Sphere-Spin-Image: A Viewpoint-Invariant Surface Representation for 3D Face Recognition⋆

Yueming Wang, Gang Pan, Zhaohui Wu, and Shi Han

Department of Computer Science and Engineering
Zhejiang University, Hangzhou, 310027, P. R. China
`{ymingwang, gpan}@zju.edu.cn`

Abstract. This paper presents a new free-form surface representation scheme, which we call *Sphere-Spin-Image*(SSI), and its application to 3D face recognition. An SSI, associated with a point on the surface, is a 2D histogram constructed from a neighborhood surface of the point using position information, which captures the characteristic of local shape. Thus, a free-form surface can be represented by a series of SSIs. Correlation coefficient is used as similarity metric for comparing SSIs. During face recognition, the SSIs of points on face surface are computed, and recognition task is achieved by SSI-comparison-based voting method. To reduce computational cost, only some particular points of face surface are involved in voting. With face database consisting of 31 different pose models, experimental result of equal error rate 8.32% demonstrates performance of the proposed method.

1 Introduction

Free-form surface representation schemes are widely used both in computer graphics and computer vision, and have been studied extensively recently. For the purpose of registration and recognition, the representation scheme should be (1)viewpoint-independent, (2)general enough to describe the sculpted objects, (3)as compact and expressive as possible. Several representation schemes have been presented during the last few years. Dorai et al. [12] used shape index($S=1/2$ $-\frac{1}{\pi}\arctan\frac{K_1+K_2}{K_1-K_2}$) along with a spectral extension of the shape measure to build a view-dependent representation of free-form surface named *COSMOS*. Johnson and Hebert [9,10] introduced *spin image* which comprises descriptive images associated with oriented points on the surface. Using a single point basis, the positions of the other points on the surface are described by two parameters. The accumulation of these parameters for many points on the surface of the object results in an image at each oriented point. *Point signature* proposed by Chua and Jarvis [2] serves to describe the structural neighborhood of a point by encoding the minimum distances of points on a 3D contour to

a reference plane. Yamany and Farag [11] generated a *signature image* where the image axes are the distance between a point basis and other points on the surface and the angle between the normal at the point basis and the vector from the point to each other point. The signature image encodes the mean curvature.

Automatic face recognition has been studied extensively in computer vision area over past decade. However, most approaches of face recognition are based on intensity images of faces, only a few exploited three dimensional information [1]. Face recognition paradigms based on range data images have more potential than intensity images based approaches because the former is invariant under the changes of lighting conditions, color and reflectance properties of face while the latter is sensitive to above changes. Previous work of face recognition based on range data is briefly reviewed in the following. Gordon [6] employed depth and curvature features from face range data and performed face recognition. It is done by computing curvatures of face surface to find fiducial regions such as eye corner cavities to calculate the face-specific descriptors(eg. eye separation etc.). Tanaka et al. [5] presented a correlation-based face recognition approach based on the analysis of maximum and minimum principal curvatures and their directions. Chua and Han [3] treated face recognition problem as a non-rigid object recognition problem. The rigid parts of the face were extracted after registering the range data of faces having different facial expressions using point signature. The most appropriate models from the database constructed with the rigid parts of face surfaces are ranked according to their similarity with the test face.

Motivated by *spin image* [10] and *point signature* [3], we present a new approach of 3D local shape representation named *Sphere-Spin-Image*(SSI) . The SSI of a point is constructed by mapping 3D coordinates of those surface's points within a sphere centered at this point into 2D space. Since SSI represents the local shape of a point, it can be utilized as a intermediate representation in 3D face recognition.

This paper is organized as follows: The SSI representation is described in Section 2. Section 3 discusses the special points selection process and the voting method based on SSI comparison for face recognition. The experimental results and conclusion are given in section 4 and section 5 respectively.

2   Sphere-Spin-Image(SSI) Representation Scheme

2.1   Definition of Sphere-Spin-Image

Given a point p with unit normal n and tangent plane TP_p through p on the surface S, each other point p_k on the surface S can be related to p by two parameters: 1) the distance $\rho_k=\|p\text{-}p_k\|$; 2) the distance from p_k to the tangent plane TP_p, $d_k=n\cdot(p_k\text{-}p)$, as shown in Fig.1 (a). Apparently, d_k could be either positive or negative while ρ_k could not be negative.

