Resolution-Aware 3D Morphable Model

Guosheng Hu
g.hu@surrey.ac.uk
Chi Ho Chan
chiho.chan@surrey.ac.uk
Josef Kittler
j.kittler@surrey.ac.uk
William Christmas
w.christmas@surrey.ac.uk

Centre for Vision, Speech and Signal Processing
University of Surrey
Guildford, UK

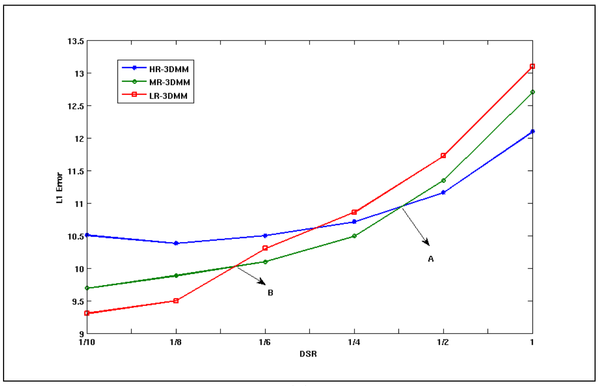
Face reconstruction models can be classified into two groups: the 2D face models and the 3D face models. The 2D-based group includes Active Appearance Models (AAMs) [4], which achieved great success. But the problem for AAMS is that the reconstruction with AAMs fails if the in-depth rotation of the face becomes large. 3D-based methods have been proposed to solve the problem. 3D morphable model (3DMM) [1] is a well-known statistical model for face reconstruction. It is very interesting to investigate how the fitting performance is affected by the resolution of 3DMM and of the input image. In this paper, the relationship between 3DMM resoluton and input image resolution is studied and a Resolution-Aware 3DMM (RA-3DMM) model is proposed. The construction of RA-3DMM is motivated by the assumption that **a high resolution 3DMM fits high resolution input images better and a low resolution 3DMM fits low resolution input images better**. Based on this assumption, a set of 3DMMs of different resolution should work better than a single 3DMM if the input images are of different resolution.

In this work, RA-3DMM consists of a selector and these 3DMMs: High Resolution 3DMM(HR-3DMM), Medium Resolution 3DMM(MR-3DMM) and Low Resolution 3DMM(LR-3DMM). The model selector automatically selects the best 3DMM to fit the input face image according to its resolution. Clearly, the selection strategies of the selector are very important for RA-3DMM. These three 3DMMs are obtained by the 4-8 mesh subdivision algorithm [5].

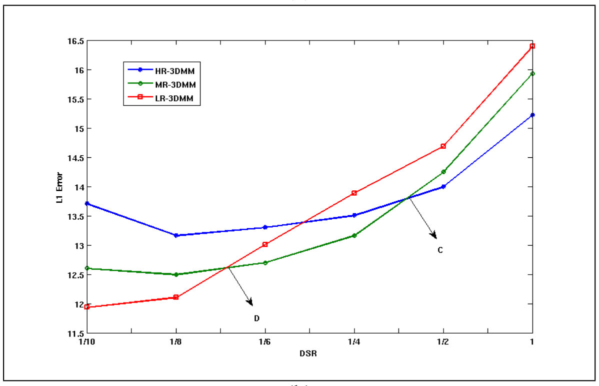
We evaluated the proposed RA-3DMM model on two face databases: XM2VTS and PIE. For building the RA-3DMM, only images with frontal pose and neutral illumination in the two databases are used. The images in XM2VTS and PIE are down-sampled to different resolutions. In this work, the down-sample rate (DSR) is 1, 1/2, 1/4, 1/6, 1/8, 1/10. All these down-sampled images were fitted by HR-3DMM, MR-3DMM and LR-3DMM. The L1-Norm is used to estimate the fitting error between the input image and the fitted image.

The fitting results on XM2VTS and PIE are shown in Fig. 1. It is clear that different models perform differently with input images of different resolutions: Obviously, HR-3DMM works best with DSR=1, 1/2; MR-3DMM with DSR=1/4, 1/6; and LR-3DMM works best with DSR=1/8, 1/10. Also it is important to know point A, B, C and D in Fig. 1 to define the selection strategies. It is not hard to conclude that the resolutions corresponding to A, B, C and D are 9397, 1344, 7422 and 985 respectively. So the selection strategies of RA-3DMM are determined as follows: under the diffused light (such as XM2VTS), HR-3DMM is selected for fitting if the input image is of high resolution (greater than 9397 pixels). MR-3DMM is selected for fitting if the input image is of medium resolution (between 1344 pixels and 9397 pixels). LR- 3DMM is selected for fitting if the input image is of low resolution (smaller than 1344). Under the point source light (such as PIE), HR-3DMM is selected for fitting if the input image is of high resolution (greater than 7422 pixels). MR-3DMM is selected for fitting if the input image is of medium resolution (between 985 pixels and 7422 pixels). LR-3DMM is selected for fitting if the input image is of low resolution (smaller than 985).

Then we apply RA-3DMM to pose correction with input images ranging from high resolution to low resolution [2]. The texture of the visible parts is extracted from the input image, and RA-3DMM automatically selects the best model to reconstruct the occluded part as discussed above. Then the Multiscale Local Phase Quantisation histogram (MLPQH) [3] descriptor is used for face verification. In this experiment, XM2VTS MPEG7 and the standard frontal datasets are used in conjugation with the Configuration 1 of the Lausanne protocol[3]. Following the Lausanne protocol, the total error rate (TER) is reported. TER is the sum of false acceptance rate (FAR) and false rejection rate (FRR) at a threshold. In our experiments the performance with RA-3DMM pose correction and without pose correction is compared. The TER of all poses with RA-3DMM pose correction is much smaller than that without pose correction for all resolutions. Even for low resolution face verification, which is a hard task in face recognition, RA-3DMM shows considerable improvement over the method without pose correction.



(a)



(b)

Figure 1: Fitting Results on (a) XM2VTS and (b) PIE

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

[2] V. Blanz, P. Grother, P.J. Phillips, and T. Vetter. Face recognition based on frontal views generated from non-frontal images. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 454–461. IEEE, 2005.

[3] C.H. Chan, J. Kittler, N. Poh, T. Ahonen, and M. Pietikainen. (multi-scale) local phase quantisation histogram discriminant analysis with score normalisation for robust face recognition. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 633–640. IEEE, 2009.

[4] T.F. Cootes, G.J. Edwards, and C.J. Taylor. Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(6):681–685, jun 2001.

[5] L. Velho and D. Zorin. 4-8 subdivision. *Computer Aided Geometric Design*, 18(5):397–427, 2001.