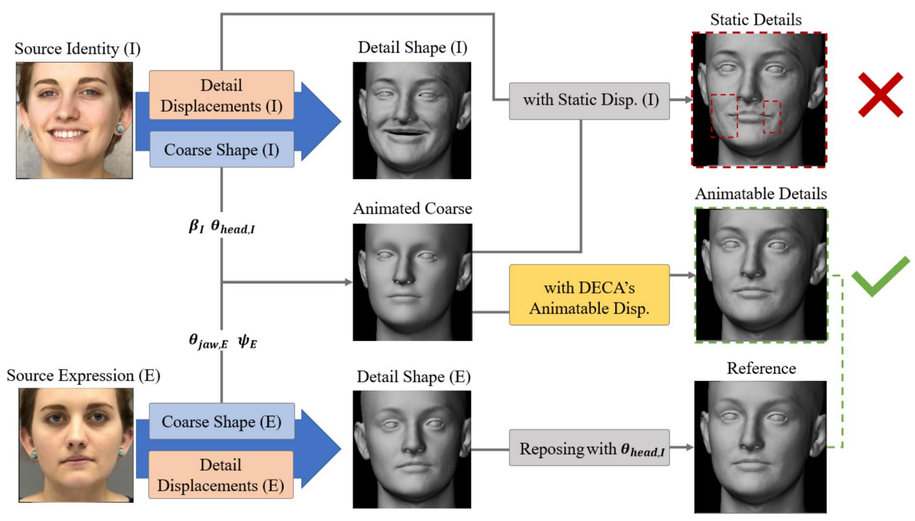
DECA: Learning an Animatable Detailed 3D Face Model from In-The-Wild Images

Yao Feng*[1,2], Haiwen Feng*[1], Michael J. Black[1], Timo Bolkart[1]
[1]Max Planck Institute for Intelligent Systems, Tübingen [2]ETH Zürich

Introduction

Goal

Given a single image of a human face, the goal is to estimate an accurate 3D model of the person's head, with detailed wrinkles, and to animate this face with natural wrinkle deformations.



Problem

Previous methods are able to extract wrinkle details but do not provide a model that can be animated such that the details vary with expression.

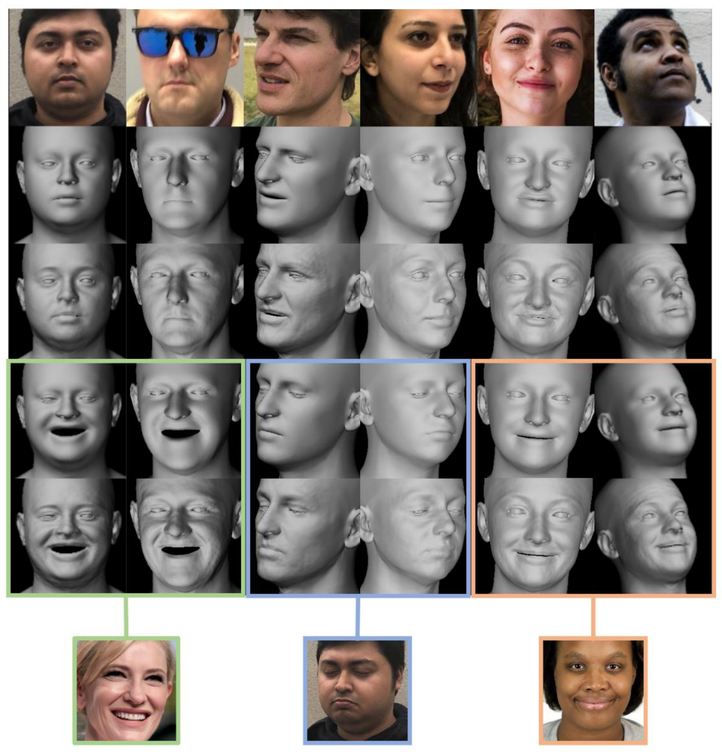
Contribution

The first approach to learn an animatable displacement model from in-the-wild images that can synthesize plausible geometric details by varying expression parameters.

