Integrating Range and Texture Information for 3D Face Recognition

Xiaoguang Lu and Anil K. Jain
Dept. of Computer Science & Engineering
Michigan State University
East Lansing, MI 48824
{Lvxiaogu, jain}@cse.msu.edu

Abstract

The performance of face recognition systems that use two-dimensional images depends on consistent conditions w.r.t. lighting, pose, and facial appearance. We are developing a face recognition system that utilizes three-dimensional shape information to make the system more robust to arbitrary view, lighting, and facial appearance. For each subject, a 3D face model is constructed by integrating several 2.5D face scans from different viewpoints. A 2.5D scan is composed of one range image along with a registered 2D color image. The recognition engine consists of two components, surface matching and appearance-based matching. The surface matching component is based on a modified Iterative Closest Point (ICP) algorithm. The candidate list used for appearance matching is dynamically generated based on the output of the surface matching component, which reduces the complexity of the appearance-based matching stage. The 3D model in the gallery is used to synthesize new appearance samples with pose and illumination variations that are used for discriminant subspace analysis. The weighted sum rule is applied to combine the two matching components. A hierarchical matching structure is designed to further improve the system performance in both accuracy and efficiency. Experimental results are given for matching a database of 100 3D face models with 598 2.5D independent test scans acquired in different pose and lighting conditions, and with some smiling expression. The results show the feasibility of the proposed matching scheme.

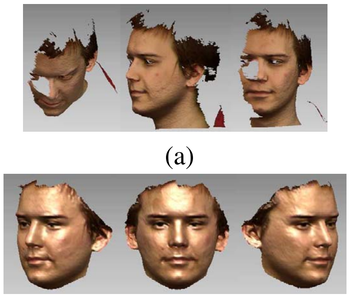
1. Introduction

Automatic human face recognition is a challenging task that has gained a lot of attention during the last decade [26]. While most efforts have been devoted to face recognition from two-dimensional (2D) images [26], a few approaches have utilized depth information provided by 2.5D range images [18, 16, 23, 12, 8, 22, 19]. Current 2D face recognition systems can achieve good performance in constrained environments, however, they still encounter difficulties in handling large amounts of facial variations due to head pose, lighting conditions and facial expressions [2]. Because the human face is a three-dimensional (3D) object whose 2D projection (image or appearance) is sensitive to the above changes, utilizing 3D facial information should improve the face recognition performance [9, 2]. Range images captured explicitly by a 3D sensor [1, 4] contain facial surface shape information. The 3D shape of facial surface represents the facial structure, which is related to the internal anatomical structure instead of external appearance and environment. It is also more difficult to fake a 3D image compared to a 2D image to fool the face recognition system.

We use 3D models to recognize 2.5D face scans, provided by a 3D sensor, such as the Minolta Vivid series [4]. A 2.5D scan is a simplified 3D (x, y, z) surface representation that contains at most one depth value (z direction) for every point in the (x, y) plane (see Figure 1), along with a registered color image. Each scan can only provide a single view point of the object, instead of the full 3D view. As the 3D imaging technology is progressing quickly [5], non-intrusive 3D data capture along with texture information will become readily available. In real world scenarios, similar to the current 2D camera capture systems, 3D sensors provide only partial views of the face. However, during the training stage, 3D face model can be constructed by taking several scans from different viewpoints. Therefore, we address the scenario that matches a 2.5D facial scan to 3D models.

Face recognition based on range images has been addressed in a number of different ways. Lee and Milios [18] segmented the range image to obtain the convex regions, which correspond to distinct facial features. The Extended Gaussian Image (EGI) is used to represent each convex region. A similarity metric between two regions is defined to match the facial features of the two images. Gordon [16] explored the face feature extraction for recognition based on depth and curvature features. Tanaka et al. [23] consid-

Proceedings of the Seventh IEEE Workshop on Applications of Computer Vision (WACV/MOTION'05)
0-7695-2271-8/05 $ 20.00 IEEE


(a)

(b)

Figure 1. Range scan and 3D face model. (a) One profile range scan viewed at different viewpoints; (b) a full 3D model.

ered the face recognition problem as a 3D shape recognition problem involving rigid free-form surfaces. Their method is based on the curvature information. Chua et al. [12] extended the use of Point Signature to recognize frontal face scans with different expressions, which was treated as a 3D recognition problem of non-rigid surfaces. Beumier and Acheroy [8] extracted the profiles (curves) both from depth and gray scale image for face verification. Pan et al. [22] utilized the partial directed Hausdorff distance to align and match two range images for verification. Work by Chang et al. [10] demonstrated that improvements can be made if a system uses a combination of texture and shape information. They applied PCA to both 2D and 3D face data.

