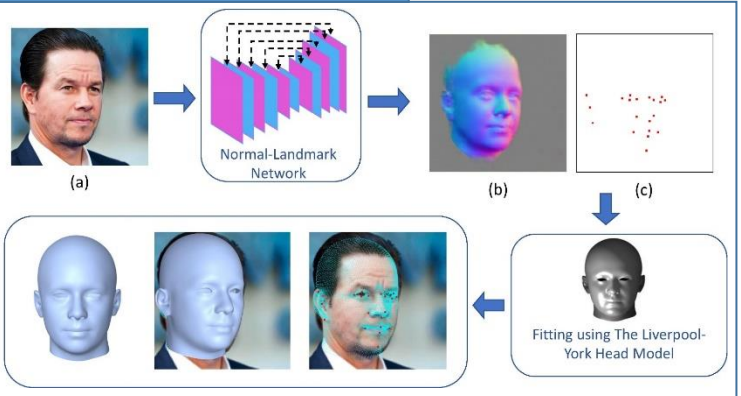
Hybrid Approach for 3D Head Reconstruction: Using Neural Networks and Visual Geometry

Oussema Bouafif[1,2], Bogdan Khomutenko[1] and Mohamed Daoudi[2]
[1]MCQ-Scan, Lille, France [2]IMT Lille Douai, Univ. Lille, CNRS UMR 9189 CRIStAL, Lille, France
Email: o.bouafif@mcq-scan.com, b.khomutenko@mcq-scan.com, mohamed.daoudi@imt-lille-douai.fr

Abstract

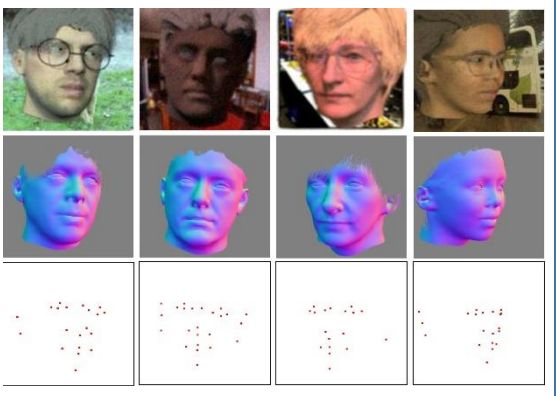
- We present a novel method for reconstructing 3D heads from a single or multiple image(s) using a hybrid approach based on deep learning and geometric techniques.
- We propose an encoder-decoder network based on the U-net architecture and trained on synthetic data only. It predicts both pixel-wise normal vectors and landmarks maps from a single input photo. Landmarks are used for the pose computation and the initialization of the optimization problem, which, in turn, reconstructs the 3D head geometry by using a parametric morphable model and normal vector fields.
- State-of-the-art results are achieved through qualitative and quantitative evaluation tests on both single and multi-view settings.

Overview of the Proposed Framework



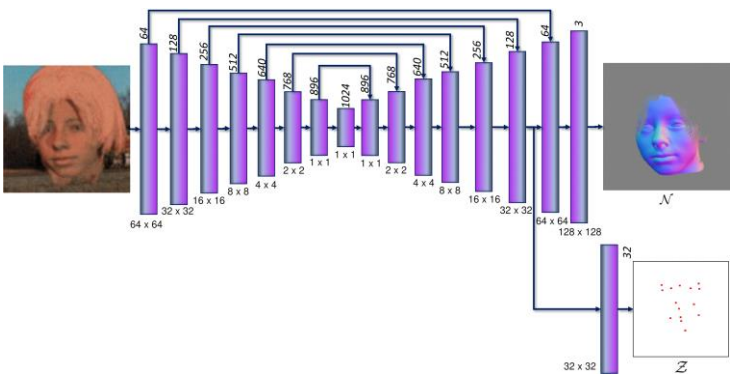
Given an input facial image (a), we estimate two different maps (normal surface map *N* (b), Landmarks map *Z* (c)) used to reconstruct the 3D facial shape via a fitting process with the LYHM [1] morphable model.

Synthetic Dataset



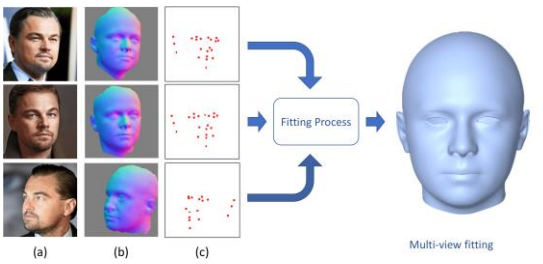
Training data samples. From top to bottom: Synthetic facial images, Normal surface maps *N* and Landmarks maps *Z*.

