Alternative face models for 3D face registration

Albert Ali Salah, Neşe Alyüz, Lale Akarun

Boğaziçi University, 34342 Bebek, İstanbul, Turkey

ABSTRACT

3D has become an important modality for face biometrics. The accuracy of a 3D face recognition system depends on a correct registration that aligns the facial surfaces and makes a comparison possible. The best results obtained so far use a one-to-all registration approach, which means each new facial surface is registered to all faces in the gallery, at a great computational cost. We explore the approach of registering the new facial surface to an average face model (AFM), which automatically establishes correspondence to the pre-registered gallery faces. Going one step further, we propose that using a couple of well-selected AFMs can trade-off computation time with accuracy. Drawing on cognitive justifications, we propose to employ category-specific alternative average face models for registration, which is shown to increase the accuracy of the subsequent recognition. We inspect thin-plate spline (TPS) and iterative closest point (ICP) based registration schemes under realistic assumptions on manual or automatic landmark detection prior to registration. We evaluate several approaches for the coarse initialization of ICP. We propose a new algorithm for constructing an AFM, and show that it works better than a recent approach. Finally, we perform simulations with multiple AFMs that correspond to different clusters in the face shape space and compare these with gender and morphology based groupings. We report our results on the FRGC 3D face database.

Keywords: 3D face recognition, registration, average face model, ICP, TPS, Procrustes analysis, automatic landmark localization, other race effect

1. INTRODUCTION

With the advances in acquisition hardware and 3D recognition algorithms, 3D face recognition has become an important biometric recognition/verification technique. The speed and accuracy bottleneck in face recognition is in the registration step. Test faces need to be aligned to the gallery faces for comparison. In 3D face classification applications, the most frequently employed[1,2,3,4,5] registration approach is the Iterative Closest Point (ICP) algorithm, which establishes a dense correspondence between two point clouds in a rigid manner. Typically, a test face is registered to each gallery face separately,[1,3,4,5] and a point set distance is adopted for classification. In the work of İrfanoğlu, an alternative and fast method was proposed to register faces.[2] An average face model (AFM) was employed to determine a single point-to-point correspondence. The gallery faces, previously used in generating the AFM, are already in dense correspondence with it. Thus, a single ICP is enough to register a test face, which is much faster for a reasonably sized gallery. This overcomes what has been reported as the major drawback of ICP,[6] at a cost: Since the registration with an AFM is (hypothetically) poorer than one-to-all registration, it is expected that the method should suffer in terms of verification accuracy.

In this paper, we inspect AFM-based registration more closely. We propose an approach that uses one AFM for each facial category, and a novel AFM generation method that aims at facilitating the subsequent classification (Section 2). The classifiers that are tuned for within-category differences stand to benefit from such a two-tiered architecture. What constitutes a facial category is an open issue; we contrast an approach based on cognitive justifications (detailed in Section 3) with one that is based on clustering on the shape space (Section 4).

The recognition engine employed in this paper is primarily used to assess the effect of AFM on registration. A point set difference (PSD) measure is adopted for nearest neighbour classification after registration. We

Further author information: (Send correspondence to A.A.S.)
A.A.S.: E-mail: salah@boun.edu.tr, Telephone: +90 212 359 4490
N.A.: E-mail: nese.alyuz@boun.edu.tr, Telephone: +90 212 359 4490
L.A.: E-mail: akarun@boun.edu.tr, Telephone: +90 212 359 6858

evaluate the quality of the AFM under rigid (ICP) and non-rigid registration methods. The coarse registration in ICP and the non-rigid thin-plate spline based registration both require a couple of fiducial points for guidance. We evaluate the effect of errors in landmark detection by using 3D ground-truth versus automatically located landmarks. This permits us to analyze the algorithms under realistic assumptions, as automatic landmarking errors are not uniformly distributed. Although we did not primarily attempt a direct comparison, we hope that our results will provide some insights into registration methods for 3D face recognition.

2. AVERAGE FACE MODEL CONSTRUCTION

In dense registration, the points on the test surface and the points on the gallery surface are put into one-to-one correspondence. ICP achieves this by iteratively locating the closest point on the test surface for each point on the gallery surface, and rigidly moving the aligned surface to minimize the total point-to-point distances.[7] Upon convergence, the distances between the points can be summed up to find a total distance to the gallery face. Usually the gallery face is cropped and cleansed from all clutter, and the number of correspondences equals the number of points on the gallery surface.

In the thin-plate spline (TPS) based non-rigid registration method used in this paper, seven landmarks are identified on the test face, and these drive the registration.[8] The TPS method describes a transformation that aligns the landmarks on the test face with the landmarks on the gallery face exactly, and all other points are interpolated. This method is much faster than the ICP alignment.

