Unconstrained 3D Face Reconstruction

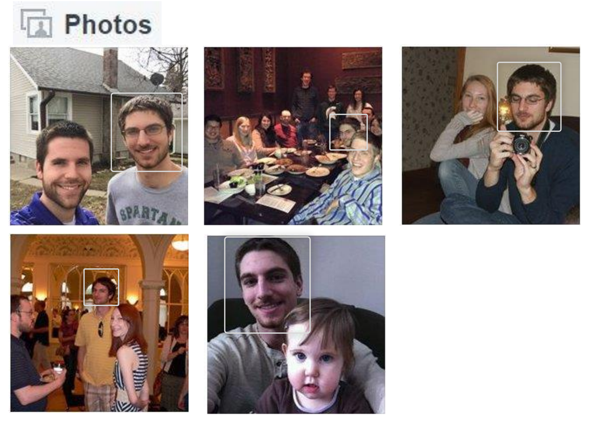
Joseph Roth, Yiying Tong, and Xiaoming Liu (Michigan State University)

Abstract

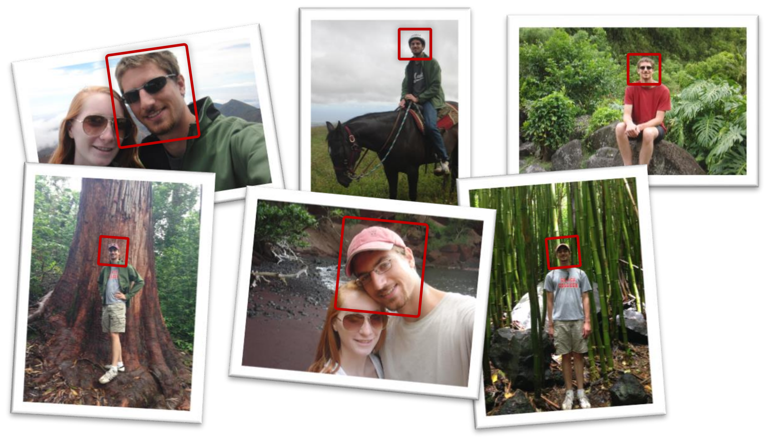
Goal
Given an unconstrained 2D photo collection of an individual with pose, expression, and illumination variations, we propose a method for reconstructing a detailed, watertight 3D surface model of the face.

Input
A photo collection may be gather in a variety of methods.