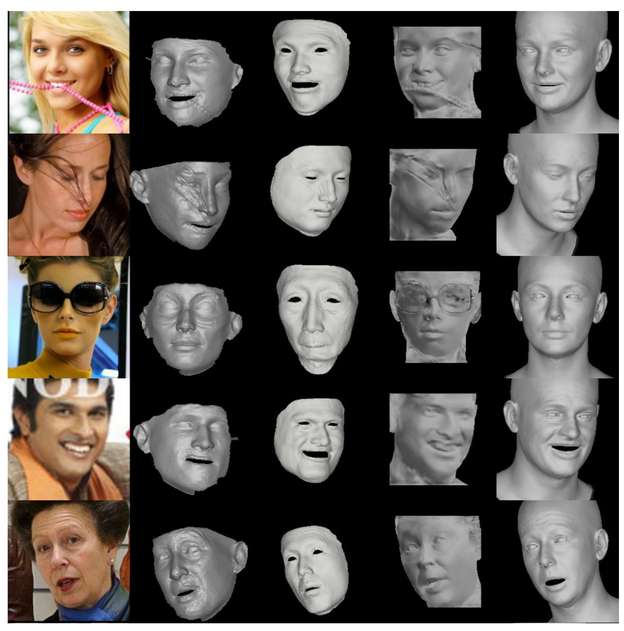
deca.is.tue.mpg.de

Results

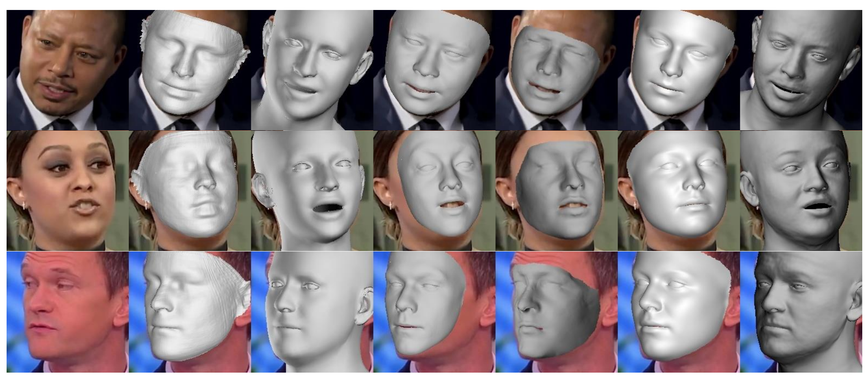
Detailed reconstruction and animation



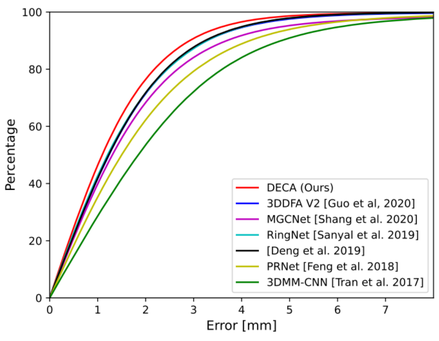
Compare to other detail reconstruction methods



Compare to other coarse reconstruction methods



cumulative error curves on NOW [1]



NoW [Sanyal et al. 2019]

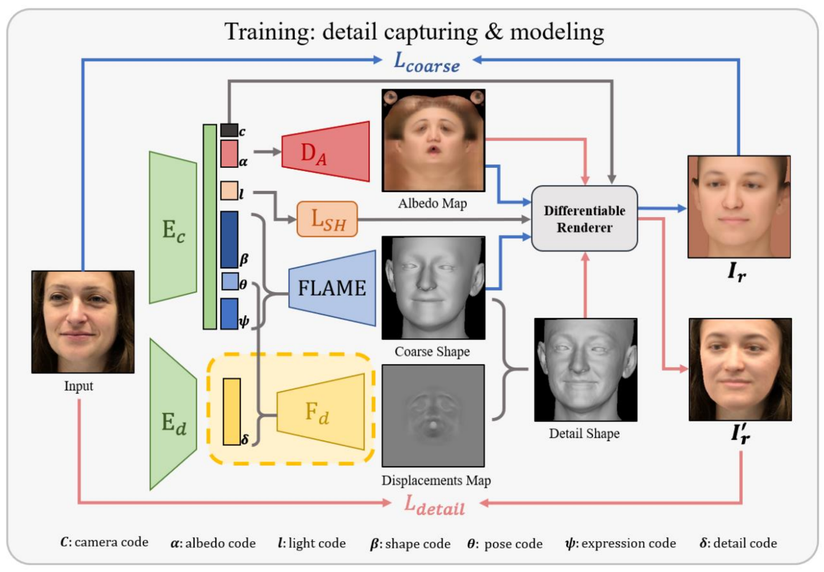
Method	Median (mm)	Mean (mm)	Std (mm)
3DMM-CNN [Tran et al. 2017]	1.84	2.33	2.05
PRNet [Feng et al. 2018b]	1.50	1.98	1.88
Deng et al.19 [2019]	1.23	1.54	1.29
RingNet [Sanyal et al. 2019]	1.21	1.54	1.31
3DDFA-V2 [Guo et al. 2020]	1.23	1.57	1.39
MGCNet [Shang et al. 2020]	1.31	1.87	2.63
DECA (ours)	**1.09**	**1.38**	**1.18**

