Automated Restyling of Human Portrait Based on Facial Expression Recognition and 3D Reconstruction

Cheng-Han(Dennis) Wu
Department of Electrical Engineering, Stanford University

Motivation

Experimented automatic restyling of an ordinary photo to a self portrait with effects that correspond to his/her facial expression.
This project wants to showcase and inspire audiences the capabilities of AI in visual effects automation in photography or film making.

Related Work

1. Depth Based Visual Effect
Matthias Ziegler, Andreas Engelhardt, Stefan Mller, Joachim Keinert, Frederik Zilly, Siegfried Foessel. Multi-camera system for depth based visual effects and compositing. CVMP, 2015.

2. Facial Relighting
Yang Wang, Lei Zhang, Zicheng Liu et.al. Face Relighting from a Single Image under Arbitrary Unknown Lighting Conditions. IEEE Transactions on Pattern Analysis and Machine Intelligence, Nov. 2015.

3. RGB-D Registration
Lembit Valgma. 3D reconstruction using Kinect v2 camera. Bachelor's thesis (12 ECTP). 2016

4. Expression Recognition
Yu, Zhiding, and Cha Zhang. Image Based Static Facial Expression Recognition with Multiple Deep Network Learning - Microsoft Research. Microsoft Research. IEEE, Nov. 2015.

System Architecture

Kinect
- RGB & Depth Calibration
- RGBD Registration/Denoising
- Face Recognition/Tracking

Expression
- CNN Model Training
- Expression Recognition

Restyling
- 3D Face Model Calculation
- Face Image Cropping/Masking
- Shading/Relighting/Restyling

Kinect

1. Kinect for Windows V2 Interface:

➢ Libfreenect2/OpenCV on MacOSX

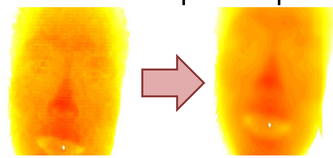
2. RGB/Depth Camera Calibration/Registration

➢ Undistort depth cam with camera matrix and distortion coefficients and project onto RGB

$$\begin{bmatrix} fx = 365.45 & 0 & cx = 254.87 \\ 0 & fy = 365.45 & cy = 205.39 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} k1 = 0.0905474 \\ k2 = -0.26819 \\ k3 = 0.0950862 \\ p1 = 0.0 \\ p2 = 0.0 \end{bmatrix}$$

3. Depth Denoising

➢ Bilateral filter depth map



4. Face Tracking/ Locating ROI

➢ OpenCV Haar-Cascade

Expression Recognition

1. Dataset/Preparation

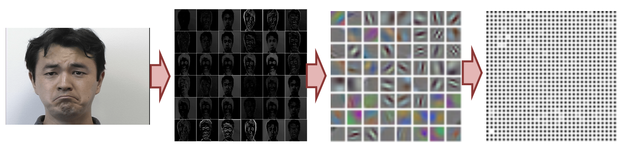
This part collaborated with Hsin Chen(hsinc@stanford.edu)

➢ Cohn-Kanade(CK+) dataset
➢ 5876 images, 8 Emotions

0 neutral
1 angry
2 contempt
3 disgust
4 fear
5 happy
6 sad
7 surprise

2. Convolutional Neural Network

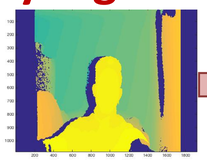
➢ Modify and fine tune a Googlenet model with 8 expression categories. Trained using Caffe on the farm, Matcaffe as local interface.

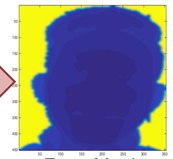


Shading/Restyling

1. Cropping ROI and Image Masking

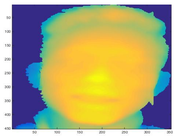
➢ Crop with OpenCV tracked face ROI
➢ Mask with dilated, hole-filled, and thresholded depthmap.

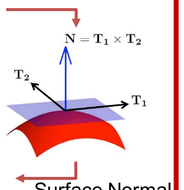

Original Depth


Face Mask

2. Shading/Relighting

➢ Calculate surface normal from depth.
➢ Calculate directional lighting intensity and color for the face.


Depth



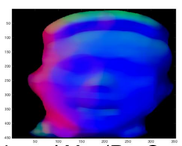
Surface Normal Calculation


Original


Left(-x) **blue** light

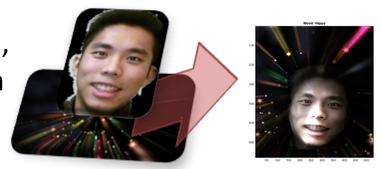

Top(+y) **red** light


Front(+z) **white** light


Normal Map(R:x G:y B:z)

2. Restyling

➢ Apply mood lighting and background, processed using above mask applied on both images.

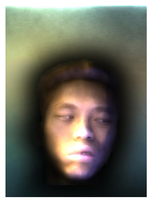


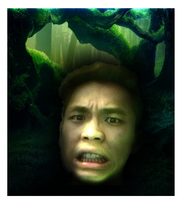
Experimental Results

1. 8 Emotion Effects

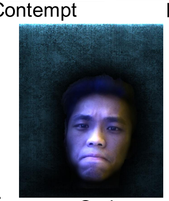

Neutral

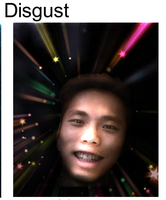

Angry


Contempt


Disgust


Fear


Sad


Happy


Surprise

2. CNN accuracy

➢ Around 60%, drop when there are large background Intensity variation or object. Fed in cropped data to improve.

Future Work

1. Lighting improvements with surface textile characteristic calculation. **2.** Improve expression recognition accuracy. **3.** Improve bilateral filtering efficiency and system efficiency.