Normal-Landmark Network



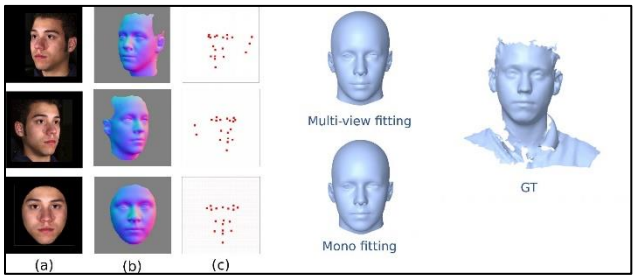
Our encoder-decoder architecture produces two different maps (*N* and *Z*) (shown on the right) from a facial input image (shown on the left). The spatial size and number of layers are shown below and above each block, respectively.

Multi-view Fitting Process



Our method can be used in mutli-view process. We estimate the two maps *N* (b) and *Z* (c) from each input image (a), then we use them in the same fitting process to get a unique 3D head reconstruction.

Stereo vs Mono Fitting



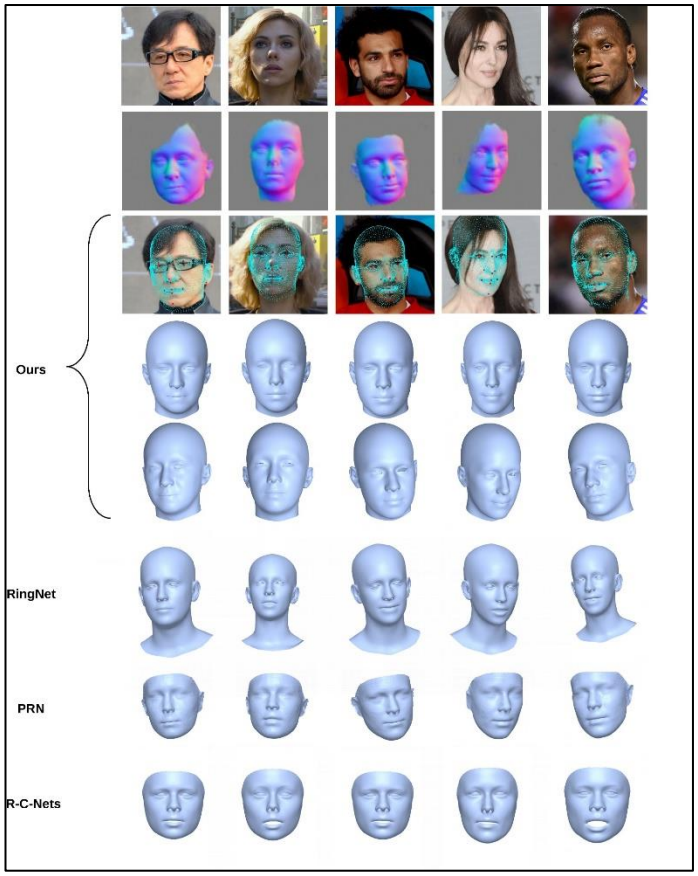
- (Multi-view fitting): 3D head reconstruction using all images in the same fitting process.
- (Mono fitting) : 3D head exploiting only the frontal image in the fitting process (third row).
- (GT) : the ground-truth of the 3D head mesh.

Quantitative Evaluation

Method	Ours (Mono)	Ours (Multi)	RingNet [3]	PRN [4]	R-C-Nets [5]
RMSE	1.74±0.44	1.67±0.43	1.90±0.49	1.86±0.47	1.60±0.41
μ	2.21	2.17	3.42	1.83	1.64
σ	1.08	1.04	1.58	1.70	1.69
m̃	2.09	2.05	3.23	1.43	1.27
$\delta_{90\%}$	3.61	3.53	5.60	3.46	3.00

Quantitative comparison on the BU-3DFE [2] dataset. Lowers are better

Qualitative Evaluation



Visual comparisons from some celebrities facial images with other methods. Rows contain in order; input image, predicted *N* map, input image with predicted landmarks (red squares) and dense alignment results. Ours (frontal view), Ours (aligned) RingNet [3], PRN [4] and R-C-Nets [5]

References

[1] H. Dai, N. Pears, W. A. P. Smith, and C. Duncan, "A 3D morphable model of craniofacial shape and texture variation," in The IEEE Inter-national Conference on Computer Vision (ICCV), Oct 2017.
[2] L. Yin, X. Wei, Y. Sun, J. Wang, and M. J. Rosato, "A 3D facial expression database for facial behavior research," in7th international conference on automatic face and gesture recognition (FGR06). IEEE,2006, pp. 211–216.
[3] S. Sanyal, T. Bolkart, H. Feng, and M. J. Black, "Learning to regress 3D face shape and expression from an image without 3D supervision" in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 7763–7772.
[4] Y. Feng, F. Wu, X. Shao, Y. Wang, and X. Zhou, "Joint 3d face reconstruction and dense alignment with position map regression network," in Proceedings of the European Conference on Computer Vision (ECCV),2018, pp. 534–551.
[5] . Deng, J. Yang, S. Xu, D. Chen, Y. Jia, and X. Tong, "Accurate 3D face reconstruction with weakly-supervised learning: From single image to image set," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2019