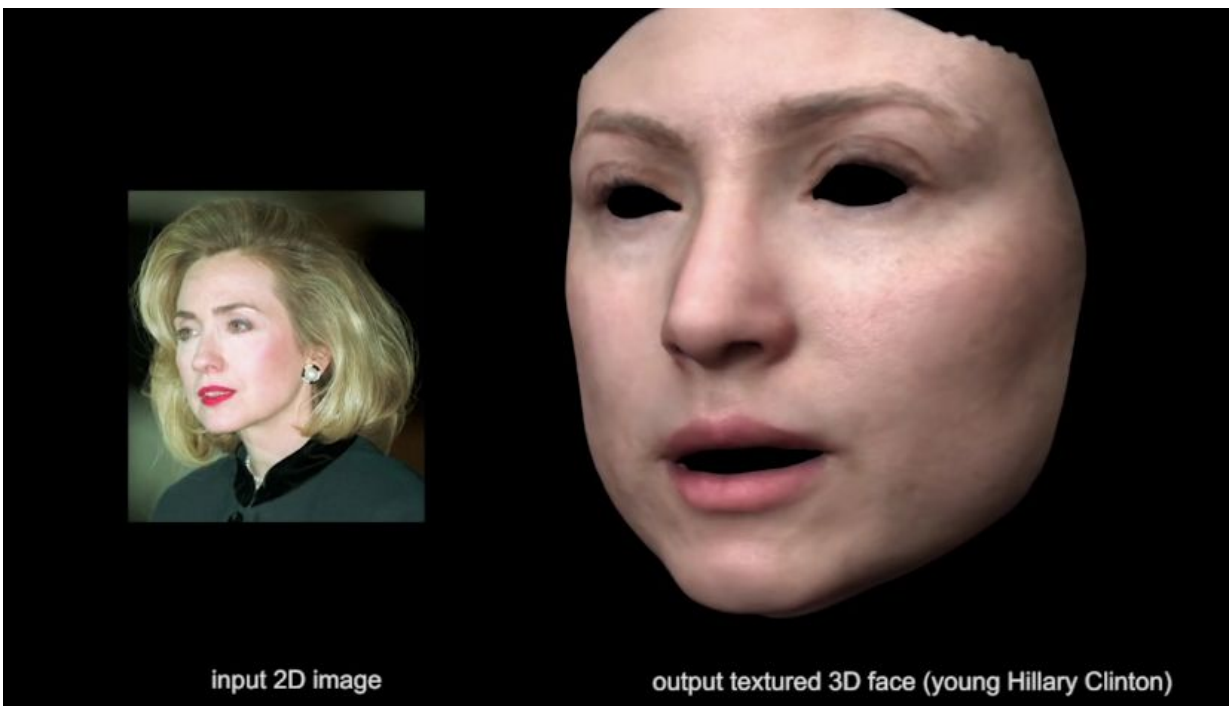


# From a 2-D view comes a 3-D face model using deep neural networks

December 12 2016, by Nancy Owano

---



(Tech Xplore)—Whether it's for gaming, entertainment, or education purposes, you can expect to see researchers aiming for techniques to bring you high visual impact in the virtual space, now that we are becoming increasingly linked to living and learning online.

Recent news is that a detailed 3-D face can be created just from a single picture. You input a 2-D image and the output is a good, textured 3-D face model.

Sidney Fussell in *Gizmodo* said the researchers' results were "remarkably accurate." The team is from the University of Southern California and USC Institute for Creative Technologies.

They have their paper on arXiv and it is [titled](#) "Photorealistic Facial Texture Inference Using Deep Neural Networks." The authors are Shunsuke Saito, Lingyu Wei, Liwen Hu, Koki Nagano and Hao Li.

Dunja Djudjic in *DIYPhotography* made the point that in general detailed and accurate face mapping is a complex task. "It requires a series of photos with ideal and consistent lighting from different angles. If you want to capture all the details and imperfections of the face, you need professional lighting and multiple [shots](#)."

Fussell in *Gizmodo* similarly noted that "Face mapping at this level usually requires a series of photos in ideal lighting to make sure you get all the curves, angles, and asymmetries of the [face](#)."

Earlier this month, the USC Institute for Creative Technologies had an item, "This Neural Network Creates 3-D Faces from Blurred Photos," calling attention to what they did, using a [video game](#) as an example.

When actors want to star in a video game, they said, "the process is incredibly complex. For this, the celebrities are scanned with dozens of cameras and scanned by [scanners](#), that map every square meter of their face."

They said the research group "believes to be able to change that."

This group has their own method. They use [deep neural networks](#) to create detailed and quite accurate 3-D models. Beau Jackson in *3-D Printing Industry* referred to their project as Annotated Faces-in-the-Wild, and said, "Saito et al. define 3-D models with depth based on light and [shadow](#), even detail the pores of the [faces](#)."

Their paper clearly shows the input and output results of faces. The authors described their approach as "a data-driven inference method."

The authors said that their proposed neural synthesis approach could handle high-resolution textures, "which is not possible with existing deep learning frameworks." They used a face database for making inferences.

According to the paper, their inference technique produced high-resolution texture maps with complex skin tones and mesoscopic-scale details (pores, stubble hair), even from very low-resolution input images.

They said in their paper that "fitting a convex combination of feature correlations from a high-resolution face database can yield a semantically plausible facial detail description of the entire face."

They said that then a photorealistic texture map can be synthesized "by iteratively optimizing for the [reconstructed](#) feature correlations."

So how could their work be used in the real world? Consider that with virtual and augmented reality becoming the next-generation platform for social interaction, "compelling 3-D avatars," as the authors worded them, could be generated.

**More information:** Photorealistic Facial Texture Inference Using Deep Neural Networks, arXiv:1612.00523 [cs.CV]  
[arxiv.org/abs/1612.00523v1](https://arxiv.org/abs/1612.00523v1)

## Abstract

We present a data-driven inference method that can synthesize a photorealistic texture map of a complete 3D face model given a partial 2D view of a person in the wild. After an initial estimation of shape and low-frequency albedo, we compute a high-frequency partial texture map, without the shading component, of the visible face area. To extract the fine appearance details from this incomplete input, we introduce a multi-scale detail analysis technique based on mid-layer feature correlations extracted from a deep convolutional neural network. We demonstrate that fitting a convex combination of feature correlations from a high-resolution face database can yield a semantically plausible facial detail description of the entire face. A complete and photorealistic texture map can then be synthesized by iteratively optimizing for the reconstructed feature correlations. Using these high-resolution textures and a commercial rendering framework, we can produce high-fidelity 3D renderings that are visually comparable to those obtained with state-of-the-art multi-view face capture systems. We demonstrate successful face reconstructions from a wide range of low resolution input images, including those of historical figures. In addition to extensive evaluations, we validate the realism of our results using a crowdsourced user study.

© 2016 Tech Xplore

Citation: From a 2-D view comes a 3-D face model using deep neural networks (2016, December 12) retrieved 3 May 2024 from <https://techxplore.com/news/2016-12-d-view-deep-neural-networks.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.