

3D Human Face Recognition Using Point Signature

Chin-Seng Chua, Feng Han, Yeong-Khing Ho
School of Electrical and Electronic Engineering
Nanyang Technological University, Singapore 639798
ECSSChua@ntu.edu.sg

Abstract

In this paper, we present a novel face recognition algorithm based on the Point Signature - a representation for free-form surfaces. We treat face recognition problem as a non-rigid object recognition problem. The rigid parts of the face of one person are extracted after registering the range data sets of faces having different facial expressions. These rigid parts are used to create a model library for efficient indexing. For a test face, models are indexed from the library and the most appropriate models are ranked according to their similarity with the test face. Verification of each model proceeds according to their ranking. In this way, the correct model face can be quickly and efficiently identified. Experimental results with range data involving six human subjects, each with four different facial expressions, have demonstrated the validity and effectiveness of our algorithm.

1 Introduction

Automatic identification of human faces is a very challenging research area, which has gained much attention during the last few years. Most of this work, however, is focused on intensity or color images of faces [2]; only a few approaches deal with range images [7, 8, 6]. However, there is evidence that range images have the potential to overcome some problems of intensity and color images. Some advantages of range images are the explicit representation of 3D shape, which is invariant under the change of color and reflectance properties of the objects. An example of the last type of invariance is images of a person with and without facial makeup. While the basic biometrics properties of the face represented by a range image remain the same, the visual appearance of a person can be altered with makeup in such a way that even human intelligence can sometimes be misled to recognize that person.

Many researchers have handled the 3D-face recognition problem using differential geometry tools of computing of

curvatures [7, 8, 6]. Lee and Milios [7] extracted the convex regions of human faces by segmenting the range images based on the signs of the mean and Gaussian curvature at each point. These regions form the basic set of features of a face. For each of these convex regions the Extended Gaussian Image is extracted which is then used to match the facial features of the two face images. Gordon [6] presented a detailed study of face recognition based on depth and curvature features. The curvatures on a face are calculated to find face-specific descriptors (e.g. nose ridge and eye features). Comparison between two faces is made based on their relationship in the feature space. Hiromi *et. al* [8] treated the face recognition problem as a 3D shape recognition problem of rigid free-form surfaces. Each face in both input images and model database, is represented as an Extended Gaussian Image (EGI), constructed by mapping principal curvatures and their directions at each surface points, onto two unit spheres, each of which represents ridge and valley lines respectively. Individual faces are then recognized by evaluating the similarities among others using Fisher's spherical correlation on EGI's of faces. However, the computation of curvature is neither accurate nor reliable [4]. As to treating the face recognition problem as a 3D recognition problem of rigid surface, such as in [7, 8, 6], it is based on the assumption that the shape of faces doesn't change significantly. However, this is usually not the case, especially when dealing with faces of different facial expressions. With the latter, the shape of many parts of the face will change considerably (see Figure 1); the face recognition problem should be treated as a 3D recognition problem of non-rigid surfaces. New methods are required to handle this difficult problem. In this paper we propose the use of Point Signature [5] to recognize non-rigid objects. This is an extension of its use in [5], which uses Point Signature to recognize 3D rigid objects. In particular, we are concerned with the identification of faces, despite they having different facial expressions.

While the face-shape of the same person may change, sometimes greatly, due to different facial expressions, there will still be regions, such as nose, eye socket region and



Figure 1. Rendered 3D faces of six human subjects, each with four facial expressions.

forehead, which will keep their shape and position or be subjected to much less deformation between any different expression. If these regions can be identified, the 3D non-rigid face recognition problem can be reduced to the rigid case. Based on this observation, our approach is to extract the rigid parts of the face and utilize them to realize the task of recognition. In this way, the greatly changing facial regions can be rejected and thus will not affect the recognition result. Moreover, only using the most stable and meaningful parts of the face will compress the information stored in the model library, which undoubtedly will improve the speed of recognition considerably.

This paper is organized as follows. Section 2 discusses the registration of two face-surfaces with different facial expressions using Point Signature. The optimal registration should be based only on the "rigid" regions of the face, which are not known *a priori*. At the same time, it is from this registration that the "rigid" regions of the face are derived. This section will propose an algorithm to solve this "catch-22" problem. Having established the approach to optimally register face-surfaces, Section 3 presents the methodology of recognizing a face from a library of face-subjects. In this section, the build-up of the face-library and

the indexing of a candidate face is discussed. The efficiency of the algorithm is verified with test results and this is presented in Section 4. The tests are based on real range data, sampled from six different faces of human subjects.

