Fitting 3D Morphable Models using Local Features

Patrik Huber (p.huber@surrey.ac.uk)
Centre for Vision, Speech & Signal Processing
University of Surrey, UK

Supervisor: Prof. Josef Kittler

www.patrikhuber.ch
3D Face Reconstruction From A Single 2D Image

2D input image 3D face representation Applications

\[ \alpha = [\alpha_1, ..., \alpha_N] \]
\[ \beta = [\beta_1, ..., \beta_M] \]

Pose normalisation
Recognition
Analysis
Videos
3D Morphable Models

- 3D scans in dense correspondence
- Apply PCA
  - Shape and albedo (color) model \( M := (\mu, \sigma, U) \)
- New model instances generated by \( S = \mu + \sum_{i}^{M} \alpha_i u_i \)
- Fitting to a 2D image: Find optimal...
  - ...shape- and color model coefficients \( \alpha, \beta \)
  - ...camera and lighting parameters

Data = \([x_0, y_0, z_0, x_1, ...]\)
Existing Fitting Algorithms

• Multiple Features Fitting (Romdhani, Tena, Schönborn):
  • minimise the L2 pixel error
  • uses landmarks, RGB pixel color, edges
  • highly nonlinear problem, Levenberg-Marquardt, MCMC sampling
  • several minutes

• Linear (Smith, Amberg):
  • minimise landmark error for shape-fit, pixel error for rest
  • uses landmarks, RGB pixel color
  • linear, closed-form solutions, iterative
  • order of seconds
• Why not use local features instead of relying on raw pixel values?
  • HoG/SIFT operator not differentiable, hard to optimise
  • Regression based methods
Fitting 3D Morphable Models Using Local Features

Supervised descent / cascaded regression for 2D landmark detection:

• Non-parametric model, learn a shape-update step $\delta x$ as a function of image features... $x = [x_1, y_1, \ldots, x_n, y_n]$

• ...using a series of linear regressors: $\delta x = A_n f(I, x) + b_n$

• Learn these regressors from data. Start from an initial location and then learn the shape-step towards the ground truth location

• Recently proposed to solve for generic vision problems
  • X. Xiong and F. De la Torre, “Supervised Descent Method for Solving Nonlinear Least Squares Problems in Computer Vision”, in submission to TPAMI

• We propose an approach to use it to fit 3D Morphable Models using local features
Fitting 3D Morphable Models Using Local Features

Fitting using cascaded regression & local features:

- Instead of (2D) landmark locations, we learn the 6 DOF and shape parameters: \( \mathbf{R}_n: \delta \theta = A_n f(I, \theta) + b_n \)

- \( \theta = [r_x, r_y, r_z, t_x, t_y, t_z, \alpha_0, \alpha_1] \)

- How does \( f(I, \theta) \) look like?
  - Project the 3D model points to 2D using the current \( \theta \)
  - Extract HoG features at all 2D positions
  - Concatenate them to one vector
Fitting 3D Morphable Models Using Local Features

Input image

Model projection using the current parameter estimates

Local feature extraction regions
Results

Pose estimation:

- Setting: Morphable Model generated renderings, random backgrounds
  - -30° to +30° yaw and pitch variation

- For reference: POSIT (Pose from Orthography and Scaling with Iterations)
  - with ground truth landmarks: average error 1.84°
  - with 5 pixel Gaussian noise: 3.68°
Results

Pose & shape fitting:

• Setting: PIE database
  • Basel Face Model (BFM) fittings as ground truth

• Runtime: ~200ms per image
Conclusions & Future Work

- Promising results so far for pose and shape fitting
- Fits the shape model using robust local features (not only to landmarks), in the order of milliseconds
- Need more «in the wild» training data (shape ground truth hard to obtain)
- The approach unifies landmark detection and 3DMM fitting and can be seen as **landmark detection with a 3DMM prior or landmarks-free 3DMM fitting**
Generic implementation of the supervised descent method: 
https://github.com/patrikhuber/superviseddescent

All infos, slides & link to paper pre-print on arXiv:
www.patrikhuber.ch
Thank you!

Time for questions
References

• Z. Feng, P. Huber, J. Kittler, W. Christmas, X.J. Wu, “Random Cascaded-Regression Copse for Robust Facial Landmark Detection”, SPL 2015

• O. Aldrian and W. A.P. Smith, “Inverse rendering of faces with a 3D morphable model”, PAMI 2013

• X. Xiong and F. De la Torre, “Supervised Descent Method and Its Applications to Face Alignment”, CVPR 2013

• X. Xiong and F. De la Torre, “Supervised Descent Method for Solving Nonlinear Least Squares Problems in Computer Vision”, in submission to TPAMI

• J. R. Tena, “3D Face Modelling for 2D+3D Face Recognition”, PhD thesis, CVSSP, University of Surrey, 2007

• S. Romdhani, T. Vetter, “Estimating 3D Shape and Texture Using Pixel Intensity, Edges, Specular Highlights, Texture Constraints and a Prior”, CVPR 2005