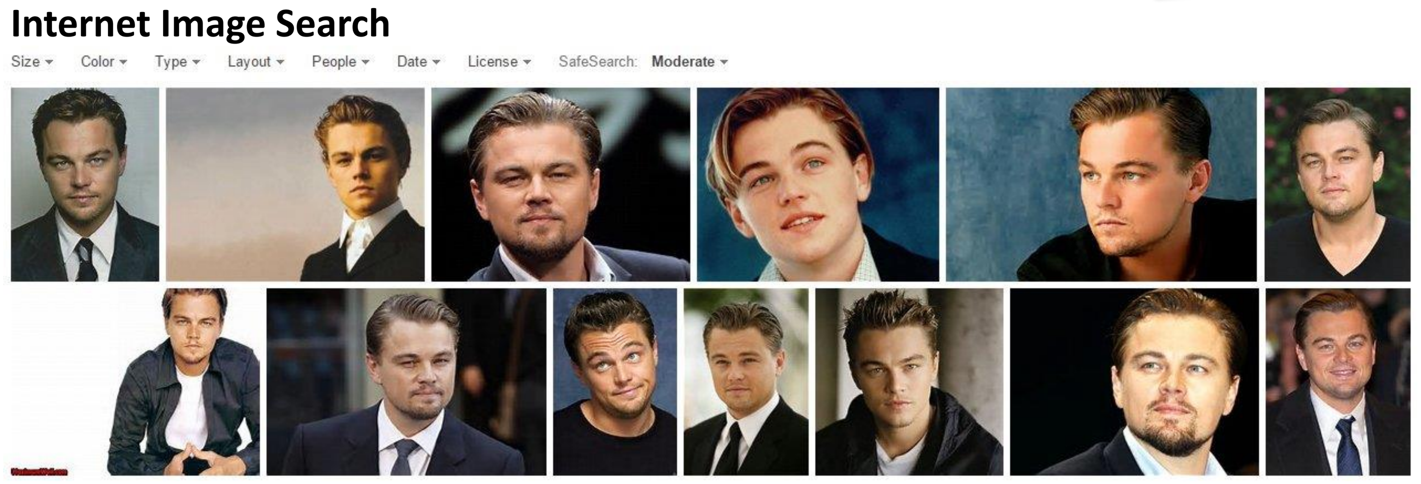
Facebook Tagged Photos



Personal Vacation Photos



Internet Image Search

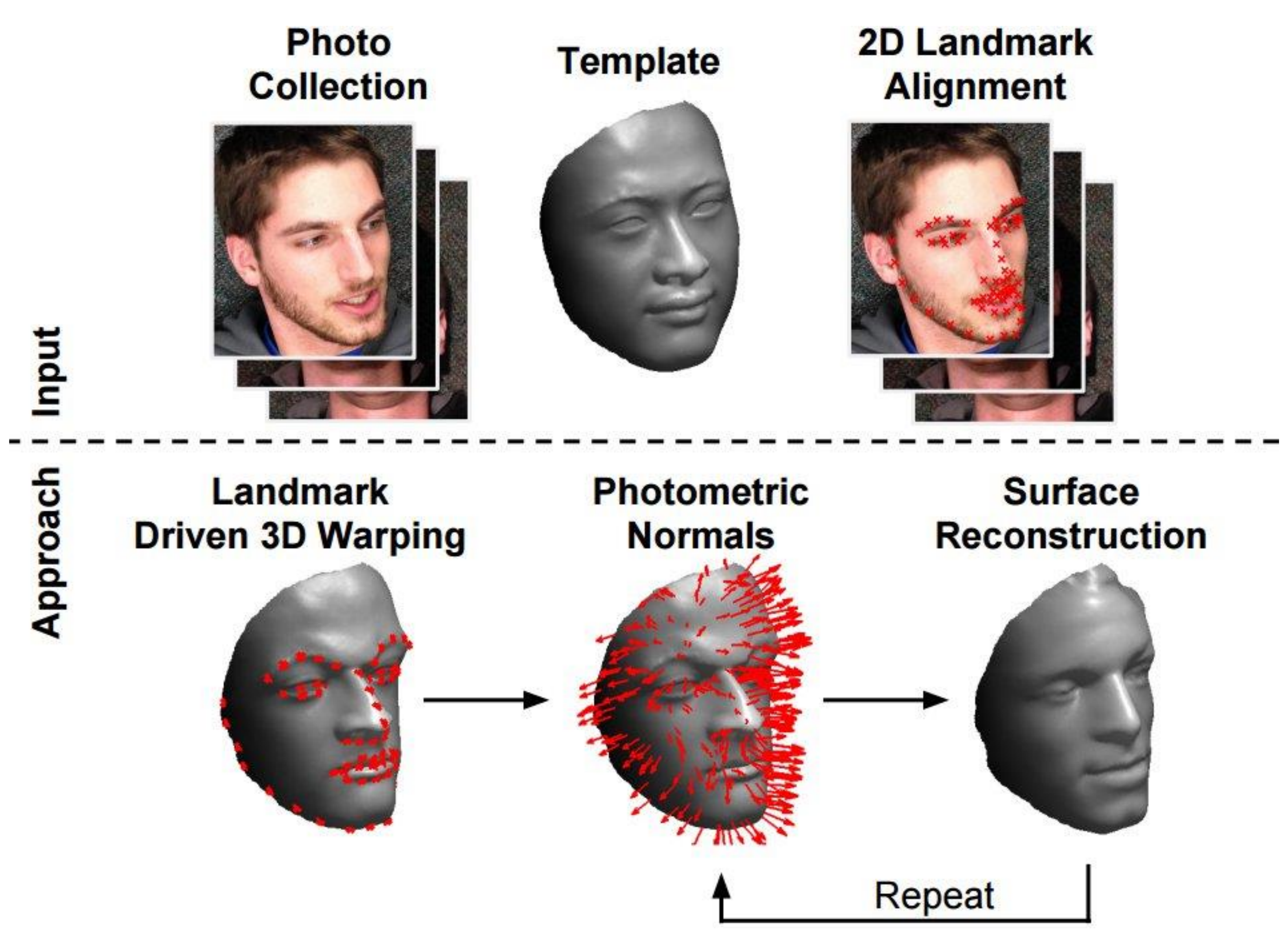


Output
- 3D triangulated mesh.

