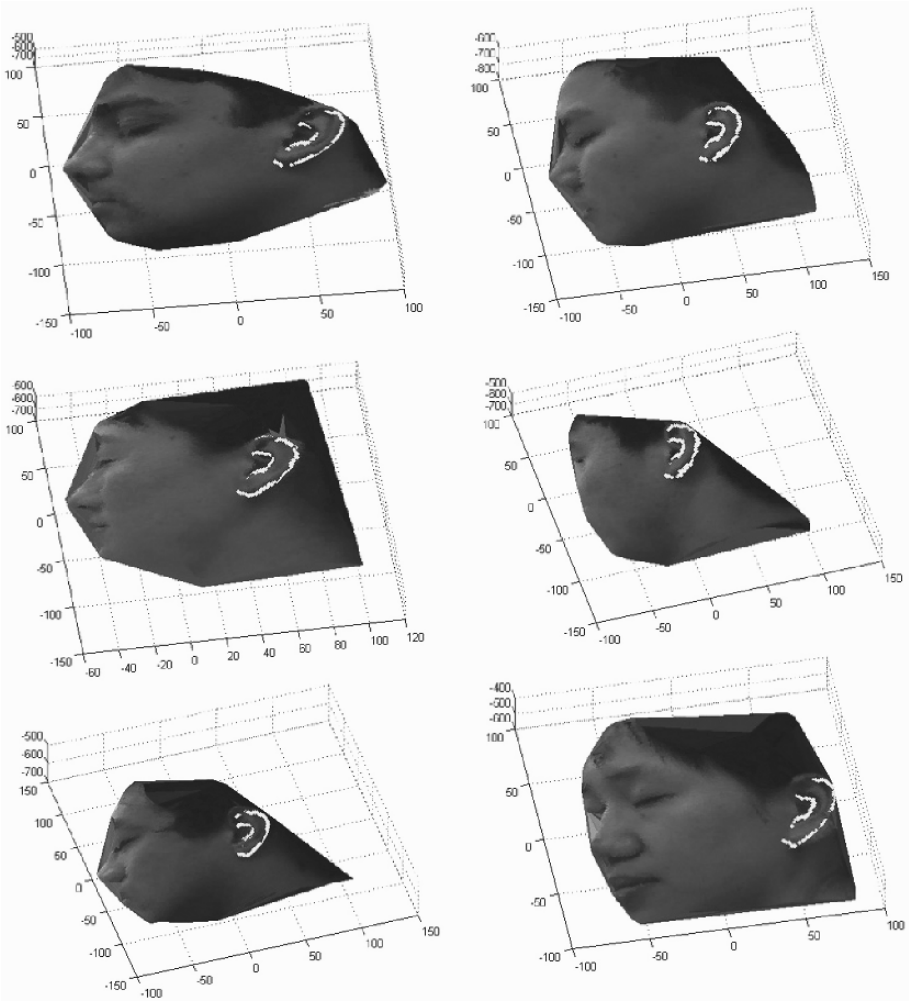

Figure 6-13. Examples of positive detection results using the shape model-based approach (The 3D axes and units are the same as in Figure 6-6).

Figure *6-15* shows the examples of detected ears with a red bounding box. From Figure *6-15*, we observe that the ears are accurately located from side face range images. For the failed cases, we notice that there are some edge segments around the ear region caused by hair, which bring more false edge segments or results in the cluster that cannot include the ear helix and anti-helix parts.

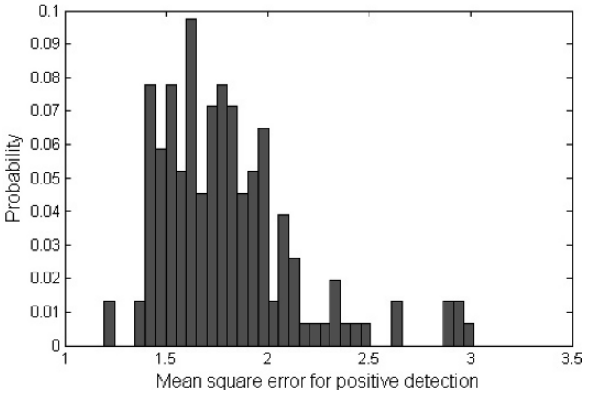

Figure 6-14. Distribution of the mean square error for positive detection using the shape model-based approach.