While different methods have been used to address face recognition based on range images, most of them have focused on the frontal view face recognition. Further, most of these methods only use the shape information. But the texture component also plays an important role in face recognition process, especially when the shapes of two faces in the gallery are similar. Although the 3D shape will not change due to pose and lighting variations, it can still change due to expression and the aging factor. Therefore, using 3D shape information alone can not yet fully handle all the variations which the face recognition system encounters.

We propose a combination scheme, which integrates surface (shape) matching and constrained appearance-based methods for multi-view face matching (see Fig. 2). The appearance-based stage is constrained to a small candidate list generated by the surface matching stage, which reduces the classification complexity. In the conventional appearance-based algorithm, all the subjects in the training database are used for subspace analysis and construction. When the number of subjects is large, this leads to a problem with high complexity. In our scheme, the 3D model is utilized to synthesize training samples with facial appearance variations, which are used for discriminant subspace analysis. The scores obtained by the two matching components are combined to make the final decision. Further, a hierarchical matching structure is designed to improve the system performance in terms of both accuracy and efficiency.

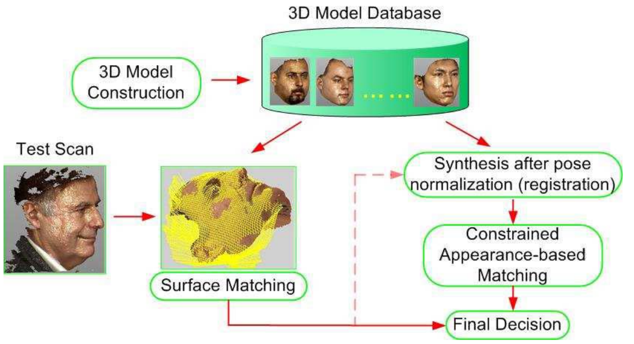

3D Model Database
3D Model Construction
Test Scan
Synthesis after pose normalization (registration)
Constrained Appearance-based Matching
Surface Matching
Final Decision

Figure 2. Face recognition based on combination of shape and appearance-based features.

2. 3D Model Construction

The 3D face model for each subject is constructed by stitching several 2.5D scans obtained from different view points. The scans were stitched together using a commercial software, called Geomagic Studio [3]. In our current setup, 5 scans are used, i.e., frontal, left 30 degrees, left 60 degrees, right 30 degrees and right 60 degrees. The 2.5 scans are first registered with each other in the same coordinate system and then merged to create a surface model. Hole-filling and basic clean-up procedures are applied to smooth the surface and remove noisy points associated with hair and clothing. The end result is a smooth full view of the face for each subject. Figure 3 demonstrates the 3D face model construction procedure. The resulting model is highly dense, containing ∼27,000 vertices and ∼50,000 polygons. It can be used to render new realistic facial appearance with pose and illumination variations.

3. Surface Matching

In order to match two facial surfaces (2.5D test scan and 3D model), we follow the coarse-to-fine strategy (see Fig. 4).

3.1. Coarse Alignment

We applied a feature based alignment for coarse registration for its simplicity and efficiency. A minimum of three corresponding points is needed in order to calculate the rigid

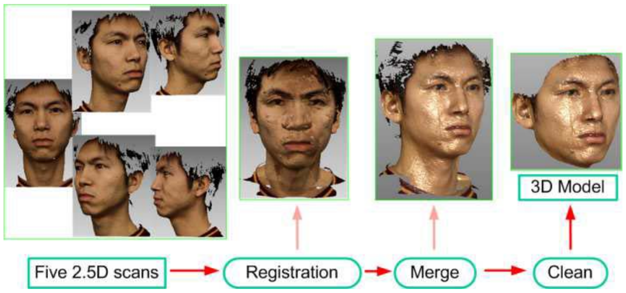


Figure 3. 3D model construction.