The registration procedures we consider here require an average face model (AFM). The idea in AFM-based registration is to register all gallery surfaces to a single AFM, which acts as an index file. The test surface is registered once and for all with the AFM, which associates one point on the test surface for each point of the AFM, and consequently, for each point of any given gallery surface.

In Ref. 2, a method for generating the AFM was described, which involves a TPS-based registration of the training faces to a consensus shape. Then, one of the faces is selected as the AFM candidate, and its vertices are trimmed if their distance to an other training image exceeds a threshold. This procedure creates a very smooth facial surface.

Using a set of landmarked training faces, we generate a more pronounced AFM with the following procedure:

- Using Procrustes analysis,[9] a consensus distribution of landmarks (`consensus shape`) is found on the training set.
- The landmarks of the consensus shape are rectified to present a fully frontal face. This heuristic is used to facilitate the use of the transformed range image in later stages. Rectification is achieved by rotating the face so that the eye and mouth planes are parallel to the x-axis and the z-axis as much as possible.
- TPS deformation is computed for the training faces, which warps the landmarks of each face to the consensus shape perfectly, and interpolates the rest of the points.
- The depth values of the interpolated face are resampled from a regular x-y grid. This ensures that all added faces have points with overlapping x and y values, and the depth values are given for matching points. For the simple range image representation, this extra offline computation leads to much faster online model comparison.
- Faces are cropped before they are added to the average face model. The first face is used to define a cropping mask used for the rest of the faces. First we calculate the maximum distance from the nose tip to any landmark in the consensus shape. We add a ten per cent margin to this distance, and retain all points closer than this value to the nose.
- After all the training faces are added, depth values are simply averaged.

Samples of AFMs generated with this method can be seen in Figure 2. Any irregularity in the surfaces is due to poor pre-processing of the depth data. The database we use in this study was collected with a laser sensor that typically generates holes (especially at the eyes and the mouth) or other artifacts.[10] The pre-processing for these files (9×9 median filtering, followed by 9×9 mean filtering, followed by polynomial interpolation of missing points at each direction) sometimes falls short of repairing larger errors.

3. COGNITIVE JUSTIFICATION FOR MULTIPLE AFMS

When humans see faces, they perceive each face individually, and not merely as specimens of a generic object category. The mechanism that operates for enhanced within-category recognition is what we can call *expertise* of that category. Tong et al. remark that expertise-related visual judgements involve enhanced within-category discrimination, as opposed to between-category discrimination, the former requiring a magnification of differences in similar objects, and the latter calling for a neglecting of differences to group similar items together.[11]

People have great difficulty recognizing faces of another race if they are not exposed to them for prolonged periods. This phenomenon is termed the *other race effect*. In the light of Tong et al.'s experiments, it seems reasonable that during the acquisition of face expertise, the transformations learned by the visual cortex serve to magnify the differences between individual faces, as indicated by the statistical distribution of the encountered facial features. By this reasoning, the other race effect suggests that different face morphologies exhibit different statistical distributions of distinguishing characteristics. Valentine has previously employed principal component analysis to find different subspace projections and obtained results that mimick the other race effect.[12]

We stress that our aim is not to detect the *race* of a person; therefore, we use the term *morphology* to denote similar facial surface morphology characteristics. Based on the cognitive cues, we predict better recognition rates if the faces are clustered into morphological or gender groups that exhibit greater intra-group similarity and the discriminative features are learned within each group separately. This is not trivially true for all pattern recognition applications, as the grouping reduces the number of training samples and consequently runs the risk of impairing learning conditions.

4. SHAPE SPACE CLUSTERING

If our hypothesis of meta-classification is correct, we expect morphology and gender to be discriminating dimensions of the face space. However, we do not wish to categorize faces into races explicitly, as this approach has ethical consequences. Can the gender and race determination during the training (and possibly, in the testing) stage be evaded? For our simulation purposes, we have roughly assigned facial images into African, Asian and Caucasian morphological face classes. The other-race effect suggests that racial-morphology based clusters exist in the face space, and an unsupervised clustering method can recover those clusters, among other meaningful structure. Thus, it is not necessary to recover the race and gender of a person; the clustering will hopefully provide us with a useful set of average faces to serve in meta-classification with increased discrimination within clusters.

We propose to take a straightforward race- and gender-blind clustering approach with the k-means algorithm. The clustering is performed on the shape space, as represented by the aligned coordinates of seven facial landmarks. We specify the number of clusters for the shapes, and initialize the cluster consensus shapes by random selection from the training samples. At each iteration, we align the training samples to the consensus shapes of the clusters via Procrustes analysis, and assign each sample to the cluster with minimum average distance to the cluster consensus. We then re-estimate cluster consensus shapes from the samples assigned to the cluster, and iterate until the total distance stabilizes.