2 Face-Surface Registration Using Point Signature

Given two face-surfaces of different facial expressions, but from the same person, the aim of registration is to find the transformation which optimally aligns the rigid regions of the two face-surfaces. From the registration the rigid regions are identified as containing range points with low registration errors. Conversely, the non-rigid, or flexible, regions will have range points of high registration errors. These non-rigid regions are rejected and are not used to build the library of the face-surfaces which will be discussed in Section 3.

2.1 Definition of point signature

The definition of the Point Signature is summarized here for completeness. For details, the reader may refer to [5].

For a given point p , we place a sphere of radius r , centered at p . The intersection of the sphere with the object surface is a 3D space curve, C , whose orientation can be defined by an orthonormal frame formed by a normal vector, $\mathbf{n}_1$, a "reference" vector, $\mathbf{n}_2$, and the vector cross-product of $\mathbf{n}_1$ and $\mathbf{n}_2$. $\mathbf{n}_1$ is defined as the unit normal vector of a plane fitted through the space curve C which need not be planar; the planar fit is used as an approximation to provide a vector direction. A new plane $\mathbf{P}'$ is defined by translating the fitted plane to the point p in a direction parallel to $\mathbf{n}_1$. The perpendicular projection of C to $\mathbf{P}'$ forms a new planar curve C' with the projection distance of points on C' forming a signed distance profile. The reference direction, $\mathbf{n}_2$, is defined as the unit vector from p to the projected point on C' which gives the largest positive distance.

Every point on C may now be characterized by (see Figure 2):

1. The signed distance from itself to the corresponding point on C .
2. A clockwise rotation angle θ about $\mathbf{n}_1$ from the reference direction $\mathbf{n}_2$.

We refer to this profile of distances, $d(\theta)$, with $0 \leq \theta \leq 360$ degrees, as the signature at point p . Practical consideration, however, do not permit us to consider all points on the space curve C but rather an angular sampling of points starting from the reference direction $\mathbf{n}_2$. Depending on the resolution of the input range data, the angular sampling interval, $\Delta\theta$, may vary from 5 to 15 degrees. (In our implementation $\Delta\theta = 10$ is used). Hence, the distance profile may now be represented by a discrete set of values $d(\theta_i)$ for

$i = 1, \dots, n_\theta$, $0 \leq \theta_i \leq 360$, where n_θ is the number of samples.

2.2 Matching of point signatures

The similarity matching between a given signature $d_s(\theta_i)$ and a candidate signature $d_m(\theta_i)$ is as follows:

If the signature $d_s(\theta_i)$ has a global maximum that is significantly different from the other local maximum, then $d_s(\theta_i)$ would be uniquely oriented by the reference vector $\mathbf{n}_2$. In this case $d_s(\theta_i)$ can be directly compared with the candidate signature $d_m(\theta_i)$ using the following constraint:

$$|d_s(\theta_i) - d_m(\theta_i)| \leq \varepsilon_{tol}(\theta_i) \quad \forall i = 1, \dots, n_\theta \quad (1)$$

where the tolerance band $\varepsilon_{tol}(\theta_i)$ is used to handle the errors in computing Point Signature and achieve a better acceptance and rejection of candidate signatures.

However, if $d_s(\theta_i)$ has more than one local maximum which are similar, then for each position of local maximum, $d_s(\theta_i)$ is phase shifted before comparison is carried out at that position:

$$|d_s(\theta_i - \theta_j) - d_m(\theta_i)| \leq \varepsilon_{tol}(\theta_i) \quad \forall i = 1, \dots, n_\theta, \quad (2)$$

$\theta_j \in \text{position of local discrete maximum}$

2.3 Registering two face-surfaces with different expressions

Given two face surface data sets, $S = \{s_i : i = 1, \dots, n_s\}$ and $M = \{m_j : j = 1, \dots, n_m\}$, the task of 3D registration is to discover the transformation (translation and rotation) which will optimally align the rigid regions of