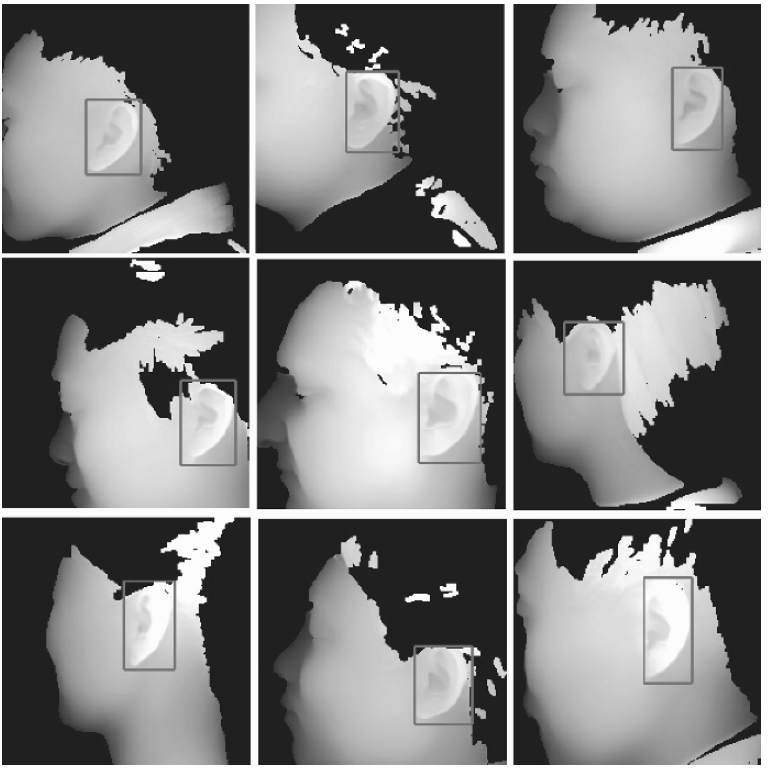

Figure 6-15. Examples of positive ear detection using the shape model-based approach.

Since ICP algorithm cannot converge due to the existence of outliers, the false detection happens; these cases are shown in Figure *6-16* and Figure *6-17*. The original face range images and corresponding edge clusters are shown in Figure *6-16*. In this figure, the first row shows face images; the second row shows edge clustering results. The textured 3D faces with overlaid detected ear helix and anti-helix are shown in Figure *6-17*.

5.4 Comparison of the Two Approaches

Table *6-2* lists the comparison of two approaches in terms of positive detection rate and average detection time. From this table, we observe that the second approach performs slightly better than the first approach and it is a little bit slower. Note that for the template matching approach, if the ear region is roughly detected, it is a positive detection; while for the ear shape model based approach, if the ear helix and anti-helix parts are aligned with the shape model, it is a positive detection. If we compare Figure *6-11* and Figure *6-15*, we see that we can locate ears more accurately by using the shape model-based approach, which can provide us more helpful cues for ear recognition.