The clustering gives us a number of cluster consensus shapes, and assigns each training face to one of these clusters. We apply our AFM generation algorithm to these reduced training sets separately, and obtain one AFM for each cluster. These models can be seen in Section 7.3.

5. AUTOMATIC LANDMARK LOCALIZATION

Registration of facial images are usually guided by a few fiducial points of established correspondence (e.g. nose tip, eye and mouth corners). These points (also called the anchor points, or landmarks) have a great influence on the resulting registration. There are very few completely automatic systems for face recognition; most research results are reported on cropped and aligned (i.e. registered) faces, or the existence of a few manually located landmarks is assumed. This is apparently necessary, because the quality of the landmarks can be of greater consequence to the final accuracy than the classification algorithm itself. To understand the extent of the dependence to landmarks, we have contrasted manually selected ground truth with the results from a recent automatic landmark localization algorithm.[13,14,15] In this section, we briefly summarize this algorithm.

The method we employ is a coarse-to-fine approach, based on local feature information. During the training phase, 7×7 patches are cropped from around each landmark in the downsampled depth image (60×80). These patches are points in a 49-dimensional space, and their distribution is modeled with a mixture of Gaussians. We employ an unsupervised feature learning method that automatically tunes the model complexity to the complexity of the data distribution.[16] During the testing phase, this generative model is used to produce likelihoods for each point on the image. The highest-likelihood locations for each landmark are passed to a structural correction algorithm that can recover from localization errors.[13] The remaining errors are locations on the face with characteristics that are statistically similar to the landmark in question, and they conform to the general face pattern within a margin, as they have passed through structural correction. For instance the corners of the eyebrows may be selected as eye corners, or the cleft between the chin and the mouth may be detected as the mouth. The landmark detection was performed on the 3D depth map, which is found to be less informative than 2D images, but much more robust to illumination changes.[15]

The coarse localization is complemented with a fine-level search on the original 480×640 image. Only a relatively small (41×41) area around the coarsely located landmark is searched. The first and second depth gradients are estimated in vertical and horizontal directions, and convolved with kernels learned during the training stage. The kernel outputs are combined, and the highest response is taken as an indicator of the presence of the sought landmark. This approach is computationally simple, fast, and improves the quality of the coarse landmarks.

Figure 1 shows the accuracy of the automatic landmarking procedure in terms of millimeter-distance from the manually selected ground truth. The evaluation considers one to ten candidates for each landmark location, as indicated by the x-axis. The curves begin at the average distance for the best candidate, and decrease as better locations are found among the candidates. The quality of landmarking also depends on the type of landmark; the nose tip is easier to find in 3D as expected.

6. REGISTRATION METHODS

We test two different registration methods. In the first method (termed *TPS-based* in the experiments section), the test face is aligned to the average face with the TPS method, and the points in correspondence with the AFM are cropped.[2] This method deforms the test face to fit the AFM, and the amount of deformation is proportional to the number (and spread) of the landmarks. At the limit of using all facial points as landmarks, the face deforms into the AFM, losing the discriminative information completely. However with a few landmarks, corresponding facial structures are aligned.

In the second method, we use the iterative closest point method to align the test face with the AFM. ICP is a rigid registration method, hence the test face is not deformed at all. TPS-based methods are completely guided by the landmarks, whereas ICP needs a coarse initialization. Previous work on ICP show that a good initialization is necessary for fast convergence and an accurate end-result. We compare several approaches for the coarse registration.

In our first approach, the point with the greatest depth value is assumed to be the tip of the nose,[17,18] and a translation is found to align it to the nose tip of the AFM. This is the fastest and simplest heuristic we can use, and we expect it to perform well in the near-frontal faces. In the second approach, we use the manually determined nose tip (i.e. ground truth) in the coarse alignment. In the third approach, we use Procrustes analysis[9] to bring seven manually determined landmark points (inner and outer eye corners, nose tip, and

