SHREC'08 Entry: Shape Based Face Recognition with a Morphable Model

Brian Amberg* Reinhard Knothe* Thomas Vetter*

Abstract

We present a method for face recognition by fitting a 3D Morphable Model to shape data. Fitting is done with a a robust nonrigid ICP algorithm. For recognition, it is possible to use either the fitted model parameters, or the correspondences induced by the model. We compare different similarity measures, and show that a 3D Morphable Model allows very robust retrieval results.

Index Terms: I.4.8 [Image Processing and Computer Vision]: Scene Analysis—Object Recognition, Surface Fitting, Range Data I.4.7: Feature Measurement—Size and Shape, Feature representation I.4.9: Applications I.5.1 [Pattern Recognition]: Models—Statistical

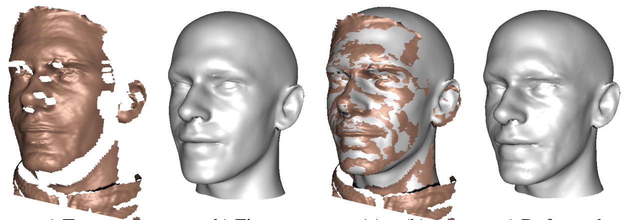

Figure 1: The robust fitting gives a good estimate (b) of the true face surface given the noisy measurement (a). It fills in holes and removes artifacts using prior knowledge from the face model. The fitted shape plus the exact correspondences found can be used to extrapolate the image by a robust poisson deformation (c).

1 Introduction

We tackle the task of textureless 3D face recognition. The system is fully automatic and can handle the typical artifacts of 3D scanners, namely outliers and missing regions. Face recognition in this setting is a difficult task, and difficult tasks need strong prior knowledge. To introduce the prior knowledge we use a 3D Morphable Model (3DMM) [3], which is a generative statistical model of 3D faces. 3DMM have been applied successfully for face recognition on different modalities. The most challenging setting is recognition from single images under varying light and illumination. This was adressed by [4, 7]. There a 3DMM with shape, texture and illumination model was fit to probe and gallery images. As the model separates shape and albedo parameters from pose and lighting, it enables pose and lighting invariant recognition. In [1] a similar approach was used to fit a pure shape model to stereo images, also enabling recognition by correlating the shape parameters. We use the same approach for shape based face recognition. We fit a 3DMM build from 170 subjects with neutral expressions to the gavabDB [6] database, and compare different distance measures which can be derived from the model fit.

An alternative to fitting a generative model is to align the probe to each example in the database using e.g. ICP [8]. But comparing the probe directly to every gallery image has the disadvantage of scaling linearly with the number of entries in the gallery, while for a model based approach only a single fit to the probe is necessary, and the comparision to the database can then be performed by a distance measure in the lower dimensional space of registered faces.

Another interesting model-less approach [5] compares surface by the distribution of geodesics, which stays constant for nonrigidly deforming (but not stretching or tearing) objects. This approach is difficult to apply in this setting though, as the scanning produces holes, disconnected regions and strong noise, which can best be handled by a method which uses specific information about the object class.

2 Fitting

The fitting algorithm used in this paper is a variant of the nonrigid ICP work in [2]. It is a robust iterated fitting algorithm. Like other ICP methods, it is a local optimization method, which does not guarantee convergence to the global mimimum, but is dependent on the initialization. It consists of the following steps

- Iterate over a sequence of regularization values $\theta_1 > \cdots > \theta_N$.
  - Repeat until convergence.
    1. Find candidate correspondences by searching for the closest compatible point for each model vertex.
    2. Weight the correspondences by their distance using a robust estimator.
    3. Fit the 3DMM to these correspondences using a regularization strength of θ_i.
    4. Continue with the lower θ_{i+1} if the median change in vertex position is smaller than a threshold.