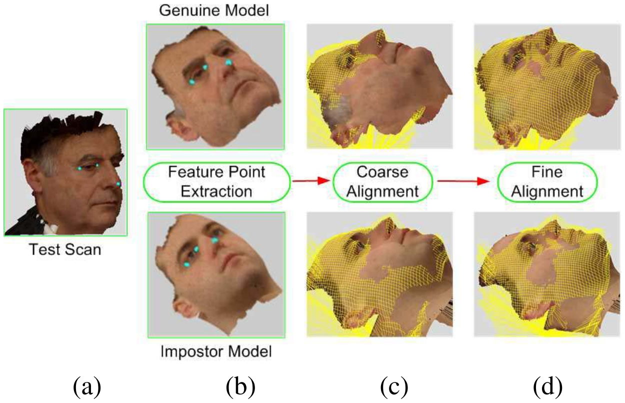


Figure 4. Surface matching streamline. The alignment results are shown by the 3D model overlaid on the wire-frame of the 2.5D test scan.

transformation between two sets of 3D points. Once the three corresponding points (anchor points) are known [1], the transformation is made using a combination of rigid transformation matrices following the guidelines described in [24]. This is done by a least squares fitting between the triangles formed from the two sets of three anchor points. We pick a combination of the inside of one eye, the outside of that eye and the nose tip as our three anchor points. See Fig. 5 for examples. These points are selected because they are relatively easy to locate in the range image and they do not change between different scans of different people across different poses. See Fig. 4(c) for an example of a 2.5D face scan coarsely aligned to a 3D face mesh model.

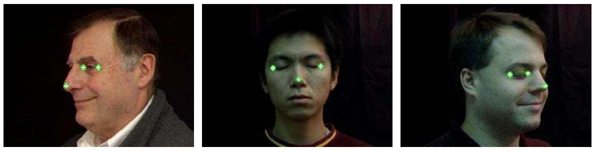


Figure 5. Anchor point labeling based on pose: left-profile, frontal and right-profile.

3.2. Fine Alignment

Our fine registration process follows the Iterative Closest Point (ICP) framework [7, 11, 25] to align two sets of control points. The basic Iterative Closest Point scheme is described as follows:

1. Select control points in one point set

2. Find the closest points in the other point set (correspondence)

3. Calculate the optimal transformation between the two sets based on the current correspondence

4. Transform the points; repeat step 2, until convergence.

Starting with an initial estimate of the rigid transformation, ICP iteratively refines the transform by alternately choosing corresponding (control) points in the 3D model and the 2.5D scan and finding the best translation and rotation that minimizes an error function based on the distance between them.

Besl [7] used point-to-point distance and a close-form solution when calculating the transformation matrix during each iteration. The point-to-plane distance used in [11] makes the ICP algorithm less susceptible to local minima than the point-to-point metric [15]. It also needs a fewer number of iterations to converge. But point-to-plane distance based ICP has to solve a non-linear optimization problem using numerical algorithms. We integrate the two classical ICP algorithms [7, 11] in a zigzag running style, and call it the hybrid ICP algorithm. Each iteration consists of two steps, using Besl's scheme to compute an estimation of the alignment, followed by Chen's scheme for a refinement. The two different distance metrics are utilized together, which leads to a better registration than the individual metrics.

Figure 6 shows the grids used for control point selection for various poses. Regions around the eyes and nose were selected because these regions are less malleable than other parts of the face (such as the region around the mouth, which changes greatly with facial expression.) The fine alignment results are demonstrated in Fig. 4(d).

3.3. Surface Matching Distance

The root mean square distance minimized by the ICP algorithm is used as the primary matching distance of face

[1] In order to evaluate the matching scheme, we study the feature extraction and matching components separately. The coarse alignment is currently performed using manually picked anchor points. Our scheme for automatic feature extraction is described in [13], which can extract anchor points with about 98% accuracy on frontal face scans and 80% on profile face scans.

