Automated Restyling of Human Portrait Based on Facial Expression Recognition and 3D Reconstruction
## EE368 Project Proposal

Cheng-Han(Dennis) Wu(chw0208@stanford.edu)

October 21, 2016

[Note: This project combines with Hsin Chen(hsinc@stanford.edu)'s CS229: Machine Learning Project on the machine learning parts in the project. There are two major sections with image processing, one is the facial 3D depth information extraction, and the other one is portrait restyling using the depth data. These parts will be done solely by Dennis Wu and expression learning will be done jointly. No Android devices will be used, but will have to borrow **Microsoft Kinect**]

1   Project Description

The main goal for this project is to automatically restyle a plane human portrait by recognizing his/her facial expression. By utilizing human expression recognition through machine learning, we will be able to perform techniques such as graphical relighting and background modulation based on the color and depth images collected. Thus turning an ordinary human portrait into a dramatic photo with cinematic effects.
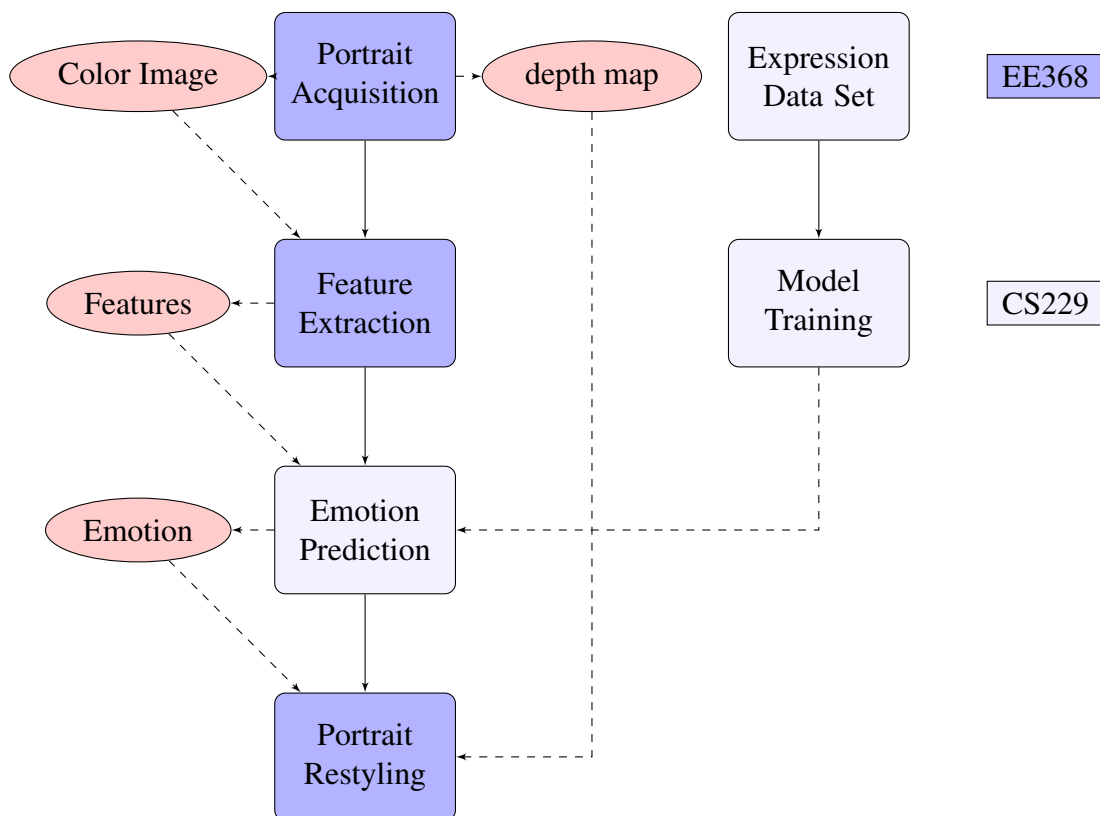
1.1 Depth Data Acquisition

We plan to use Microsoft Kinect to capture half body portraits of the user. We use not only color but depth information for restyling purposes.

1.2 Facial Feature Extraction

With the photos of facial expression collected as our dataset, we then extract the features from them for later model training. This process can be done in several ways, one involves filtering certain colors of the pictures to obtain the shape of facial features (such as mouth, eyes, and eyebrows), then calculating their angle, momentum, or other characteristics. The results can be used as our training features. Using Scale-invariant feature transform (SIFT) algorithm is another way to collect the feature descriptions.In this project, We plan to integrate existing approches and develop our own feature collecting stategy. The relevant research papers are listed below.

1.3 Expression Recognition(CS229)

There are a variety of methods used to achieve the goal of facial expressions recognition, in this project, we plan to use convolutional neural network as our main approach. We choose CNN for its great performance shown in existing works [4]. Also, [5] provided a framework of combining multiple CNN models to boost the facial expression predicting result, and we aim to follow similar method to seek for further improvements of our model. On top of the techniques above, we found in [6] a new way to build neural network with a few examples only, and we will try to integrate this method into our emotion recognition process.

1.4 Portrait Restyling

By utilizing the depth map of the portrait, the system is able to mimic real world lighting condition instead of a plain filter applied to the image. By structuring different color temperature lighting as well as partial filters, the system will be able to automatically reconstruct a dramatic portrait from a plan portrait. As described in [1] and [2], a 3D facial model obtained by Kinect and be used to calculate rendering of light sources from various directions. Relighting style will be similar to [3] while incorporating a expression based automatic restyling.

2 Timeline

Date	Goal
October 21	Project Proposal
October 28	Finish paper survey and data collection
	Start extracting/collecting features
November 11	Finish data processing and feature extraction
	Start training data
November 18	Milestone Report
November 25	Get first training done and start tweaking parameters & features
December 1	Finish model training
December 7	Project wrap up and Poster
December 9	Final Writeup

References

[1] 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 (Volume: 31, Issue: 11, Nov. 2009)

[2] Zhen Wen, Zicheng Liu, Thomas S. Huang. *Face Relighting with Radiance Environment Maps.* CVPR, 2003.

[3] Matthias Ziegler, Andreas Engelhardt, Stefan Mller, Joachim Keinert, Frederik Zilly, Siegfried Foessel. *Multi-camera system for depth based visual effects and* compositing CVMP, 2015.

[4] Fasel, B. *Robust Face Analysis Using Convolutional Neural Networks.* Object Recognition Supported by User Interaction for Service Robots (2001): 1-48.

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

[6] Vinyals, Oriol, Blundell, Charles, Lillicrap, Timothy, Kavukcuoglu, Koray, and Wierstra, Daan. *Matching networks for one shot learning.* arXiv preprint arXiv:1606.04080, 2016.