The search for the closest compatible point takes only points in account which have conforming normals. Note, that it is necessary to balance robustness and regularization, as the right balance depends on the noise characteristic of the data. Suitable values were determined manually for a single scan and kept constant for all experiments. In step 3 the 3DMM is fit to 3D-3D point correspondences. This is done with a gauss-newton least squares optimization, using an analytic Jacobian and first-order Hessian.

As the database is pose normalized, we initialize the registration such that the tip of the nose and pose coincides. This initialization is good enough to fit the complete database fully automatic. For non pose-normalized databases, we would either need three landmarks, or – to keep the algorithm fully automatic – repeated random initialization.

3 Retrieval

We evaluated four different distance measures.

3.1 Model Based Measures

We begin with measures which are acting in the parameter space of the model. These have the advantage of being extremely cheap to calculate, once the model has been fit.

3.1.1 Mahalanobis distance of shape coefficients

The first method calculates the distance between two vectors of shape coefficients $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$ expressed in Mahalanobis space as

$$s_1(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \|\boldsymbol{\alpha}_1 - \boldsymbol{\alpha}_2\| \quad . \tag{1}$$

*e-mail: {brian.amberg, reinhard.knothe, thomas.vetter}@unibas.ch

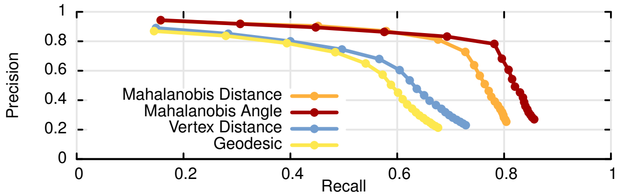

Figure 2: Measures in Mahalanobis Space outperform vertex based measures.

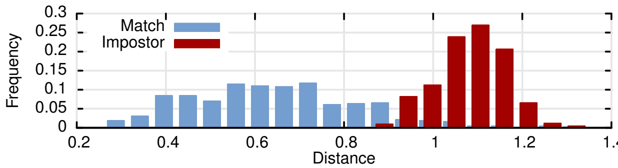

Figure 3: Impostor detection is reliable, as the minimum distance to a match is smaller than the minimum distance to a nonmatch.

3.1.2 Angular distance of shape coefficients

In face space, caricatures lie along the rays from the origin. Mapping all caricatured versions of a face onto a canonical face gives a method which has proven to have very high recognition rates [4]. To do this we use the angle between the shape coefficients in mahalanobis space as the distance measure.

$$s_2(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) = \arccos\left(\frac{\boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_2}{\|\boldsymbol{\alpha}_1\| \|\boldsymbol{\alpha}_2\|}\right) \quad . \tag{2}$$

For the angular measure shown in Figure 3.1 the distribution of distances towards the first match and the first nonmatch in the database. This shows, that by choosing a suitable threshold it is possible to perform face recognition with impostors, where we decide if the identity is in the database or not.

These two measures give similar results, but the caricature invariance of the angular distance improves recall a bit. This tells us, that in face space, the direction alone codes the identity.

3.2 Shape Based Measures

The second type of measures acts in vertex space. If we want to compare two model instances, it does not make much sense to measure in vertex space instead of parameter space, as there is a one to one mapping between the spaces and the parameter space is of much lower dimensionality. But in our case we do have additional information which can not be expressed by the model. After fitting, some residual will remain, which is caused by three reasons. 1) The individuals used to train the model were not from the gallery, and we can not span the complete face space with 170 training examples. 2) The probe images have expressions, while the database was build using neutral expressions. 3) The aquisition process introduces noise. Therefore, we add more flexibility to the model by allowing smooth nonrigid deformations of the final fit to minimize the remaining residual. This is achieved by robustly fitting a poisson deformation with soft boundary and zero right hand side, where the boundary is given by the correspondences found by the fitting algorithm, and the deformed shape is the fitted head. Results can be seen in Figure 2. With these correspondence established we can use different distance measures.