Applications

- 3D-assisted face recognition (Blanz & Vetter '03, Hu *et al.* '04).
- Facial animation (Cao *et al.* '14).
- 3D expression recognition (Wang *et al.* '06).
- Consumer entertainment, e.g., personalized bobbleheads.

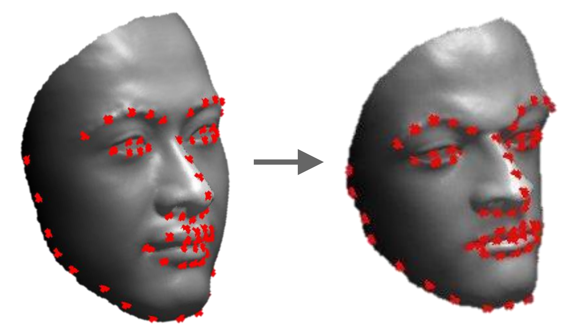
Overview



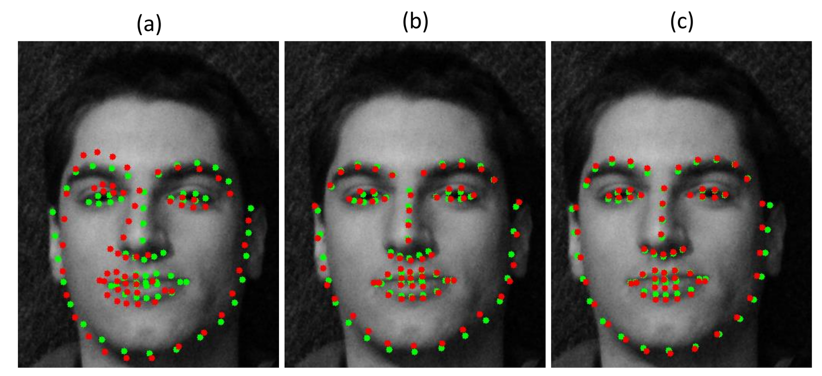
Approach

Landmark Driven 3D Warping
Deform the generic template to match the *overall* structure of the subject.

Takes advantage of profile images.



Which projection (red) best fits the true landmarks (green)? Image (c), but this is not the template, but a subject specific deformation.



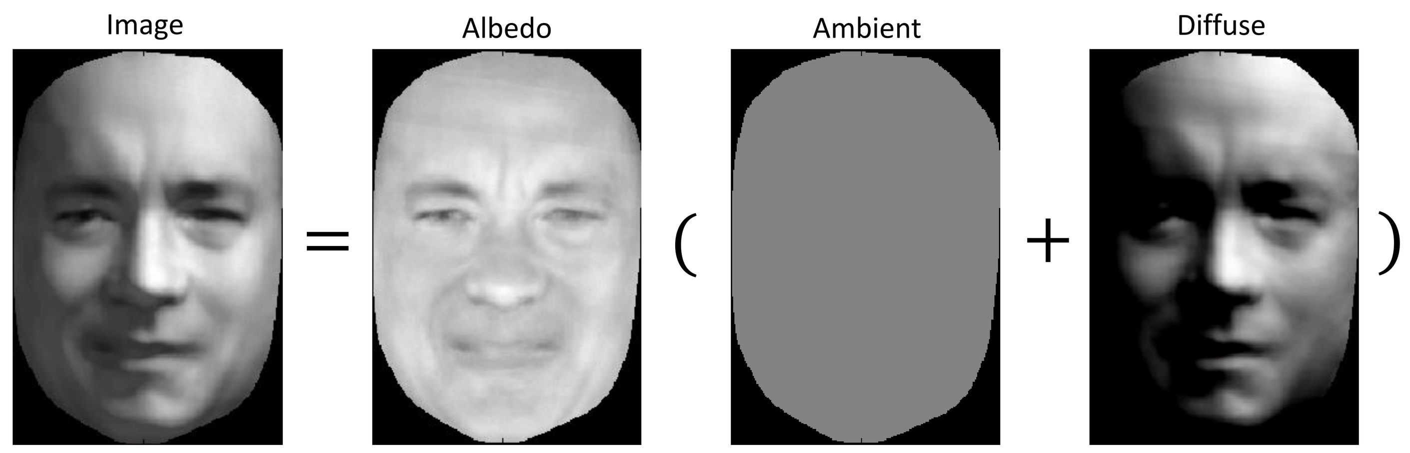
Template Projection

Project 3D template landmarks to best fit the visible 2D landmarks

$$\arg\min_{\mathbf{P}_i} \| \mathbf{P}_i \mathbf{D}_i \mathbf{X} - \mathbf{W}_i \|^2 \quad \text{subject to } \mathbf{P}_i \mathbf{P}_i^{\intercal} = s\mathbf{I}$$