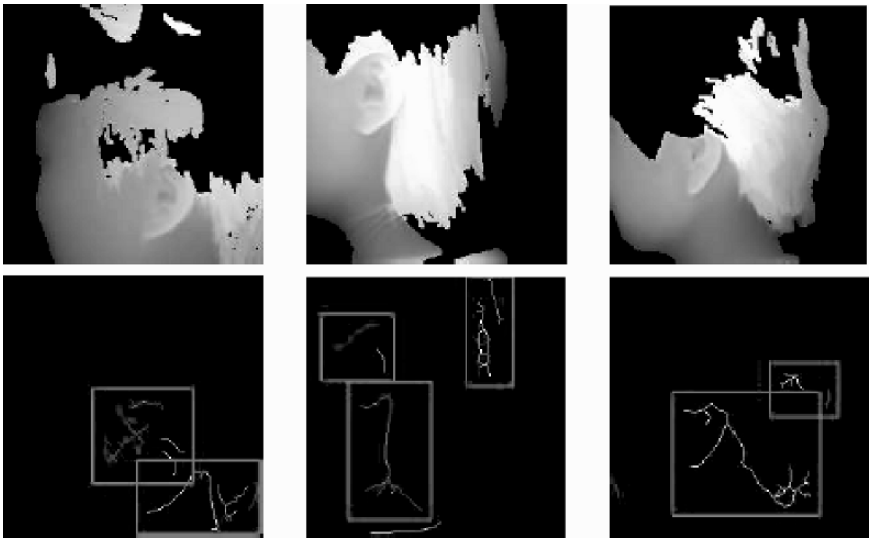

Figure 6-16. Examples of failed cases using the shape model-based approach. Each column shows the range image and the edge clustering result respectively.

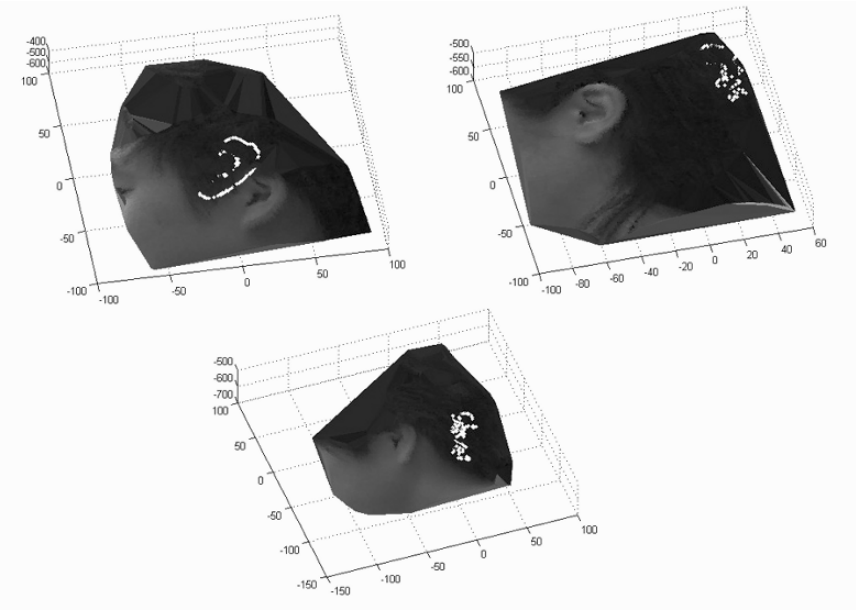

Figure 6-17. Examples of false detection results using the shape model-based approach.

Table 6-2. Comparion of two approaches.

	Detection Rate	Detection Time
Template matching	92.4%	5.2sec
Ear shape model	92.6%	6.5sec

6. CONCLUSIONS

We have proposed two techniques; template matching based detection and ear shape model based detection, to locate ears from side face range images. The success of the first approach relies on two facts: 1) there is a sharp step edge around the ear helix that can be easily extracted; 2) shape index is a good measurement to capture the geometric characteristics of ears since the ear has much ridge and valley areas. The first approach is also simple, effective and easy to implement. For the second approach, the ear shape model is represented by a set of discrete 3D vertices corresponding to ear helix and anti-helix parts. Given side face range images, step edges are extracted, dilated, thinned and grouped into different clusters which are potential regions containing ears. For each cluster, we register the ear shape model with the edges. Our method not only detects the ear region, but also

identifies the ear helix and anti-helix parts. Experimental results on real side face range images demonstrate the effectiveness of our proposed approaches.