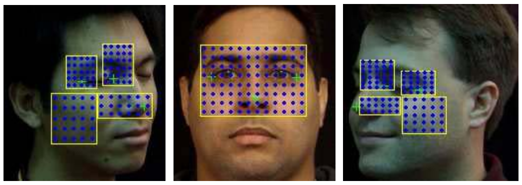


Figure 6. Automatic control point selection, based on three anchor points, for a left profile, frontal, and right profile scan. (About 100 control points are selected in each scan).

scans. We use the point-to-plane distance metric MD_{ICP} defined in [11].

$$MD_{ICP} = \sqrt{\frac{1}{N_c}\sum_{i=1}^{N_c} d^2(\Psi(p_i), S_i)}, \qquad (1)$$

where $d(\cdot)$ is the point-to-plane metric; $\Psi(\cdot)$ is the rigid transformation applied to each control point p_i in the 2.5D test scan; S_i is the corresponding tangent plane in the 3D model w.r.t. p_i; N_c is the number of control points. The smaller the MD_{ICP}, the better the surface matching.

4. Constrained Appearance-based Matching

The appearance-based algorithm requires the training and test samples to be aligned. In our approach, the test scan and the 3D model are aligned by the ICP registration procedure, i.e., the pose is normalized. By synthesizing new appearance from the constructed 3D model, additional appearance-based training samples become available. We applied the linear discriminant analysis (LDA) based method for appearance-based matching [6, 20]. Instead of using all the subjects in the database, the LDA is applied only to a small list of candidates, generated dynamically by the surface matching stage for each test scan. We call this as the constrained appearance-based matching.

4.1. Appearance Synthesis

Each subject only has one face model with neutral expression in the database. In order to apply the subspace analysis based on the facial appearance, many training samples, which are aligned with the test sample, are needed [6, 20]. After the surface registration (pose normalization), the 3D model is aligned with the test scan and so it is easy to synthesize new appearance with lighting variations. As the alignment may not be perfect, small pose variations are also synthesized in our framework. In principle, the number of available synthesized samples can be arbitrarily large.

Pose variation synthesis is straightforward by simply rotating and shifting the 3D model. Lighting is simulated by adding a virtual light source around the reconstructed face surface. Different illumination variations are generated by changing the position of the light source. Phong shading technique is employed to render lighting effects on the face surface [14].

Based on the anchor points (eye corners and the nose tip) and registration results, the critical area in the face is determined, which is used to automatically crop the synthesized images. Examples of the cropped synthesized images are shown in Fig. 7.

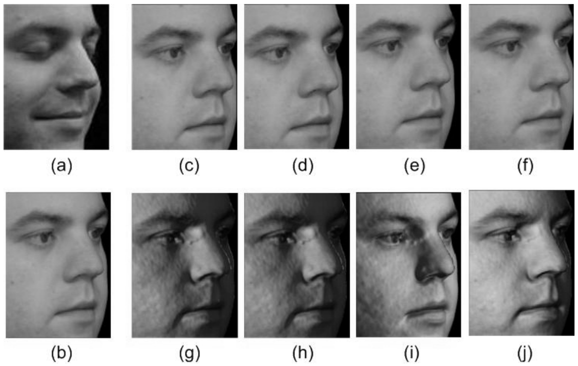


Figure 7. Cropped synthesized training samples for discriminant subspace analysis. (a) 2.5D test (scan) image; (b) 3D model after pose normalization (alignment); (c-f) synthesized images of (b) with shift displacement in horizontal and vertical directions; (g-j) synthesized images with lighting changes.

4.2. Dynamic Candidate Selection in LDA

In the conventional LDA, all the subjects in the training database (gallery) are used for subspace construction. When the number of subjects is large, the complexity of the recognition problem is increased due to large intra-class variations and large inter-class similarities, resulting in low recognition accuracy. Therefore, in our approach, for each test scan, the gallery used for subspace analysis and matching is dynamically generated based on the output of the surface matching. Only a small fraction of the subjects in the database is selected for the appearance-based matching, so that the number of subjects to be matched to the test scan is small. In our experiments, the top M candidates in the sorted matching list are selected ($M = 30$).