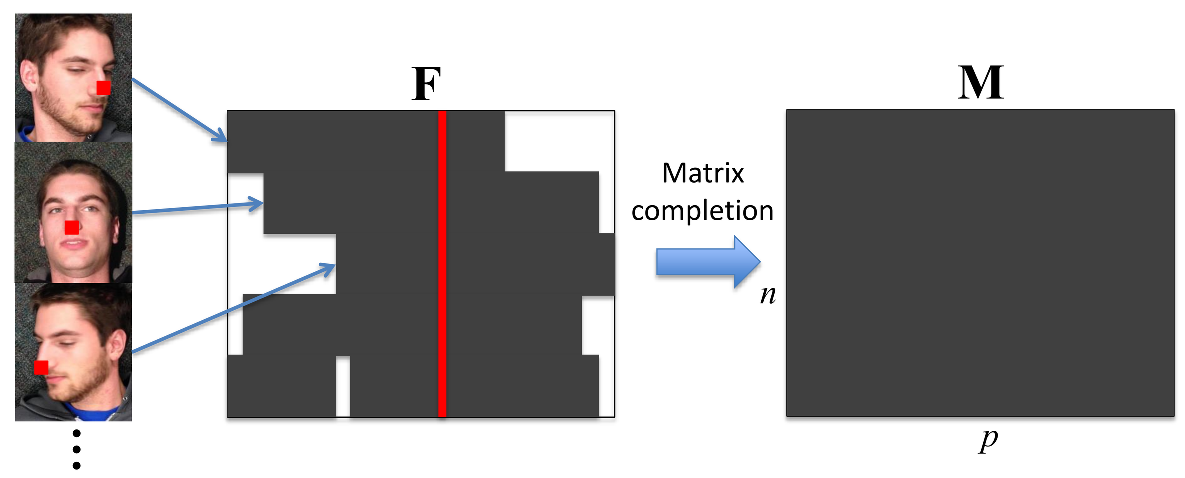
2x3 camera projection — Landmark selection — Template vertices — 2D landmark locations

Photometric Normals
Lambertian reflectance model. Intensity in image is a linear combination of the surface normals weighted by the lighting.

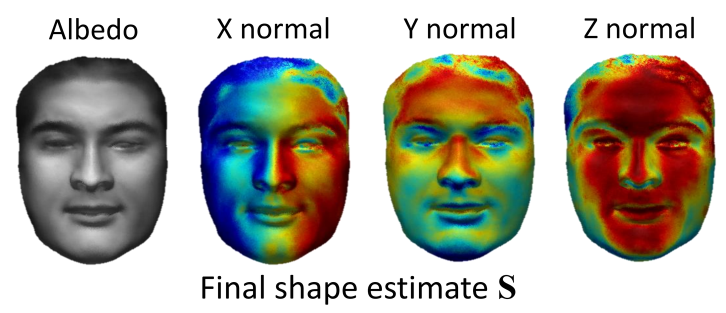


Illumination Correspondence Matrix

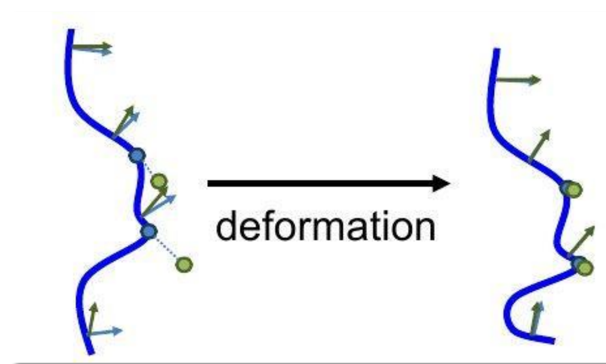
Project template onto each face to find vertex correspondence across all images. Some parts of the face may be obscured in a given image.



Decompose Lighting and Shape
- $\mathbf{M}$ is rank-4 under Lambertian assumption.
- SVD to find light $\mathbf{L}$ and shape $\mathbf{S}$.
- Resolve ambiguity with template.
- Refine estimate based on a subset of images which agree locally.



Surface Reconstruction
- Deform the surface to better match the landmark constraints and the surface normal constraints.
- Additional boundary constraint to maintain consistency.