REFERENCES

1. A. Iannarelli, Ear Identification, Forensic Identification Series, Paramont Publishing Company, 1989.
2. A. Jain, Personal Identification in Network Society, Kluwer Academic, 1999.
3. D. Hurley, M. Nixon, and J. Carter, Automatic ear recognition by force field transformations, IEE Colloquium on Visual Biometrics, 7/1 --7/5, 2000.
4. M. Burge and W. Burger, Ear biometrics in computer vision, Proc. Int. Conf. on Pattern Recognition, vol. 2, 822-826, 2000.
5. K. Chang, K. Bowyer, S. Sarkar, and B. Victor, Comparison and combination of ear and face images in appearance-based biometrics, IEEE Trans. Pattern Analysis and Machine Intelligence, 25(9), 1160--1165, 2003.
6. B. Bhanu and H. Chen, Human ear recognition in 3D, Workshop on Multimodal User Authentication, 91--98, 2003.
7. H. Chen and B. Bhanu, Contour matching for 3D ear recognition, 7th IEEE Workshops on Application of Computer Vision, vol. 1, 123--128, 2005.
8. P. Yan and K. W. Bowyer, Multi-Biometrics 2D and 3D ear recognition, Audio and Video based Biometric Person Authentication, 503-512, 2005.
9. B. Bhanu, Representation and shape matching of 3-D objects, IEEE Trans. Pattern Analysis and Machine Intelligence, 6(3): 340-351, 1984.
10. B. Bhanu and L. Nuttall, Recognition of 3-D objects in range images using a butterfly multiprocessor, Pattern Recognition, 22(1): 49-64, 1989.
11. H. Chen and B. Bhanu, Human ear detection from side face range images, Proc. Int. Conf. on Pattern Recognition, vol.3, 574--577, 2004.
12. H. Chen and B. Bhanu, Shape model-based 3D ear detection from side face range images, Proc. IEEE Conf. Computer Vision and Pattern Recognition workshop on Advanced 3D Imaging for Safety and Security, 2005.
13. J. Keller, P. Gader, R. Krishnapuram, and X. Wang, A fuzzy logic automatic target detection system for LADAR range images, IEEE International Conference on computatinoal intelligence, pp. 71-76, 1998.
14. E. Meier and F. Ade, Object detection and tracking in range images sequences by separation of image features, IEEE International conference on Intelligent Vehicles, 176-181, 1998.
15. J. Sparbert, K. Dietmayer, and D. Streller, Lane detection and street type classification using laser range images, IEEE Intelligent Transportation Systems conference proceedings, 454-459, 2001.
16. J. Garcia, J. Valles, and C. Ferreira, Detection of three-dimensional objects under arbitrary rotations based on range images, Optics Express, 11(25), 3352-3358, 2003.
17. B. Heisele and W. Ritter, Segmentation of range and intensity image sequences by clustering, Proc. IEEE Conf. on Information Intelligence and Systems, 223-227, 1999.
18. C. Boehnen and T. Russ, A fast Multi-Modal approach to facial feature detection, 7th IEEE Workshops on Application of Computer Vision, 1:135-142, 2005.
19. F. Tsalakanidou, S. Malasiotis, and M. G. Strintzis, Face localization and authentication using color and depth images, IEEE Trans. on Image Processing, 14(2):152-168, 2005.

20. C. Dorai and A. Jain, COSMOS-A representation scheme for free-form surfaces, Proc. Int. Conf. on Computer Vision, 1024-1029, 1995.
21. J. J. Koenderink and A. V. Doorn, Surface shape and curvature scales, Image Vision Computing, 10(8), 557--565, 1992.
22. P. Flynn and A. Jain, On reliable curvature estimation, Proc. IEEE Conf. Computer Vision and Pattern Recognition, 110-116, 1989.
23. N. Yokoya and M. D. Levine, Range image segmentation based on differential geometry: A hybrid approach. IEEE Trans. Pattern Analysis and Machine Intelligence, 11(6), 643-649, 1989.
24. B. Schiele and J. Crowley, Recognition without correspondence using multidimensional receptive field histograms, International Journal of Computer Vision, 36(1), 31-50, 2000.
25. P. Besl and N. D. Mckay, A method of registration of 3-D shapes, IEEE Trans. Pattern Analysis and Machine Intelligence, 14(2), 239-256, 1992.
26. G. Turk and M. Levoy, Zippered polygon meshes from range images, Proceedings of Conf. on Computer Graphics and Interactive Techniques, 311--318, 1994.