5. Integration

5.1. Weighted Sum Rule

Surface matching and appearance-based matching provide two scores based on different cues. Each of them can be considered as a classifier. Since they explore different properties of the face, namely, shape and texture, these two classifiers are not highly correlated. A combination of these two classifiers has the potential to outperform each individual classifier [17]. We applied the weighted sum rule to integrate the surface matching and appearance-based matching distances as follows:

$$MD_{comb} = MD_{ICP} + \alpha \cdot MD_{LDA}, \qquad (2)$$

where $MD_{LDA} = (1 - MS_{LDA})/2$, MS_{LDA} is the matching score generated by the appearance-based matching component, converting the matching score (similarity) to matching distance (dissimilarity). The weighting parameter α balances the two matching components, which can be set beforehand or learned from an independent validation dataset.

5.2. Hierarchical Matching

The surface matching in Section 3 focused on the face region that is more robust to deformation due to expression changes. We call it the 'local' scheme. But to solve the ambiguity between shapes, larger facial area may provide more evidence, especially for the faces with the same expression as that of the 3D models (neutral expression in our experiments). Therefore, a hierarchical matching framework is designed, where a 'global' surface matching component is introduced. Figure 8 illustrates the hierarchical system and Fig. 9 shows the global control point sampling scheme . Only those test scans for which the surface matching component does not have sufficient evidence to make the decision, are fed to the combination stage. This cascading framework also provides the potential to reduce the total computation cost. In our current implementation, if the shape matching distance (MD_{ICP} in Eq. (1)) is below a pre-defined threshold δ, then it is considered as a good surface matching. Since the surface matching distance is measured by the root mean square distance among the control points, it has a physical meaning. We choose δ equal to 1 in units of millimeters. The value of δ depends on the noise level of the scans and the performance of the anchor point locator.

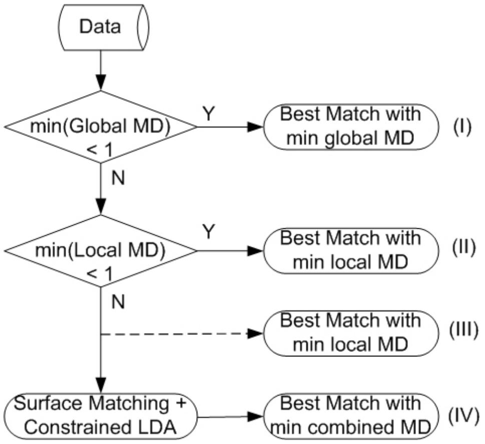

Data
min(Global MD) < 1
Y
N
Best Match with min global MD
(I)
min(Local MD) < 1
Y
N
Best Match with min local MD
(II)
Best Match with min local MD
(III)
Surface Matching + Constrained LDA
Best Match with min combined MD
(IV)

Figure 8. Hierarchical matching design.

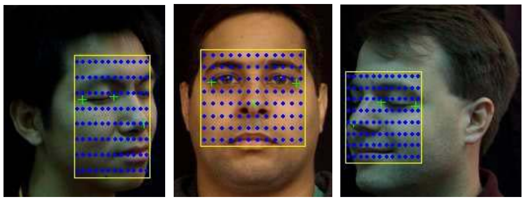


Figure 9. Global control point sampling based on three anchor points, for a left profile, frontal, and right profile scans.

6. Experiments and Discussion

6.1. Data

There is no publicly available multi-view range (with registered texture) face database, along with expression variations. All range images (downsampled to 320×240 with a depth resolution of $\sim 0.1mm$) were collected using a Minolta Vivid 910 scanner [4] in our laboratory. This scanner uses structured laser light to construct the face image in less than a second. Each point in a scan has a texture color (r, g, b) as well as a location in 3D space (x, y, z). Each facial scan has around $18,000$ effective points (excluding the background).

There are currently 100 subjects in our database. Five scans with neutral expression for each subject were captured to construct the 3D model. For each subject, another six scans are captured for testing, including 3 scans with neutral expression and 3 with smiling expression. For a few subjects fewer than 6 test scans are available. So, there are a total of 100 3D models stored in the gallery database and 598 independent 2.5D scans for testing. The representative 3D models and test scans are shown in Fig. 10 and Fig. 11, respectively.

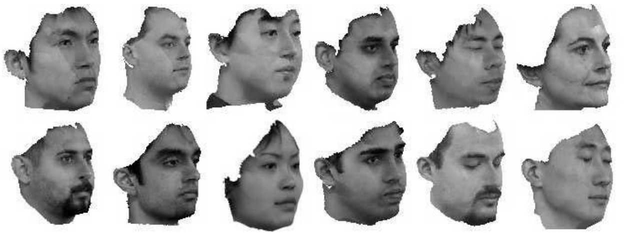


Figure 10. Some of the 3D face models.

