# 3D Face Morphable Models “In-the-Wild”

Chaonan Song

May 27, 2018

# 1 Model Fitting

Professor James propose to fit the 3DMM on an input image using a Gauss-Newton iterative optimisation. To this end, herein, James first formulate the cost function and then present two optimisation procedures.

## 1.1 Cost Function

The overall cost function of the proposed 3DMM formulation consists of a texture-based term, an optional error term based on sparse 2D landmarks, and an optional regularisation terms on the parameters.

**Texture reconstruction cost.** The main term of the optimisation problem is the one that aims to estimate the shape, texture and camera parameters that minimise the $\mathcal{L}_2^2$ norm of the difference between the image feature-based texture that corresponds to the projected 2D locations of the 3D shape instance and the texture instance of the 3DMM.

**Regularisation.** In order to avoid over-fitting effects, James augment the cost function with two optional regularisation terms over the shape and texture parameters. James formulate the regularisation terms as the $\mathcal{L}_2^2$ of the parameters vectors weighted with the corresponding inverse eigenvalues.

**Overall cost function.** The landmarks term as well as the regularisation terms are optional and aim to guide the optimisation procedure to converge faster and to a better minimum. Note that thanks to the proposed “in-the-wild” feature-based texture model, the cost function does not include any parametric illumination model similar to the ones in the relative literature Blanz and Volker [1], Blanz *et al.* [2], which greatly simplifies the optimisation.

## 1.2 Gauss-Newton Optimisation

Inspired by the extensive literature in Lucas-Kanade 2D image alignment Baker *et al.* [3] we formulate a Gauss-Newton optimisation framework. Specifically, given that the camera projection model is applied on the image part of Eq. 1, the proposed optimisation has a “forward” nature.

$$\begin{aligned} \arg\min_{p,c,\lambda} \parallel F\left(\mathcal{W}(p,c)\right) - \mathcal{T}(\lambda) \parallel^2 \\ +c_l \parallel \mathcal{W}_l(p,c) - s_l \parallel^2 + c_s \parallel P \parallel^2 + c_t \parallel \lambda \parallel^2 \end{aligned} \tag{1}$$

# 2 KF-ITW Dataset

For the evaluation of the 3DMM, James have constructed KF-ITW, the first dataset of 3D faces captured under relatively unconstrained conditions. The dataset consists of 17 different subjects recorded under various illumination conditions performing a range of expressions (neutral, happy, surprise). James employed the KinectFusion [4, 5] framework to acquire a 3D representation of the subjects with a Kinect v1 sensor.

The fused mesh for each subject serves as a 3D face ground-truth in which we can evaluate our algorithm and compare it to other methods. A voxel grid of size $608^3$ was utilised to get the detailed 3D scans of the faces. In order to accurately reconstruct the entire surface of the faces, a circular motion scanning pattern was carried out. Each subject was instructed to stay still in a fixed pose during the entire scanning process. The frame rate for every subject was constant to 8 frames per second.

# 3 Experiments

## 3.1 3D Shape Recovery

James use the ground-truth annotations provided in the KFITW dataset to initialize and fit all three techniques to the "in-the-wild" style images in the dataset. Our error metric is the per-vertex dense error between the recovered shape and the model-specific corresponded ground-truth fit, normalized by the inter-ocular distance for the test mesh. Fig. 1 shows the cumulative error distribution for this experiment for the three models under test.

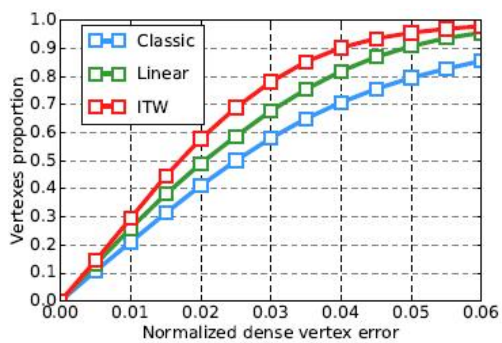


Figure 1: Accuracy results for facial shape estimation on KF-ITW database. The results are presented as Cumulative Error Distributions of the normalized dense vertex error. Table 1 reports additional measures.

Table 1 reports the corresponding Area Under the Curve (AUC) and failure rates.

| *Method* | *AUC* | *Failure Rate (%)* |
|---|---|---|
| **ITW** | **0.678** | **1.79** |
| Linear | 0.615 | 4.02 |
| Classic | 0.531 | 13.9 |

Table 1: Accuracy results for facial shape estimation on KF-ITW database. The table reports the Area Under the Curve (AUC) and Failure Rate of the Cumulative Error Distributions of Fig. 1.

Figure 2 demonstrates qualitative results on a wide range of fits of "in-the-wild" images drawn from the Helen and 300W datasets that qualitatively highlight the effectiveness of the proposed technique.

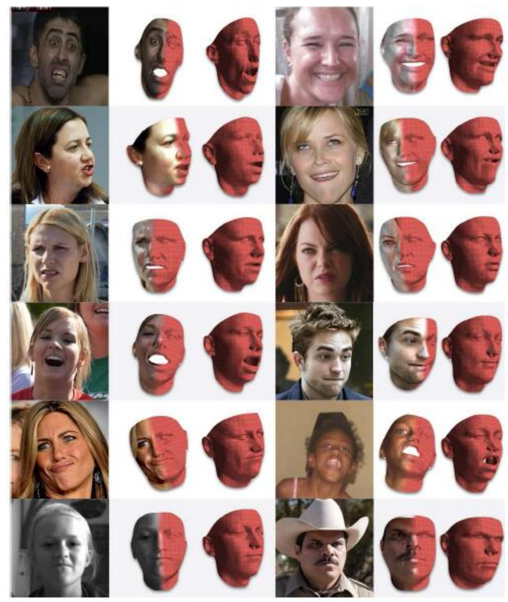

Figure 2: Examples of in the wild fits of our ITW 3DMM taken from 300W [6]

# 4 Conclusion

Professor James capitalise on the annotated "in-the-wild" facial databases to propose a methodology for learning an "in-the-wild" feature-based texture model suitable for 3DMM fitting without having to optimise for illumination parameters.

# References


[1] V. Blanz, "A morphable model for the synthesis of 3d faces," *Acm Siggraph*, pp. 187–194, 1999. 1

[2] V. Blanz and T. Vetter, "Face recognition based on fitting a 3d morphable model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1063–1074, 2003. 1

[3] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International*

*Journal of Computer Vision*, vol. 56, no. 3, pp. 221–255, 2004. 1

[4] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, and A. Davison, "Kinectfusion:real-time 3d reconstruction and interaction using a moving depth camera," in *ACM Symposium on User Interface Software and Technology, Santa Barbara, Ca, Usa, October*, 2011, pp. 559–568. 1

[5] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohi, J. Shotton, S. Hodges, and A. Fitzgibbon, "Kinectfusion: Real-time dense surface mapping and tracking," in *IEEE International Symposium on Mixed and Augmented Reality*, 2012, pp. 127–136. 1

[6] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: database and results," *Image Vision Computing*, vol. 47, pp. 3–18, 2016. 2