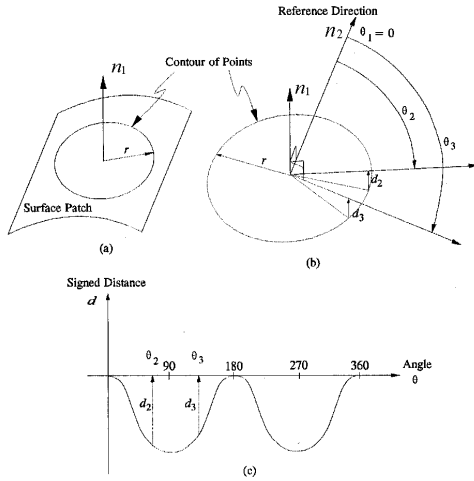


Figure 2. Definition of point signature.

S with those of M . The two face shapes need not be identical, which makes this registration different from the traditional one [1, 4, 9, 5, 3].

First, we select three points s_α , s_β and s_γ in S using the approach given below. Next, three points m_α , m_β and m_γ in M are hypothesized to correspond to s_α , s_β and s_γ using the constraints discussed below. If the hypothesis is acceptable, the transformation from the 3-tuple $\langle s_\alpha, s_\beta, s_\gamma \rangle$ in S to the 3-tuple $\langle m_\alpha, m_\beta, m_\gamma \rangle$ in M represents a possible registration solution. This solution is verified by the criterion given below. If it meets the need of this criterion, registration is achieved with measurable quality.

A) Selecting s_α, s_β and s_γ

Our purpose is to register the rigid regions of surface S and surface M . So s_α , s_β and s_γ should be chosen on the rigid regions of surface S . If any one of them is on the deforming regions, the transformation produced will definitely be incorrect. From the definition of Point Signature we can see that the signature of one point describes the local underlying surface structure in the neighborhood of this point. Almost all the points on the rigid regions of S will find points with the same signature on M , because their corresponding points on M will have similar local underlying surface structure and signature. At the same time, almost all the points on the non-rigid regions of S will not find points with the same signature on M because their corresponding points on M will have different local underlying surface structure; their signatures will not be similar. Based on this observation, we pick up all the points on S , which can find points with the same signature on M , to form a point set, namely S^r . Selecting s_α , s_β and s_γ from S^r , we can make it more likely that they will produce correct registration and avoid unnecessary computation.

From S^r , many 3-tuples of $\langle s_\alpha, s_\beta, s_\gamma \rangle$ can be formed. If the triangle formed by s_α , s_β and s_γ is too small, the computed transformation will be very sensitive to noise [3]. To prune this redundant set of S^r , we select only those 3-tuples where the lengths of all three sides of $\langle s_\alpha, s_\beta, s_\gamma \rangle$ exceed a predefined threshold. Also, we do not allow close neighbours to be included once a particular point is chosen. The results of all valid selections consist of a set $S^t = \{\langle s_{\alpha i}, s_{\beta i}, s_{\gamma i} \rangle : i = 1, \dots, t_s\}$, from which the needed s_α , s_β and s_γ can be chosen.

B) Selecting m_α, m_β and m_γ

Having chosen s_α , s_β and s_γ on S , the next step is to find all the possible m_α , m_β and m_γ on M , that may correspond to s_α , s_β and s_γ . To select any set of three model points as possible correspondents is not only computationally expensive but also unnecessary. Correspondence can be established based on the following constraints:

1. Point Signature constraint

Let $d_{s_p}(\theta_i)$ and $d_{m_p}(\theta_i)$ denote the Point Signature of point s_p and m_p . According to section 2.2, for correspon-

dence they should match with each other within an error tolerance band $\varepsilon_{tol}(\theta)$.

2. Distance Constraint

Let $d(s_i, s_j)$ represent the Euclidean distance between range points s_i and s_j . For correspondence, the distance between any two points of s_α , s_β and s_γ should match, within an error of tolerance, ε_d , to the corresponding two points of m_α , m_β and m_γ .