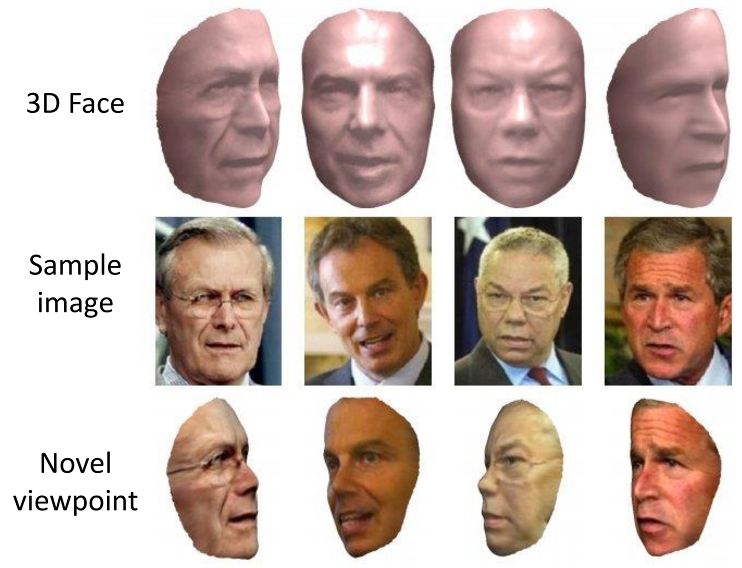


$$X^{k+1} = \arg\min_{\mathbf{X}} \underbrace{\lambda_l \sum_i \| \mathbf{P}_i^k \mathbf{D}_i^k \mathbf{X} - \mathbf{W}_i \|^2}_{\text{Landmarks}} + \underbrace{\| \mathcal{L}\mathbf{X} - H^k \mathbf{n}^k \|^2}_{\text{Surface Normals}} + \underbrace{\lambda_b \| \mathcal{L}_b \mathbf{X} - \mathcal{L}_b \mathbf{X}^k \|^2}_{\text{Boundary}}$$

Results

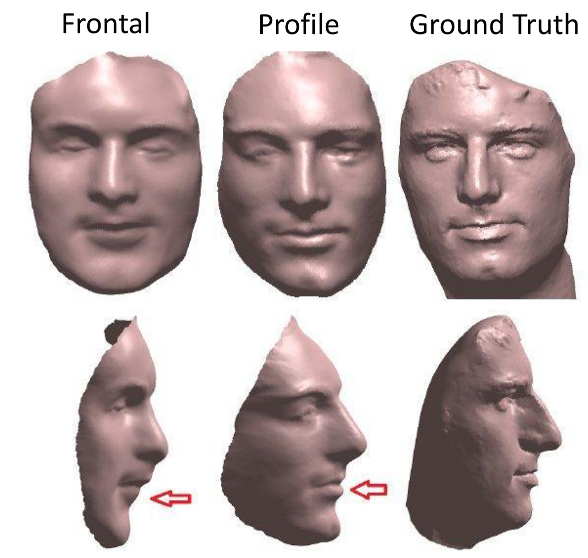
Labeled Faces in the Wild

Popular database for face recognition.



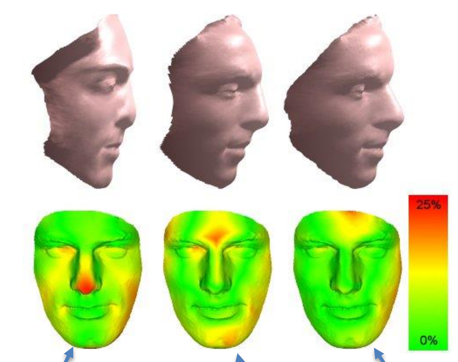
Pose Variation

Benefits from non-frontal images.