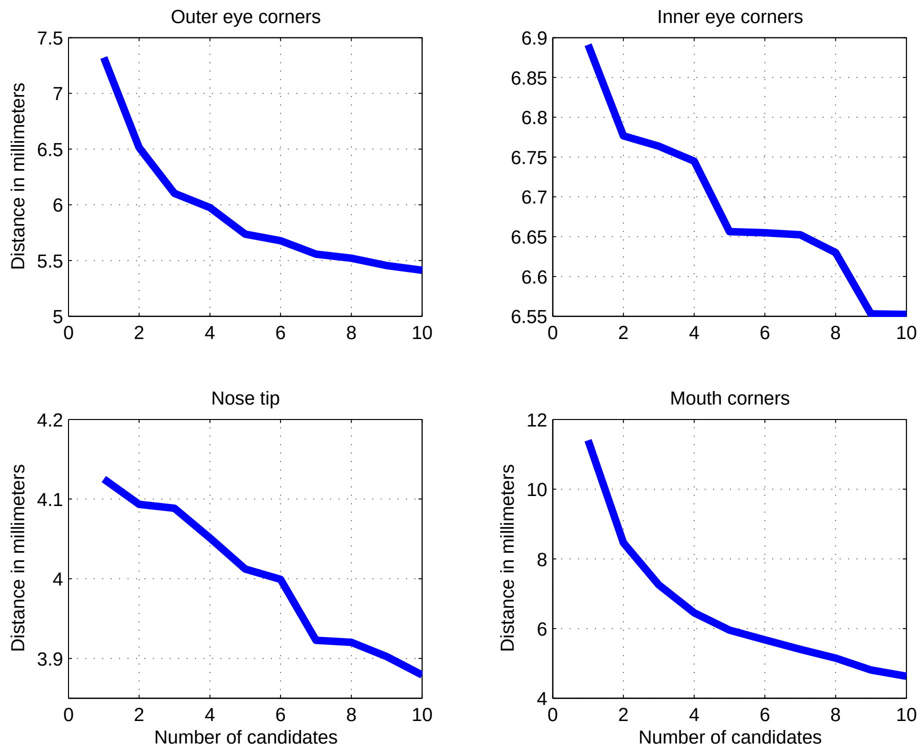

Figure 1. Accuracy of the automatical landmarking method, as indicated by average distance to true landmark in millimeters.

the mouth corners) into alignment with the average face model. Finally, we use Procrustes analysis to align automatically determined landmarks with the average face model. The automatic landmarking errors are not random, and cannot be simulated by injecting noise to the manually determined landmarks, except by modeling the specific landmarking procedure.

Intuitively, ICP will benefit from using category-specific AFMs, as the rigid registration is not able to cope with shape differences very well. A more similar average face will ensure that the dense correspondence will be established between points that have better structural correspondence. The TPS-based method will also benefit from category-specific AFMs, albeit for another reason. A more similar average face means that the test surface will be less deformed, and discriminatory information will not be lost.

7. EXPERIMENTS

We use a subset of the FRGC 2D-3D ver.1 face database in our experiments.[10] The database contains 943 near-frontal depth images from over 250 subjects, stored in a point cloud representation. We use images from 195 subjects, with one training face in the gallery and 1-4 test faces (for a total of 659 test faces). We only work with 3D information; 2D is not used at all. We design a number of experiments to answer various questions. Each subsection deals with a particular question, and reports relevant simulation results. The overall system has many dimensions, ruling out a factorial experimentation protocol.

For each method, we run a recognition experiment and a verification experiment. For the recognition experiments, the rank-1 recognition rate (R1) is reported. In the verification experiments, each of the 659 test faces is used for one genuine and 194 false claims, thus the number of false claims is two orders of magnitude higher. The equal error rate (EER) is reported under these conditions.

7.1. Coarse registration

Table 1 shows the effect of coarse alignment methods on ICP-based registration. As we have suspected, the nose-tip heuristic performs worst. Automatic localization of seven landmarks and using Procrustes alignment works better than the nose-tip heuristic. For ICP, using the nose ground truth works well in this dataset, because the faces we deal with are mostly upright and frontal. Ideally three landmarks should be used to accomodate greater pose differences in 3D. Finally, the manual landmarks with the Procrustes analysis give us an upper-bound on the performance of the ICP-PSD method.

We have also contrasted our AFM construction method with the method of Irfanoğlu et al.[2] on ICP. Manual landmarks were used, and initialization was by Procrustes alignment. With their smoother AFM, a rank-1 recognition rate of 86.34 was achieved, as opposed to our 92.11 per cent. Similarly, the EER was higher with their AFM by more than two per cent.

Table 1. Effect of coarse alignment on ICP

	Nose-tip heuristic	Automatic landmarks + Procrustes	Nose ground truth	Manual landmarks + Procrustes
R1	82.85	87.86	90.60	92.11
EER	14.25	8.12	6.60	6.20

7.2. Meta-classification

Does meta-classification and more specialized individual experts increase discrimination? We have tested this hypothesis by employing the average faces that are generated from groups of training faces. We have grouped the training samples for gender and morphology, and generated average face models (AFM) for each group. Figure 2 shows range images obtained for the average face, for male and female averages, for three morphological group averages (roughly Caucasian, Asian, and African), and for all combinations, respectively. The morphology does not correspond to a clear-cut distinction; the morphological group of a given face will be determined by its proximity to the average face of the group.