The Sphere-Spin-Image(SSI) is defined as follows: for each point p on the surface S, shown in Fig. 1 (b), a sphere with radius r can be placed centrally at

(a)　　　　　　　　(b)　　　　　　　　(c)
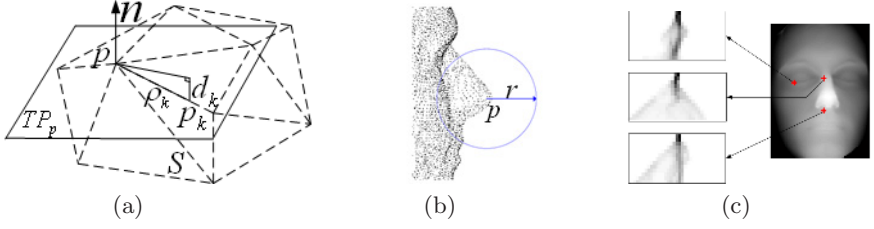
Fig. 1. (a) Relative position of point $\boldsymbol{p_k}$ to point $\boldsymbol{p}$. (b) A surface patch surrounded by a sphere centered at $\boldsymbol{p}$. (c) Examples of SSIs taken at different point location on face surface. Notice the difference between the SSI corresponding to a location and SSI of other location.

$\boldsymbol{p}$. The SSI of $\boldsymbol{p}$ is a 2D histogram of ρ_k and d_k formed by those points both on the surface S and in the sphere. It can be formalized as:

$$\boldsymbol{SSI}: S^r(\boldsymbol{p}) = (S^r_{ij}(\boldsymbol{p}))_{m \times n, 1 \le i \le m, 1 \le j \le n} \tag{1}$$

$$S^r_{ij}(\boldsymbol{p}) = \sum_k s^k_{ij} \tag{2}$$

where

$$s^k_{ij} = \begin{cases} 1, & \text{if } \boldsymbol{p_k} \in S \text{ and } \|\boldsymbol{p_k} - \boldsymbol{p}\| \le r \\ & \text{and } i * \Delta b_1 \le \rho_k < (i+1) * \Delta b_1 \\ & \text{and } j * \Delta b_2 \le d_k + r < (j+1) * \Delta b_2 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

In Eq. 3, Δb_1 and Δb_2 are the sizes of bins in the 2D histogram. Although we can notice that there is a missing degree of freedom which is the cylindrical angular parameter around unit normal with SSI representation scheme of free-form surface, each SSI uniquely defines the location of corresponding point on the surface using small sizes of bins, as shown in Fig. 1 (c). Obviously, the SSI is insensitive to pose variance.

The SSI representation scheme is inspired by *spin-image* [9,10] and *point signature* [2]. There are two main differences between SSI and spin-image. Firstly, an SSI of a point $\boldsymbol{p}$ uses points within the sphere centered at $\boldsymbol{p}$ on the surface to construct 2D histogram, while a spin-image of a certain point on the surface uses all points on the surface to construct 2D histogram. Generally, most 3D scanning methods can only capture part of the objects at a time [8]. This could be due to limitation in the physical design and the technology the scanner is built upon. Using these scanning methods parts of an object's surface may not be recoverable because the object itself shadows the structured light from the detector in the sensor. As shown in Fig. 2, each face surface contains regions that are invisible in the other. Thus, spin image constructed by all surface's points, including points in those invisible regions, will bring ambiguity during matching. Contrastively, SSIs of two corresponding points of different views will cover the same regions of the surface by using the sphere of some appropriate radius. Secondly, SSI employs
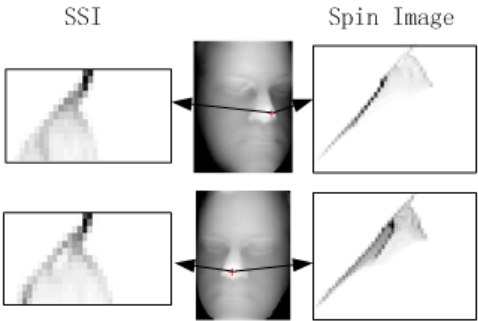
Fig. 2. Two depth images in different pose for the same person and their SSIs and spin images at the same location

the distance from other points to p as a dimension, instead of the distance from other points to the normal of point p as a dimension in spin-image. Hence, it is consistent with the use of sphere to define point's neighborhood.