3. Direction Constraint

Let $n_1(\cdot)$, $n_2(\cdot)$ and $n_3(\cdot)$ denote finding, at the point of interest, the unit vectors representing the directions of the normal vector, reference vector and the vector cross-product of normal vector and reference vector. Let $\hat{\mathbf{p}}_1$ and $\hat{\mathbf{p}}_2$ represent the unit vectors from s_α to s_β and s_γ . Similarly, let $\hat{\mathbf{q}}_1$ and $\hat{\mathbf{q}}_2$ represent the unit vectors from m_α to m_β and m_γ . To provide a tolerance for error, ε_θ , for correspondence (see Figure3):

$$|\Delta\theta_1| = |\arccos(\mathbf{n}_1(s_\alpha) \cdot \hat{\mathbf{p}}_1) - \arccos(\mathbf{n}_1(m_\alpha) \cdot \hat{\mathbf{q}}_1)| \leq \varepsilon_\theta \quad (3)$$

$$|\Delta\theta_2| = |\arccos(\mathbf{n}_2(s_\alpha) \cdot \hat{\mathbf{p}}_1) - \arccos(\mathbf{n}_2(m_\alpha) \cdot \hat{\mathbf{q}}_1)| \leq \varepsilon_\theta \quad (4)$$

$$|\Delta\theta_3| = |\arccos(\mathbf{n}_3(s_\alpha) \cdot \hat{\mathbf{p}}_1) - \arccos(\mathbf{n}_3(m_\alpha) \cdot \hat{\mathbf{q}}_1)| \leq \varepsilon_\theta \quad (5)$$

The same condition applies for $\hat{\mathbf{p}}_2$ and $\hat{\mathbf{q}}_2$.

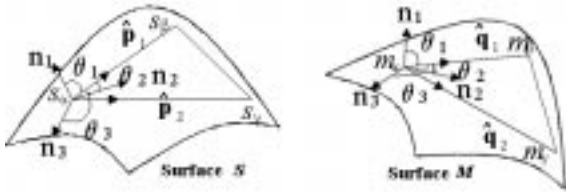


Figure 3. The direction constraints for correspondence.

Constraint 1 is first used to find a set of points on M : $m_{\alpha i}, i = 1, \dots, p$ whose signatures match with that of s_α . For each association of s_α and $m_{\alpha i}$, the directional frame for s_α can be aligned with that of $m_{\alpha i}$ by a rigid transformation [4]. Using this transformation, we can transform s_β and s_γ to the space of M . The positions of these two transformed points indicate the approximate regions in which $m_{\beta i}$ and $m_{\gamma i}$ should be searched with the above three constraints as the corresponding points to s_β and s_γ .

Repeating the above procedure, we obtain a set of n_t 3-tuples $\langle m_{\alpha k}, m_{\beta k}, m_{\gamma k} \rangle, k = 1, \dots, n_t$. For each association of $\langle s_\alpha, s_\beta, s_\gamma \rangle$ and $\langle m_{\alpha k}, m_{\beta k}, m_{\gamma k} \rangle$ a transformation is computed using the method in [4]. For the set of

n_t 3-tuples, a set of n_t transformations, $\tau = \{T_k : k = 1, \dots, n_t\}$, is formed.

C) Selecting the optimal transformation and verifying

A set of n_t possible transformations $\tau = \{T_k : k = 1, \dots, n_t\}$ have been found based on the correspondences between the 3-tuple $\langle s_\alpha, s_\beta, s_\gamma \rangle$ and the 3-tuples $\langle m_{\alpha k}, m_{\beta k}, m_{\gamma k} \rangle$. The problem here is to single out the optimal transformation and determine if it is good enough to be an accepted solution for the registration task. To attenuate the disturbance of the points on the non-rigid regions, we only apply the transformation to the points in S^r , from which we select s_α, s_β and s_γ . When the optimal transformation is applied to the points of S^r , it will closely register most of the points on S^r with the points on M .

Let ϑ be the estimation of the maximum distance between one point in S^r and its closest points on M for a good estimate of the best registration (it depends on the noise level of the range data). With the transformation T_k , all the points $s_i : i = 1, \dots, n_s$ in S^r can be brought to new positions $s'_i : i = 1, \dots, n_s$. We count the number of occurrences, namely n_r , that the distance between s'_i and its nearest point on M is smaller than ϑ . We call the ratio of n_r to n_s the *registration rate*. The transformation in τ with the highest *registration rate* is selected as the optimal one. If the *registration rate* of the optimal transformation is greater than a threshold, this transformation will be verified as one accepted solution to the registration task. Otherwise, another s_α, s_β and s_γ will be picked up from S^t to repeat the process.