In Table 2 the authentication experiment results with or without specific average faces are shown for TPS-based registration. We have supplied both the generic-AFM based system and the specific-AFM based system with the categorical information. Then, any improvement in the specific system is strictly due to better registration. We have computed distances between the test face and the gallery faces with an L_1 distance metric, and trimmed the worst 2 per cent of correspondences. The EER results reported under these conditions show that specific AFM usage is beneficial in this case. We stress that the main purpose of the PSD is to evaluate the effect of AFM usage; the EER is generally too high for these experiments because of deformations in the registration.

Table 2. Simulation of TPS registration with deformation, EER

	None	Gender	Morphology	Gender & Morphology
Generic AFM	20.10	16.79	18.50	13.87
Specific AFM	20.10	14.64	16.97	11.47

The point set distance after aligning the face to the female and male averages can be used in gender classification. This simple method works with 80 per cent accuracy.

7.3. Shape Space Clustering

Our cognitively motivated subspace hypothesis is confirmed, if the shape space clustering automatically creates clusters with a dominant gender, or puts samples from one race into a single group. This is what more or less

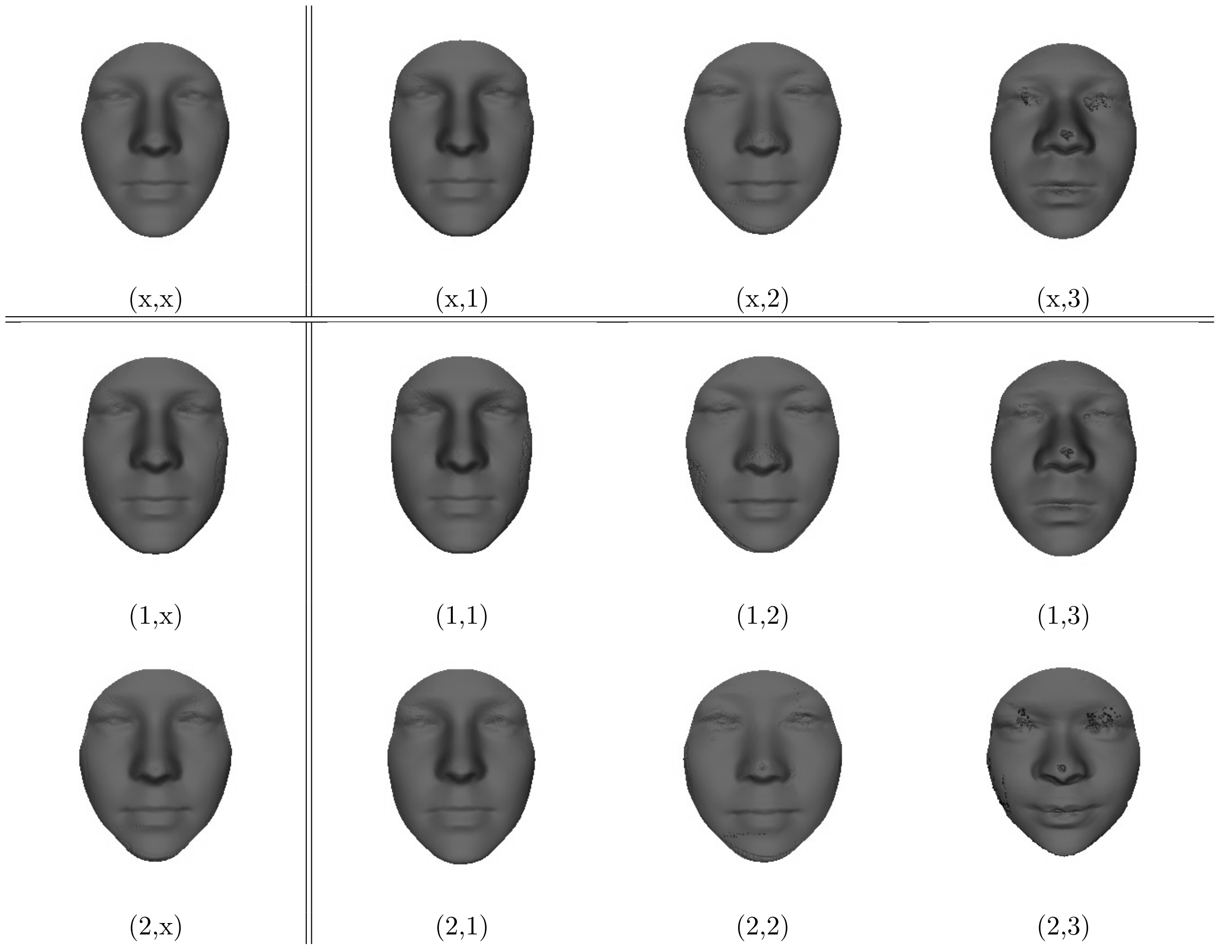

Figure 2. Average faces for different morphology and gender combinations. In (i, j) notation i is 1 (male) or 2 (female), and j is 1 (Caucasian), 2 (Asian) or 3 (African). Generic averages are denoted with x.