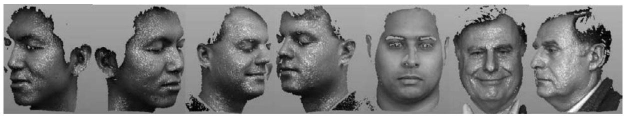


Figure 11. Representative 2.5D test scans.

6.2. Surface Matching

In order to test the matching scheme, the three anchor points (eye corners and the nose tip) are manually labeled in the current implementation. (Our scheme for automatic feature extraction is described in [13].) Coarse alignment finds the rotation and translation parameters to align the two triangles (built by the anchor points) from the test scan and the 3D model. Based on the anchor points, control points are automatically sampled for the ICP registration. Figure 6 shows the control point sampling scheme. Examples of the registration results are given in Figs. 4(c) and 4(d).

Our matching process is conducted in the identification mode. Each test scan is matched to all the 3D models stored in the gallery. The surface matching distance distributions for genuine users and impostors are provided in Fig. 12.

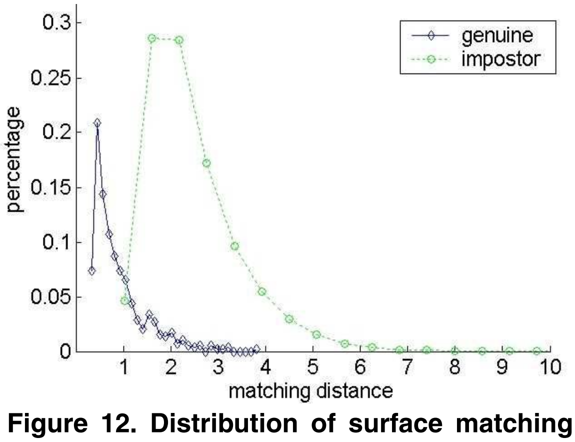


Figure 12. Distribution of surface matching distance.

6.3. Combination of Surface and Appearance-based Matching

In the constrained appearance-based matching, 4 images with different shift displacements and 4 images with different lighting conditions are synthesized. Hence, 9 images for each model are used for the LDA calculation (8 synthesized version plus the original one, see Figs. 7(b)-(j) for an example). The LDA is only applied to the first 30 matched candidates based on the surface matching distance. By applying surface matching and constrained appearance-based schemes separately to this set, we found that the sets of misclassified test scans are significantly different for these two matching schemes, implying that these two schemes are not highly correlated.

A summary of the experimental results is given in Table 1. Out of the 54 errors in 598 test scans (corresponding to 91% accuracy), 49 scans are with smiling expression. So, almost all the errors are due to expression changes that nonlinearly deform the 3D shape of the test scan. Figure 13 shows some of the test scans that are correctly classified. The rank-one matching accuracy for 312 neutral expression test scans (frontal and non-frontal) is $\sim 98\%$. The cumulative match score curves for the three different matching schemes are provided in Fig. 14.

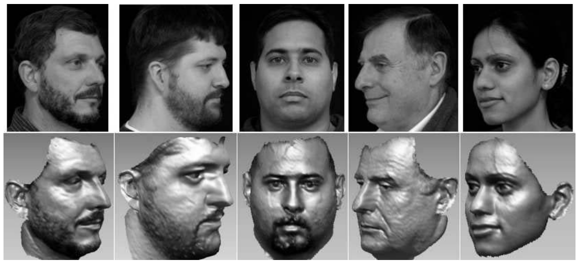


Figure 13. Test scans (top row), and the corresponding 3D models (bottom row) correctly matched. The 3D model is shown roughly in the same pose as the corresponding test scan.

The performance change with respect to α is shown in Fig. 15. In practice, the 'optimal' value of α can be learned from the validation data.