The registration provided by our method is very good and accurate to some extent. This result can be further refined by some well-known iterative procedures [1, 9]. In our method, the ICP method in [9] is used for refinement with two improvements:

1. Since the two given surface are not globally rigid, it does not make sense to register based on the whole point-set of S , as was originally done in [1, 9]. In our case, we only make use of points in S^r , which can be considered as "locally rigid".

2. In ICP, "pseudo-correspondences" must be established between points on S and M . For a particular transformation, T_k and a given point s on S , the corresponding point on M is hypothesized as $m \in M$ which is closest (in Euclidean 3D space) from the transformed point of $s \in S$ (that is $T_k s$). However, due to the non-rigid of S and M , this measure used in ICP is no longer valid. In many cases, the closest point on M may not have the same structure (i.e. the point signature) as the transformed point $T_k s$. To ensure the conformity of the underlying structure, we preselect the points on M having similar point signature as s , and from these choose the nearest point on M for the "pseudo-correspondence".

The registration result is shown in Figure 4.

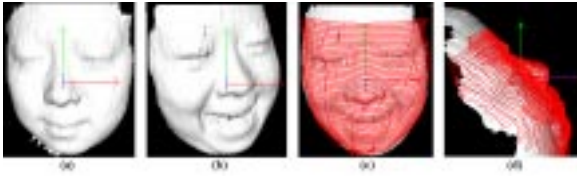


Figure 4. Registration of two face-surfaces. (a), (b) rendered surfaces of two facial expressions. (c) registration of surface (a) (wire-frame) onto (b) (rendered). (d) side-view of the registration.

2.4 Finding the rigid parts of the face surface

After registering two different face surfaces, it is crucial to separate the rigid parts from the non-rigid parts. Although the rigid parts of two face surfaces will be nearer to each other than the non-rigid parts after registration, it is difficult to define a constant threshold to discriminate this difference. To handle this, we model the point distances between two registered face surfaces as a Gaussian Distribution:

$$f(x) = \frac{1}{(2\pi\sigma)^{1/2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < +\infty \quad (6)$$

where x is the distance between any point on S and its nearest point on M after registration, μ is the mean of the distance distribution and σ is the standard deviation. For each registration, we compute an adaptive threshold as $(\mu + \sigma)$ to distinguish between rigid and non-rigid parts. Figure 5 shows the rigid parts extracted after registration for each of the first three candidates shown in Figure 1.

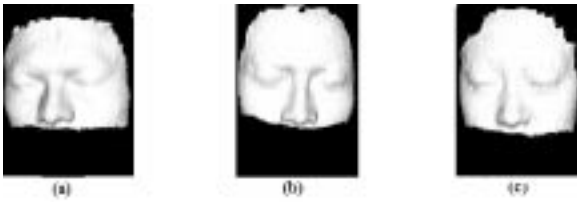


Figure 5. The extracted rigid face regions of each of the first three persons in Figure 1. (a) person 1. (b) person 2. (c) person 3.

3 Face Recognition

3.1 Creating the model library

Prior to recognition, the model library must be built up. This is done by computing the point signature for every point on the rigid regions of each model, and for every model in the library. For fast retrieval of every point signature, a three dimensional indexing table indexed by the number of peaks (maximum and minimum), maximum, minimum of the point signatures is tabulated for all signatures. Each entry of the table serves as a quick reference to the group of signatures satisfying the peak number, maximum, and minimum constraints; and each signature within the entry references the model index and from which the signature was extracted. Because the library doesn't contain all the information of the model faces, but only the most stable and meaningful part, this considerably deduces the complexity of the indexing table and thus the time complexity of indexing process.

3.2 Selecting the most likely model and verifying

For a given range image of a test face, we want to select a suitable set of scene points that can be used to hypothesize and "vote" for likely model faces. The scene data is pre-processed by computing the point signature of each range point and selecting only those points with different point signatures. This reduced set of scene points, S_c , are used to "vote" for the model faces. By voting, we mean that a scene point is used to index into the model library (Section 3.1) and every model having points of similar signature receives a vote. This is repeated for the other points in S_c . For n_c points in S_c , we would expect n_c votes for the correct model, assuming that (1) both the test and model face are in similar orientation, (2) points of s_c are chosen over rigid regions of the face, and (3) negligible errors in both source range data, computation and matching of point signatures.