happens, although the non-uniformity of the training set (many more Caucasian males than, say African females) introduces a bias. We have specified six clusters, as in the full morphology-gender combination case, and ran our algorithm on the training part of the FRGC ver.1 dataset. Figure 3 shows the cluster means and Figure 4 shows the distribution of morphology and gender in each group as pie charts. We have clusters dominant in a single gender or a single morphology, as expected.

The number of training samples are evenly divided into clusters for this shape-space method. For the morphology-gender approach, the distribution was uneven, as we have too few samples from the some of the categories. However, this sort of handicap is expected, and it is necessary to see how the algorithms cope with it.

Table 3 shows recognition rates for ICP and TPS based systems with manual or automatic landmarks. The first row shows the results obtained with a single generic AFM. The next three rows show results with gender-, morphology-, and gender + morphology-based specific AFMs. The results for the last row are obtained with six shape-space derived clusters. For this last case, the registration does not benefit from the injection of categorical information, and each test sample is compared with all the training samples. The best result is obtained with shape space derived specific AFM and ICP.

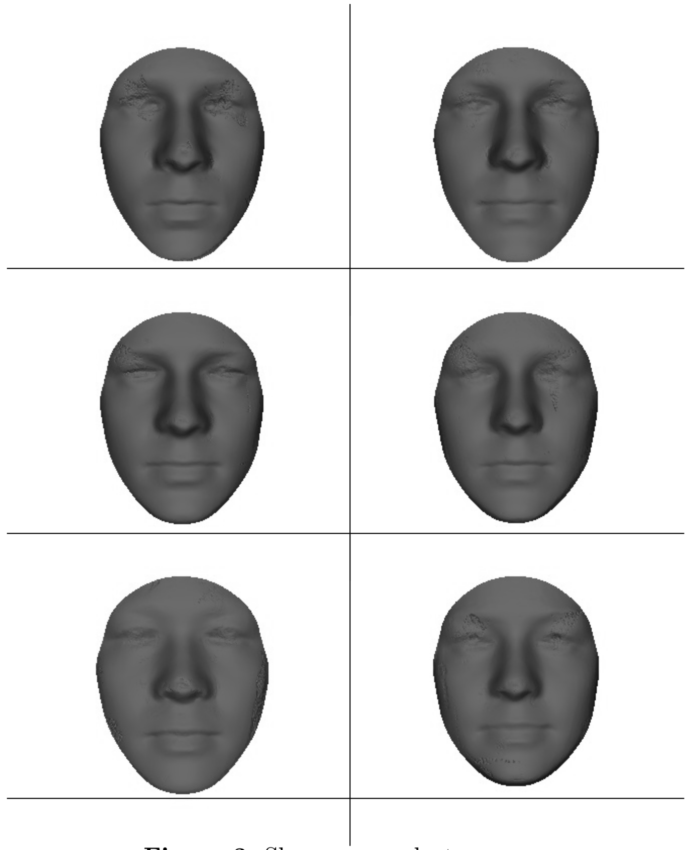

Figure 3. Shape space cluster means.

Table 3. Comparison of specific AFMs, rank-1 recognition rates

	Manual lm. + ICP	Automatic lm. + ICP	Manual lm. + TPS	Automatic lm. + TPS
Generic	92.11	87.86	52.20	42.64
Gender	90.14	86.65	54.63	45.52
Morphology	89.98	86.80	53.87	44.92
Gender & morphology	91.05	86.49	56.90	47.95
Shape space derived	93.78	91.20	47.65	41.58

8. CONCLUSIONS

We have evaluated ICP and TPS based registration of 3D faces with generic and specific average face models. Our proposed AFM generation method produces good models, and speeds up registration. The shape space clustering method revealed natural groups depending on morphology and gender in the face space, but also incorporated other factors that were useful for registration. Subsequently, the specific AFMs obtained with shape space clustering increased the accuracy of ICP.