6.4. Hierarchical Matching

In order to explore additional shape information contained in the facial area, especially for those scans with the same expression as that of the 3D models, a hierarchical matching framework from global to local is designed as shown in Fig. 8. Fig. 16 provides detailed experimen-

Matching scheme	Rank-one accuracy (598 test scans)	Rank-one accuracy (312 neutral expression test scans)
Surface matching (ICP only)	87%	97%
Appearance-based (LDA only)	77%	84%
Surface matching + Appearance-based	91%	98%

Table 1. Matching accuracy with equal weights for ICP and LDA components (i.e., $\alpha = 1$ in Eq. (2)).

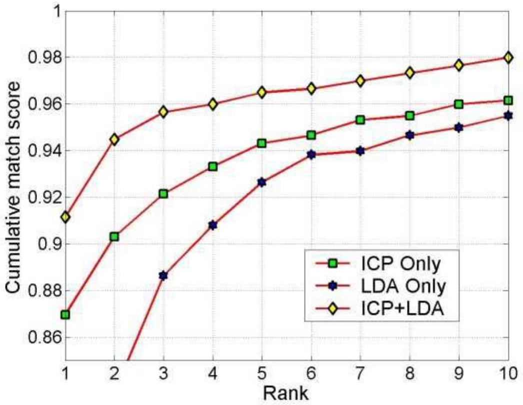


Figure 14. Cumulative matching performance.

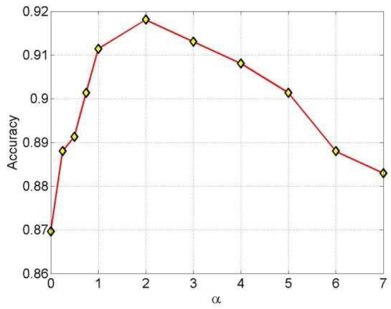


Figure 15. Identification accuracy based on the combination strategy with respect to α.

tal results when applying the proposed framework to our database. The comparative study is shown in Table 2, which demonstrates that the hierarchical matching scheme slightly improves the system recognition accuracy. Furthermore, for our database, due to the introduction of the hierarchical structure, 432 out of 598 test samples are not fed into the appearance-based matching stage, which reduces the computation cost with a good tradeoff in the accuracy (only 3 errors with smiling expression are generated at this stage). The remaining 166 test scans, for whom the surface matching component does not have sufficient evidence to make the decision, are fed to the next combination stage.

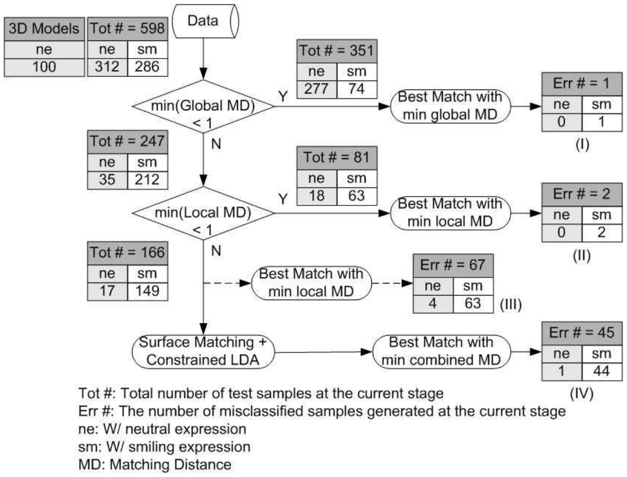


Figure 16. Hierarchical matching scheme ($\alpha = 1$).

Scheme	w/o hierarchical structure	w/ hierarchical structure
Surface matching	87%	88%
Surface matching + Appearance-based	91%	92%

Table 2. Rank-one matching accuracy ($\alpha = 1$) with and without hierarchical structure.

7. Conclusions and Future Work

We have presented a face recognition system that matches 2.5D scans of faces with pose, lighting and expression variations to a database of 3D models. A combination scheme is proposed, which integrates surface (shape) matching and a constrained appearance-based method. The surface matching is achieved by a hybrid ICP scheme. The appearance-based identification component is constrained to a small candidate list generated by the surface matching component, which reduces the classification complexity. The registered 3D model is utilized to synthesize training samples with facial appearance variations, which are used for discriminant subspace analysis. The scores obtained by the two matching components are combined using the weighted sum rule to make the final decision. Given the anchor points, the entire matching scheme is fully automatic, including surface registration/matching, dynamic candidate list selection, 3D model-based synthesis, sample image cropping, LDA, and appearance-based matching. In our current implementation, matching one 2.5D test scan to a 3D model takes about 30 seconds. Fast algorithms are being pursued to improve the speed. A hierarchical matching framework is designed to further improve the system performance in both accuracy and efficiency.