2.2 SSI Generation

Prior to SSI comparison and face recognition with SSI, we present a scheme to derive the SSI representation from object models represented by triangular meshes.

$$n_p = \frac{n_1 * A_1 + n_2 * A_2 + n_3 * A_3 + n_4 * A_4}{A_1 + A_2 + A_3 + A_4}$$
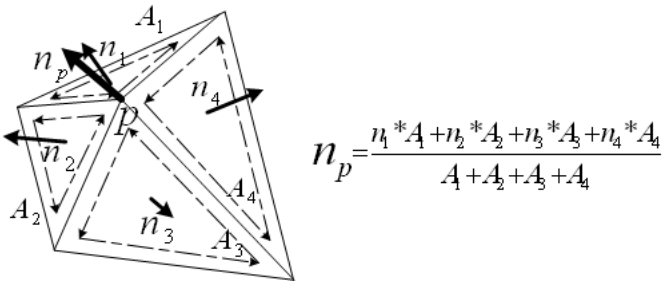
Fig. 3. Normal of vertex p calculation using the weighted means of triangles' normal adjacent to p. n_p is the normal of vertex p, n_i (i=1,2,3,4) is triangle's normal, A_i (i=1,2,3,4) is triangle's area.

Given a free-form surface consisting of triangular meshes, the normal of a vertex on the surface can be determined by the triangular patches adjacent to it. For SSIs of points on different models, the normals of vertices on different

models should be consistent, that is, all normals of vertices on different models must be oriented to the outside of the object. First of all, we sort edges of all triangular meshes of surface to be counter-clockwise viewed from one side of object. The consistent triangle's normal can be calculated based on the sorted edges, as shown in Fig. 3. Then normals of vertices whose orientation consistently extend to one side of the object is obtained by the weighted mean of triangle's normals adjacent to it. Finally, using the method proposed by Johnson [9], the orientation of all normals on the surface vertices is determined by calculating the scalar products of surface normal at each vertex and the vector from the centroid of the object to the vertex.

Before the construction of SSI, the radius r of sphere and Δb_1 and Δb_2 must be determined. In our experiments, the r is set as fifty times the mesh resolution and Δb_1 and Δb_2 are predetermined to be two to four times the mesh resolution which meets enough precision requirements. Then, the size of SSI, m and n defined in Eq.1 , can be easily computed by the equations:

$$m = \lfloor \frac{r}{\Delta b_1} \rfloor + 1, n = \lfloor 2 * \frac{r}{\Delta b_2} \rfloor + 1 \qquad (4)$$

where $\lfloor \ \rfloor$ is the floor operator.

Additionally, to deal with the different resolution of range images, we resample the data of all models with same method to construct triangular meshes and normalizing the SSI after the SSI construction. For normalization, we simply let the 2D histogram divided by the total number of those points in the sphere.

2.3   SSI Comparison

Since correlation coefficient between images is constantly used to compare the similarity of images, it is also used to SSIs matching in our experiments. The correlation coefficient between a given SSI S_r^1 and a candidate SSI S_r^2 is:

$$CorCoe(S_r^1, S_r^2) = \frac{Cov(S_r^1, S_r^2)}{\sqrt{D(S_r^1) * D(S_r^2)}} \qquad (5)$$

where

$$Cov(S_r^1, S_r^2) = \frac{1}{N} \sum_i S_r^1(i) S_r^2(i) - \frac{1}{N^2} \sum_i S_r^1(i) \sum_i S_r^2(i) \qquad (6)$$

and

$$D(S_r) = \frac{1}{N} \sum_i S_r(i)^2 - (\frac{1}{N} \sum_i S_r(i))^2 \qquad (7)$$