However, the above can not be assured and usually the number of the votes for even the most appropriate model will be less than n_c . Nevertheless, the models can be ranked according to their similarity with the test scene. For each model, we term the ratio of number of votes received, say n_m , to the number of voting scene points, n_c , as the voting rate. The voting rate represents the likelihood of each model being correctly matched with the scene. Using the voting rate for each model, inappropriate models with low voting rate are rejected. The remaining candidate models are ordered according to the voting rate. With this ordering, verification can be carried out starting with the most likely candidate.

Noting that each model contains the rigid parts of the face, if it is correctly matched with the scene, a transformation with a very small registration error can be found using the approach of Section 2.3. So we can verify that the current candidate model is the correct one by testing if its registration error with the scene is small enough. It may happen that another similar face model has higher voting rate than the correct one, but such cases seldom occur as can be seen from the experiment results. Even then, the rejection of the incorrect model can be quite fast because of its large registration error with the scene.

4 Experimental Results

We have implemented the 3D Point Signature-Based face recognition algorithm on a Pentium III 450. We have tested the algorithm for 6 face range images extracted from a range finder with a resolution of 1 part in 7500 of the field of view (Figure 1). Four face range images with different expressions are captured for each of the six persons (Figure 1). These range images are used to extract the rigid parts and subsequently to build the model library. Another range image is extracted for each person to represent the test scene (Figure 6).



Figure 6. Six test faces.

For each test scene, the models are voted using the indexing table created from the model library. Table 1 shows the voting results (in percentage) of the test scenes, Face Scene 1 to 6. From the results, the correct model is considerably identified (highest voting rate and highlighted in bold). With this ranking, verification is carried out by registering the test scene with the hypothesized model face. With a low registration error, the hypothesized face is quickly identified.

5 Conclusion

We have proposed and implemented a 3D face recognition approach using Point Signature. In our approach, we don't assume limiting assumptions of rigid (or approximately rigid) surfaces but have demonstrated with faces having widely different facial expressions for each human subject. Four different facial expressions are used. The extracted range data sets are registered together, the registration being defined by the rigid parts of the face. In this way, a library

of faces can be built up. Recognition of a test face is relative fast with the efficient indexing into the model library. The most appropriate models are ranked according to their similarity with the test scene. Verification of each model proceeds according to their ranking. In this way, the correct model face can be quickly and efficiently identified. Experimental results involving six human subjects, each with four different facial expressions, have demonstrated the validity and effectiveness of our algorithm.

References

- [1] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, February 1992.
- [2] R. Chellapa, C. L. Wilson, and S. Sirohey. Human and machine recognition of faces: A survey. in *Proceedings of the IEEE*, 83(5):705–740, May 1995.
- [3] C.-S. Chen, Y.-P. Huang, and J.-B. Cheng. A fast automatic method for registration of partially-overlapping range images. *Sixth International Conference on Computer Vision*, pages 242–248, 1998.
- [4] C. S. Chua and R. Jarvis. 3d free-form surface registration and object recognition. *International Journal on Computer Vision*, 17(1):77–99, 1996.
- [5] C. S. Chua and R. Jarvis. Point signature: A new representation for 3d object recognition. *International Journal on Computer Vision*, 25(1):63–85, 1997.
- [6] G. G. Gordon. Face recognition based on depth and curvature features. in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 808–810, 1992.
- [7] J. C. Lee and E. Milios. Matching range image of human faces. *Third International Conference on Computer Vision*, pages 722–726, 1990.
- [8] H. T. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. *Third IEEE International Conference on Automatic Face and Gesture Recognition*, pages 372–377, 1998.
- [9] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal on Computer Vision*, 13(2):119–152, 1994.

	voting rate of each model					
	Scene Face 1	Scene Face 2	Scene Face 3	Scene Face 4	Scene Face 5	Scene Face 6
Model 1	82.58	60.61	36.36	46.21	34.09	28.79
Model 2	57.02	83.97	23.97	48.76	37.19	28.93
Model 3	73.44	70.94	93.75	31.25	56.25	46.88
Model 4	40.00	63.57	20.71	80.00	41.43	33.57
Model 5	60.14	76.08	47.55	62.94	90.21	49.65
Model 6	76.92	60.19	42.31	70.19	71.15	78.85

Table 1. Voting rate of each model using the different point signatures of scene (face 1 to face 6).