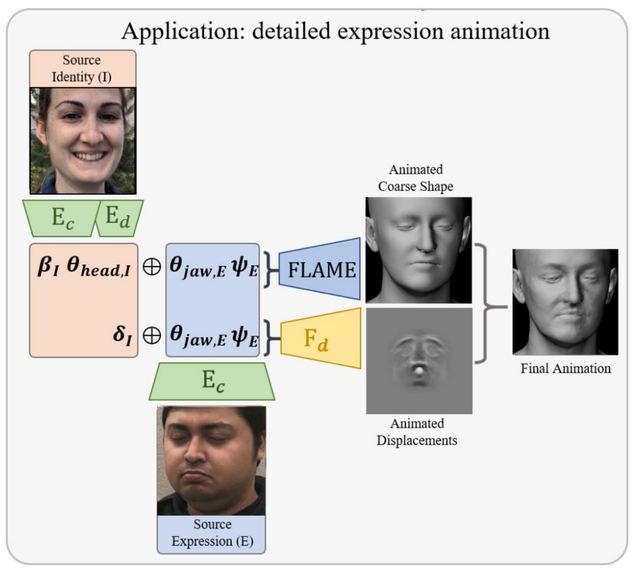
Reconstruction error on the NoW [1] benchmark.

Method

Training and Animation

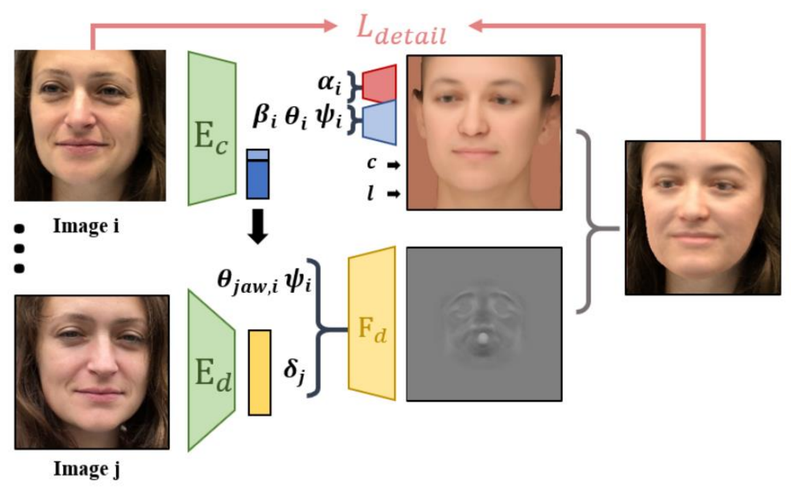
- estimates parameters to reconstruct face shape for each image
- learns an expression-conditioned displacement model by leveraging detail consistency information





Detail consistency

- uses multiple images of the same person during training to disentangle static person-specific details from expression-dependent details.



References

[1] NoW challenge. 2019. https://ringnet.is.tue.mpg.de/challenge
[2] Extreme 3D face reconstruction: Seeing through occlusions, CVPR 2018
[3] FaceScape: a Large-scale High Quality 3D Face Dataset and Detailed Riggable 3D Face Prediction, CVPR 2020
[4] Cross-modal Deep Face Normals with Deactivable Skip Connections, CVPR 2020
[5] Joint 3D Face Reconstruction and Dense Alignment with Position Map Regression Network, ECCV 2018
[6] Learning to Regress 3D Face Shape and Expression from an Image without 3D Supervision, CVPR 2019
[7] Accurate 3D Face Reconstruction with Weakly-Supervised Learning: From Single Image to Image Set, CVPRW 2019
[8] FML: Face Model Learning from Videos, CVPR 2019
[9] Towards Fast, Accurate and Stable 3D Dense Face Alignment, ECCV 2020

Discussions

- Rendering quality is limited by albedo model
- Do not explicitly model facial hair
- DECA uses a weak perspective camera model, for selfies, we would need to extend the method to include focal length
- Future work can extend the model over time, for tracking and learn more personalized model