3   Face Recognition

Before face recognition with SSI, the model database must be built up. For each model, computing SSIs of every point on the surface has a very high time complexity. Selecting some special points which serve as landmarks of the object to compute SSIs can considerably reduce the cost while not block the performance of face recognition markedly. We select these points through two steps:

(a)   (b)   (c)
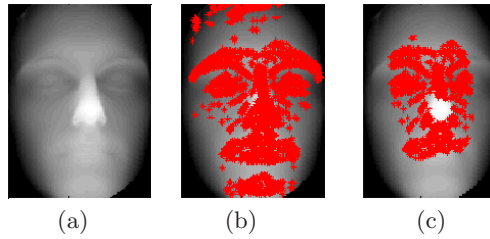
Fig. 4. (a) range image of a model. (b) Special points selection after step(1). (c) Special points selection after step(2). In (b) and (c), red parts denote the selected points.

- Firstly, ridge lines of a face surface are extracted by thresholding the minimum principal curvature of surface points, $k_{min} \leq threshold_c$. It is shown in Fig. 4(b).
- Secondly, from Fig. 4(b), it can be seen that ridge lines of a face surface contains many marginal points which are little help for face recognition but increase the computing complexity. Thus, we choose points both on ridge lines and in middle area of face surface as special point set P_{sp} of which SSIs are computed to construct model database. Marginal points are eliminated from ridge lines by counting the numbers of points within the point's related sphere and discarding the point whose related sphere contains points less than $threshold_s$. Fig. 4 (c) shows the set P_{sp}.

By the construction of SSIs of points in set P_{sp} for each model, we complete a model database in which each model has a SSI set S_{sp}. Face recognition is then implemented by voting method similar to[3]. For a given scene face, we want to select the most similar model in the database. Set P_{sp} and S_{sp} of scene face can be generated using the same construction method as models. For each SSI in S_{sp} of a scene face, correlation coefficients with every SSI in a model's S_{sp} in database are calculated. If the maximum of those correlation coefficients is larger than a $threshold_{CorCoe}$, the model receives a vote. Suppose the SSI number in set S_{sp} of the scene face is N_c and the model receives N_m votes, a voting rate associated to the model is N_m/N_c. The voting rate represents the likelihood of each model being correctly matched with the scene face. Using the voting method, the candidate models are ordered according to similarity.
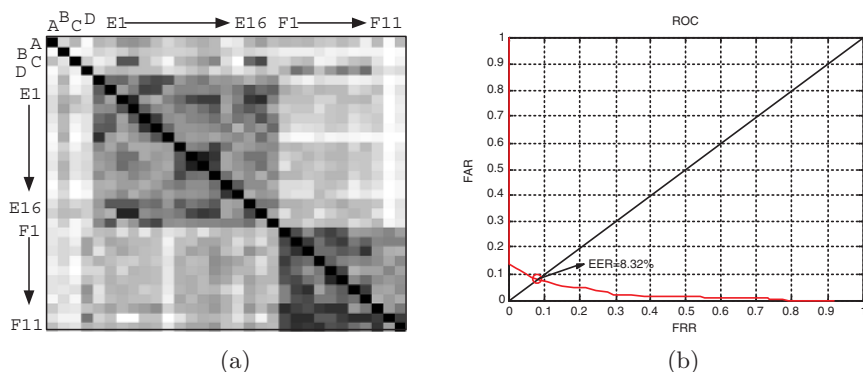
4   Experimental Results

In order to test the effectiveness of the presented approach, we perform several experiments with models obtained by Delaunay triangulation of range images downloaded from SAMPL(Signal Analysis and Machine Perception Laboratory) [7] at Ohio State University. The difference between the models from the range images of SAMPL and general model represented by polygon mesh is that the former have no points information in the regions occluded by itself when the

Fig. 5. Sample images from SAMPL. Six range images in model database. Notice the variance of the model pose.

model keeps certain special pose. Thus, the range images contain less 3D information. Our experiments involve 31 models of six human subjects and two human subjects have several views with significantly different poses. Fig. 5 shows six samples of the range images.