Our experimental results have also confirmed that ICP is sensitive to initialization, and automatical landmarking as a pre-processing step is beneficial to ICP. The nose-tip heuristic may be useful in frontal faces, but the hair, clothing and sometimes the chin can be erroneously detected as the nose tip. The error due to incorrect nose localization can be gauged by looking at the results of the simulations that use the ground-truth for the

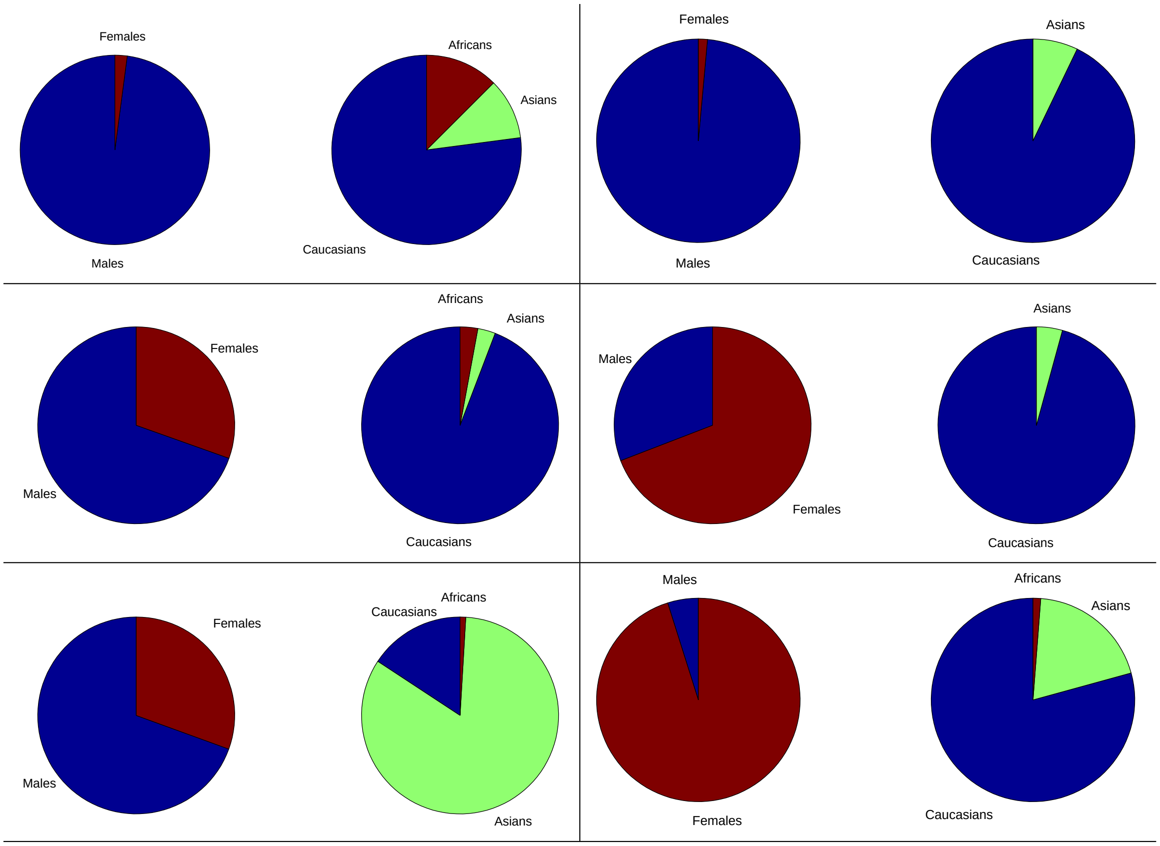

Figure 4. Shape space clustering distributions. For each cluster, the gender and morphology distributions are shown in separate pie charts. For six clusters, we have two dominantly male, and two dominantly female clusters, one dominantly Asian cluster, and almost all the males labeled as African are clustered into a single group.

nose in initialization. We should also keep in mind that the database we use is made up of near-frontal faces. The nose-tip heuristic will perform worse in any other pose settings.

The results show that the TPS based method is much inferior to ICP in accuracy. The beneficial effect from specific AFMs is evident in TPS methods that use either automatic or manual landmarks. Improvement is observed in the ICP method for the shape space clustering method, but not for the gender & morphology-specific AFMs. It is interesting to note that although the gender & morphology-based specific AFM method reduces the number of candidates for classification (whereas shape space derived AFM method does not), there is no accuracy improvement.

There may be several reasons for this lack of improvement. The categorical information is apparently not as beneficial as one would hope; closer inspection reveals that cross-gender and cross-morphology confusions are relatively rare. We can also argue that the distribution of categories in the training and test sets were uneven, and this reduced the quality of some of the specific AFMs. The TPS based method suffers less, because it only uses the landmark locations during registration, and not the actual AFM surface.

Another point to consider is that the other race effect pre-supposes feature learning within the particular manifold on the face space. Our point set difference method does not strictly correspond to learning features. We conjecture that a method using PCA subspace projection after registration would benefit more from the specific AFMs, as for each category a different subspace projection will be learned. We leave this as a future work.