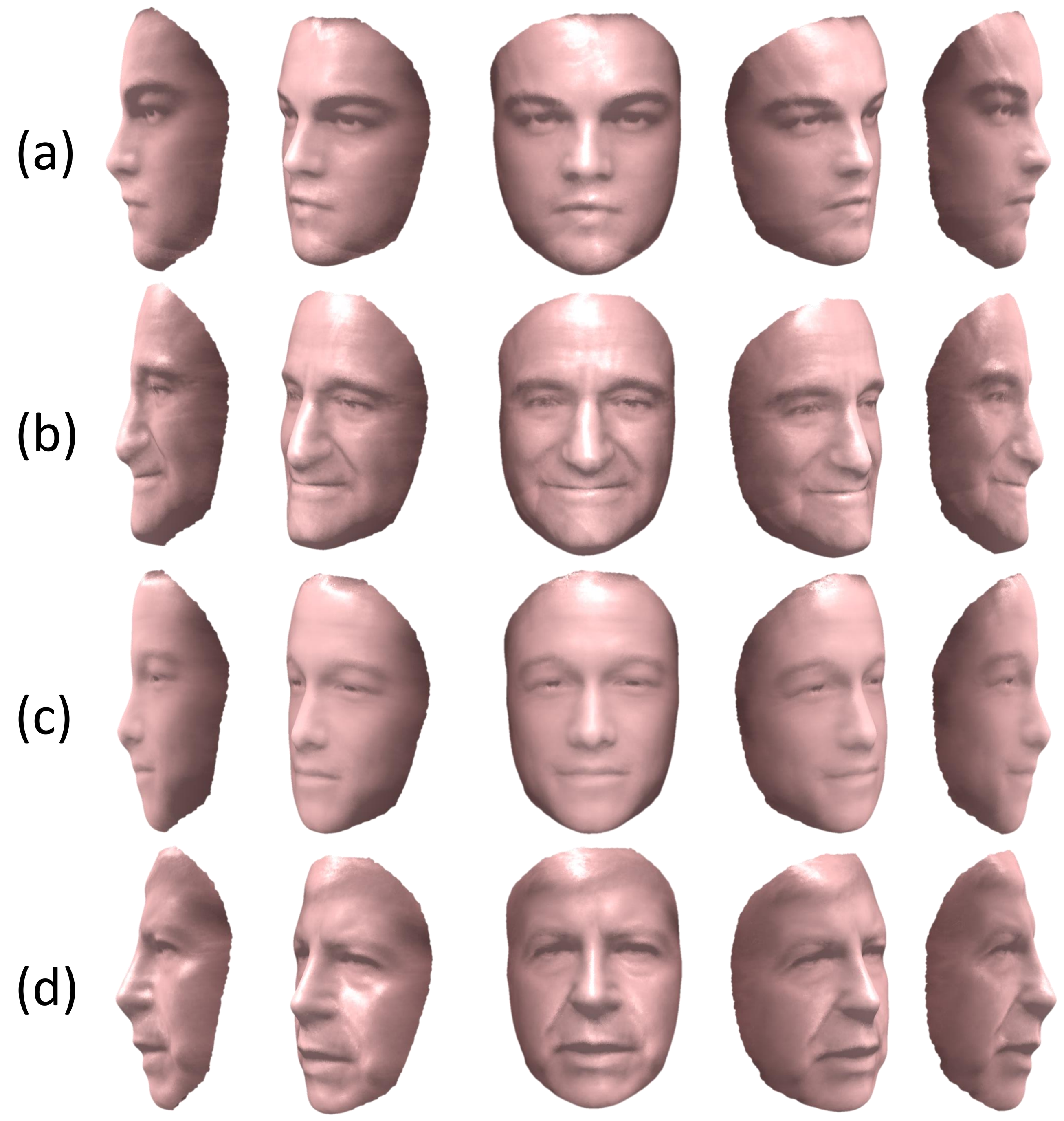
Comparison with Prior Art [2]

2.5D: [2] Warps all images to frontal before integrating depth.
2.5D Improved: Use proposed landmark driven warping before applying [2].
3D: Proposed approach which allows use of non-frontal images.



Methods	2.5D	2.5 Improved	3D
Mean	7.86%	7.79%	**5.42%**
RMS	9.71%	9.04%	**6.89%**

Guess Who?



(a) Leonardo DiCaprio, (b) Robin Williams, (c) Joseph Gordon-Levitt, (d) Rick Snyder

Conclusions

- Presented a method for 3D *face reconstruction* from an *unconstrained* photo collection.
- Deformation of a true 3D surface rather than a simple height field.
- Leverages faces from all possible poses
- Combination of 2D landmark driven constrain and a photometric stereo based normal field.

[1] Joseph Roth, Yiying Tong, Xiaoming Liu, "Unconstrained 3D Face Reconstruction," in Proceedings of IEEE Computer Vision and Pattern Recognition (CVPR 2015), Boston MA, June 7-12 2015.
[2] I. Kemelmacher-Schlizerman and S. M. Seitz. "Face Reconstruction in the Wild," in ICCV, 2011.