(a)   (b)

Fig. 6. (a) Similarity matrix for 31 models. Lightness indicates the dissimilarity between models. (b) ROC(Receiver Operator Characteristic) curve of recognition results. EER=8.32%.

In our experiments, we investigated the ability of SSI representation scheme and the voting method to recognition among models in the constructed model database. We calculated voting rates for all pairs of 31 models and the resulting voting rates are shown as a matrix in Fig. 6 (a). The lightness of each element(i,j) is proportional to the magnitude of the dissimilarity between models i and j. The darker represent better matches, while the lighter indicate worse matches. The symbols at the left and top side in Fig. 6 (a) denote different models where the same letter with different number means the models are of same human subject with different pose.

False reject rate(FRR) and false acceptance rate are quoted in many cases to evaluate recognition performance. A false acceptance rate(FAR) is the percentage of imposters wrongly matched, and a false rejection rate(FRR) is the

percentage of valid users wrongly rejected. Given a threshold, a FRR and a FAR can be calculated from the matrix. With different thresholds, a plot of numerous FAR and FRR combination, named receiver operator characteristic(ROC) curve, is obtained, shown in Fig. 6 (b). Equal Error Rate(EER) is the error rate when FRR equals FAR which can evaluate the performance of our recognition algorithm. With the significantly different pose models, EER in our experiment still maintains 8.32% which is a promising result.

5   Conclusion

This paper introduced a new free-form surface representation named SSI. An SSI of a given surface's point encodes neighborhood shape of the point through constructing a 2D histogram. It is done by mapping 3D coordinates of surface's points within a sphere centered at the given point into 2D space. The key feature of this approach is that it is viewpoint-invariant. Based on SSIs comparison of special points on models, face recognition experiments were performed by voting method. Although our model database contains relatively less face subjects due to the difficulty of obtaining the facial range images, significant pose variance in our 3D face database make our experimental results convincing.

References

1. W. Zhao, R. Chellappa, P. J. Phillips, A. Rosenfeld. Face recognition: a literature survey. ACM Computing Survey, 2003, 35(4): 399-458
2. C. S. Chua, R. Jarvis. Point signatures: A new representation for 3D object recognition. *IJCV*, 1997, 25(1): 63-85
3. C. S. Chua, F. Han, Y. K. Ho. 3D Human Face Recognition Using Point Signature. *Proc. of Int'l Conf. on Automatic Face and Gesture Recognition*, 2000: 233-238
4. R. Osada, T. Funkhouser, B. chazelle et al. Shape distributions. *ACM Transactions on Graphics*, 2002, 21(4): 807-832.
5. H. T. Tanaka, M. Ikeda, H. Chiaki. Curvature-based Face Surface Recognition Using Spherical Correlation-Principal Directions for Curved Object Recognition-. *Proc. of Int'l Conf. on Automatic Face and Gesture Recognition*, 1998: 372-377
6. G. G. Gordon. Face recognition from depth maps and surface curvature . *Proc. of SPIE Conf. on Geometric Methods in Computer Vision*, 1991, 1570: 234-247.
7. Range Imagery, SAMPL (Signal Analysis and Machine perception Laboratory at the Ohio State University). [Online]. Available: http://sample.eng.ohio-state.edu
8. R. J. Campbell, P.J.Flynn. A Survey Of Free-Form Object Representation and Recognition Techniques. *CVIU*, 2001, 81(2): 166-210
9. A. E. Johnson, M.Hebert. Surface matching for object recognition in complex three-dimensional scenes. *Image Vision Computing*, 1998, 16: 635-651.
10. A. E. Johnson, M.Hebert. Using spin images for efficient object recognition in cluttered 3D scenes.*IEEE Trans. PAMI*.1999, 21(5): 433-449
11. S. M. Yamany, A.A.Farag. Surface signatures: an orientation independent free-form surface representation scheme for the purpose of objects registration and matching. *IEEE Trans. PAMI*. 2002, 24(8): 1105-1120
12. C. Dorai, A.K.Jain. COSMOS-A representation scheme for 3D free-form objects. *IEEE Trans. PAMI*, 1997, 19: 1115-1130