

## **UC Merced**

### **Proceedings of the Annual Meeting of the Cognitive Science Society**

#### **Title**

Inverse Rendering Best Explains Face Perception Under Extreme Illuminations

#### **Permalink**

<https://escholarship.org/uc/item/5jv5t06f>

#### **Journal**

Proceedings of the Annual Meeting of the Cognitive Science Society, 42(0)

#### **Authors**

Egger, Bernhard

Siegel, Max

Arora, Riya

et al.

#### **Publication Date**

2020

#### **Copyright Information**

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Inverse Rendering Best Explains Face Perception Under Extreme Illuminations

**Bernhard Egger**

MIT, Cambridge, Massachusetts, United States

**Max Siegel**

MIT, Cambridge, Massachusetts, United States

**Riya Arora**

Computer Science and Artificial Intelligence Laboratory, MIT, Cambridge, Massachusetts, United States

**Amir Soltani**

MIT, Cambridge, Massachusetts, United States

**Ilker Yildirim**

Yale University, New Haven, Connecticut, United States

**Josh Tenenbaum**

MIT, Cambridge, Massachusetts, United States

## Abstract

Humans can successfully interpret images even when they have been distorted by significant image transformations. Such images could aid in differentiating proposed computational architectures for perception because while all proposals predict similar results for typical stimuli (good performance), they differ when confronting atypical stimuli. Here we study two classes of degraded stimuli – Mooney faces and silhouettes of faces – as well as typical faces, in humans and several computational models, with the goal of identifying divergent predictions among the models, evaluating against human judgments, and ultimately informing models of human perception. We find that our top-down inverse rendering model better matches human percepts than either an invariance-based account implemented in a deep neural network, or a neural network trained to perform approximate inverse rendering in a feedforward circuit.