This research is an encouraging first-step in designing a system that is capable of recognizing faces with arbitrary pose and illumination. More sophisticated surface matching schemes are being pursued to improve the surface matching accuracy, including exploring models that can be deformed to deal with non-rigid variations [21], which are caused by changes in expression and aging effects. To make the whole matching system fully automatic, a robust and accurate anchor point locator is being developed.

References

[1] *Cyberware Inc.* <http://www.cyberware.com/>.

[2] *Face Recognition Vendor Test (FRVT).* <http://www.frvt.org/>.

[3] *Geomagic Studio*. <http://www.geomagic.com/products/studio/>.

[4] *Minolta Vivid 910 non-contact 3D laser scanner.* <http://www.minoltausa.com/vivid/>.

[5] *The 4th International Conference on 3-D Digital Imaging and Modeling (3DIM).* <http://www.3dimconference.org/>, 2003.

[6] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 19(7):711–720, Jul. 1997.

[7] P. Besl and N. McKay. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[8] C. Beumier and M. Acheroy. Automatic 3D face authentication. *Image and Vision Computing*, 18(4):315–321, 2000.

[9] V. Blanz and T. Vetter. Face recognition based on fitting a 3D morphable model. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 25(9):1063–1074, 2003.

[10] K. I. Chang, K. W. Bowyer, and P. J. Flynn. Multi-modal 2D and 3D biometrics for face recognition. In *Proc. IEEE Workshop on Analysis and Modeling of Faces and Gestures*, France, Oct. 2003.

[11] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[12] C. Chua, F. Han, and Y. Ho. 3D human face recognition using point signature. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pages 233–238, Grenoble, Mar. 2000.

[13] D. Colbry, X. Lu, A. Jain, and G. Stockman. 3D face feature extraction for recognition. Technical Report MSU-CSE-04-39, Computer Science and Engineering, Michigan State University, East Lansing, Michigan, September 2004.

[14] J. Foley, A. van Dam, S. Feiner, and J. Hughes. *Computer Graphics: Principles and Practice*. Addison-Wesley, New York, 2nd edition, 1996.

[15] N. Gelfand, L. Ikemoto, S. Rusinkiewicz, and M. Levoy. Geometrically stable sampling for the icp algorithm. In *Proc. International Conference on 3D Digital Imaging and Modeling*, Banff, October 2003.

[16] G. Gordon. Face recognition based on depth and curvature features. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 108–110, 1992.

[17] J. Kittler, M. Hatef, R. Duin, and J. Matas. On combining classifiers. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 20(3):226–239, 1998.

[18] J. Lee and E. Milios. Matching range images of human faces. In *Proc. International Conference on Computer Vision*, pages 722–726, 1990.

[19] X. Lu, D. Colbry, and A. Jain. Three-dimensional model based face recognition. In *Proc. International Conference on Pattern Recognition*, pages 362–366, Cambridge, UK, 2004.

[20] A. Martinez and A. Kak. PCA versus LDA. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 23(2):228–233, Feb. 2001.

[21] I. Matthews and S. Baker. Active appearance models revisited. *International Journal of Computer Vision*, 60(2):135–164, 2004.

[22] G. Pan, Z. Wu, and Y. Pan. Automatic 3D face verification from range data. In *Proc. ICASSP*, volume 3, pages 193–196, 2003.

[23] H. Tanaka, M. Ikeda, and H. Chiaki. Curvature-based face surface recognition using spherical correlation. In *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pages 372–377, 1998.

[24] D. M. Weinstein. The analytic 3-D transform for the least-squared fit of three pairs of corresponding points. *School of Computing Technical Report, No. UUCS-98-005, University of Utah*, March 1998.

[25] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(1):119–152, 1994.

[26] W. Zhao, R. Chellappa, A. Rosenfeld, and P. Phillips. Face recognition: A literature survey. *CVL Technical Report, University of Maryland*, Oct. 2000. <ftp://ftp.cfar.umd.edu/TRs/CVL-Reports-2000/TR4167-zhao.ps.gz>.