3.2.1 Distance

Denote the N vertices of the registered scan i as $\boldsymbol{v}_1^i, \ldots, \boldsymbol{v}_N^i$. We use the distance after removing the rigid transformation, measuring only in a mask defined on the model, which includes the parts which are visible in most of the scans. We compute

$$s_3(\mathbf{v}^1, \mathbf{v}^2) = \min_{\mathbf{R}, \boldsymbol{t}} \sum_i \|\mathbf{R}\boldsymbol{v}_i^1 + \boldsymbol{t} - \boldsymbol{v}_i^2\| \quad . \tag{3}$$

where $\mathbf{R}, \boldsymbol{t}$ describe a similarity transform. While this measure is straight forward, the recognition results are unsatisfying. The scaling of face space which is learned from the example faces results in an improved clustering of scans, which enables better classification than comparisions in the original vertex space.

3.2.2 Geodesics

Inspired by [5] we tried a geodesic based measure, which should be invariant against expression changes. We classify by comparing the distances of a selected set of vertices, which were assumed not to change under expressions. Denote the selected vertex pairs by $\mathscr{P}$. Note that this is different from [5], as we have already brought the meshes into correspondence during fitting. The measure is then

$$s_4(\mathbf{v}^1, \mathbf{v}^2) = \sum_{(i,j)\in\mathscr{P}} \left| \|v_i^1 - v_j^1\| - \|v_i^2 - v_j^2\| \right| \tag{4}$$

Experiments with different sets of distances showed that the best classification results were achieved by selecting all neighboring edges of the model in the face area. Still, exactly because this method introduces some invariance, it also reduces the precision in the retrieval experiments. Also, it does not incorporate the knowledge about face space, which was exploited in the first two methods.

4 Conclusion

We have shown that 3D Morphable Models provide a valuable tool for face recognition with 98.8% recognition rate on this database. The strong prior knowledge allows robust handling of noisy data. Four distance measures were compared, and it turns out that the angular distance in Mahalanobis space is the most accurate classifier. It was noticably better than measuring the difference between the registered scans. This is because the face space learned from our training examples is constructed such that the identities are better separated than in vertex space.

In the future we wish to make the system expression invariant, by using a model which separates expression and identity. As we do establish correspondence between the model and the scans, it is trivial to add image based classification for datasets where a calibrated photo is available, by comparing the rectified textures.

Acknowledgements

The authors wish to thank P. Paysan for the data acquisition. This work was supported in part by a grant from Microsoft Research and the Swiss National Science Foundation (200021-103814 and NCCR COME project 5005-66380).

References

[1] B. Amberg, A. Blake, A. Fitzgibbon, S. Romdhani, and T. Vetter. Reconstructing high quality face-surfaces using model based stereo. In *ICCV 2007*, pages 1–8, 2007.

[2] B. Amberg, S. Romdhani, and T. Vetter. Optimal step nonrigid ICP algorithms for surface registration. In *CVPR 2007*, pages 1–8, 2007.

[3] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d faces. In *SIGGRAPH '99*, pages 187–194. ACM Press, 1999.

[4] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *PAMI*, 25(9):1063–1074, September 2003.

[5] A. Bronstein, M. Bronstein, and R. Kimmel. Three-dimensional face recognition. *IJCV*, 64(1):5–30, August 2005.

[6] A. B. Moreno and A. Sánchez. Gavabdb: a 3d face database. In *Workshop on Biometrics on the Internet*, pages 77–85, Vigo, March 2004.

[7] S. Romdhani, J. Ho, T. Vetter, and D. J. Kriegman. Face recognition using 3-d models: Pose and illumination. *Proceedings of the IEEE*, 94(11):1977–1999, 2006.

[8] P. Yan and K. W. Bowyer. A fast algorithm for ICP-based 3d shape biometrics. In *Automatic Identification Advanced Technologies, 2005. Fourth IEEE Workshop on*, pages 213–218, 2005.