A final observation is that the specific AFM models have different numbers of points. A male face usually contains 20 per cent more points than a female face. When we align a face to the female and the male AFMs, the distribution of distances is different in the centre of the face and at the periphery. Using a smaller AFM (the one for the females, or Asians, for instance) will effectively remove the points close to the periphery from the distance calculation. This can be an issue for one-to-all ICP approaches as well.

As a future work, we intend to compare AFM-based ICP with one-to-all ICP. The TPS-based method we have used does not perform well, although it is very fast. In another deformation-based methodology, the AFM is only used to label the points of the test face, which are then transformed rigidly into alignment with the original position of the AFM and used for classification.[19] The method gives a potentially better labeling of facial points, although the quality of landmarks plays a crucial role. Another possible drawback of this latter method is the costly iterative optimization used for local alignment.

ACKNOWLEDGMENTS

This work is supported by the Scientific & Technological Research Council of Turkey research grant 104E080 and by the European FP6 Network of Excellence BIOSECURE.

REFERENCES

1. Chua, C.S., F. Han, Y.K. Ho, "3D human face recognition using point signature," in *Proc. IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 233–238, 2000.
2. İrfanoğlu, M.O., B. Gökberk, L. Akarun, "3D Shape-Based Face Recognition Using Automatically Registered Facial Surfaces," in *Proc. Int. Conf. on Pattern Recognition,* vol. 4, pp. 183–186, 2004.
3. Lu, X. A.K. Jain, D. Colbry, "Matching 2.5D Face Scans to 3D Models," *IEEE Trans PAMI*, vol. 28, no. 1, 2006.
4. Pan, G., Z. Wu, "3D Face Recognition From Range Data," *Int. Journal of Image and Graphics*, vol. 5, no. 3, pp. 573–583, 2005.
5. Papatheodorou, T., D. Rueckert, "Evaluation of automatic 4D face recognition using surface and texture registration," in *Proc. AFGR*, pp. 321–326, 2004.
6. Hüsken, M., M. Brauckmann, S. Gehlen, C. von der Malsburg, "Strategies and Benefits of Fusion of 2D and 3D Face Recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2005.
7. Besl, P., N. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans PAMI*, vol. 14, no. 2, pp. 239–256, 1992.
8. Bookstein, F.L., *Morphometric Tools for Landmark Data: Geometry and Biology*, Cambridge University Press, 1991.
9. Goodall, C., "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society B,* vol. 53, no. 2, pp. 285–339, 1991.
10. Phillips, P.J., P.J. Flynn, W.T. Scruggs, K.W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, W.J. Worek, "Overview of the Face Recognition Grand Challenge," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 1, pp. 947–954, 2005.
11. Tong, M.H., C.A. Joyce, G.W. Cottrell, "Are Greebles special? Or, why the Fusiform Fish Area would be recruited for sword expertise (if we had one)," in *Proc.* 27^{th} *Annual Cognitive Science Conference*, La Stresa, Italy. Mahwah: Lawrence Erlbaum, 2005.
12. Valentine, T., "A unified account of the effects of distinctiveness, inversion and race in face recognition," *Quarterly Journal of Experimental Psychology*, vol. 43A, pp. 161–204, 1991.
13. Salah, A.A., H. Çınar Akakın, L. Akarun, B. Sankur, "Robust Facial Landmarking for Registration," accepted for publication in *Annals of Telecommunications*, 2006.
14. Çınar Akakın, H., A.A. Salah, L. Akarun, B. Sankur, "2D/3D Facial Feature Extraction," *Proc. SPIE Conference on Electronic Imaging*, 2006.
15. Salah, A.A., L. Akarun, "3D Facial Feature Localization for Registration," in B. Günsel et al. (eds.), *Lecture Notes in Computer Science,* vol. 4105/2006, International Workshop on Multimedia Content Representation, Classification and Security, pp. 338-345, 2006.

16. Salah, A.A., E. Alpaydın, "Incremental Mixtures of Factor Analyzers," in *Proc. Int. Conf. on Pattern Recognition*, vol. 1, pp. 276–279, 2004.
17. Colbry, D., G. Stockman, A.K. Jain, "Detection of Anchor Points for 3D Face Verification," in *Proc. IEEE Workshop on Advanced 3D Imaging for Safety and Security*, 2005.
18. Yan, Y., K. Challapali, "A system for the automatic extraction of 3-D facial feature points for face model calibration," in *Proc. Int. Conf. on Image Processing*, vol.2, pp.223-226, 2000.
19. Mao, Z., X. Ju, J.P. Siebert, W.P. Cockschott, A. Ayoub, "Constructing dense correspondences for the analysis of 3D facial morphology," *Pattern Recognition Letters*, vol. 27, pp. 597–608, 2006.