

FACIAL MOVEMENT BASED PERSON AUTHENTICATION

Pengqing Xie

Yang Liu (Presenter)

Yong Guan

Iowa State University

Department of Electrical and Computer Engineering

OUTLINE

- Introduction
- Literature Review
- Methodology
- Experiment
- Conclusion and Future Work

INTRODUCTION

- Face recognition has received significant attention these days especially in the field of authentication.
- Most of the face recognition techniques are achieved by extracting static facial features such as eyes vertical position, mouth width, nose height, etc. (Figure 1.)
- Can we exploit dynamic features during recognition?

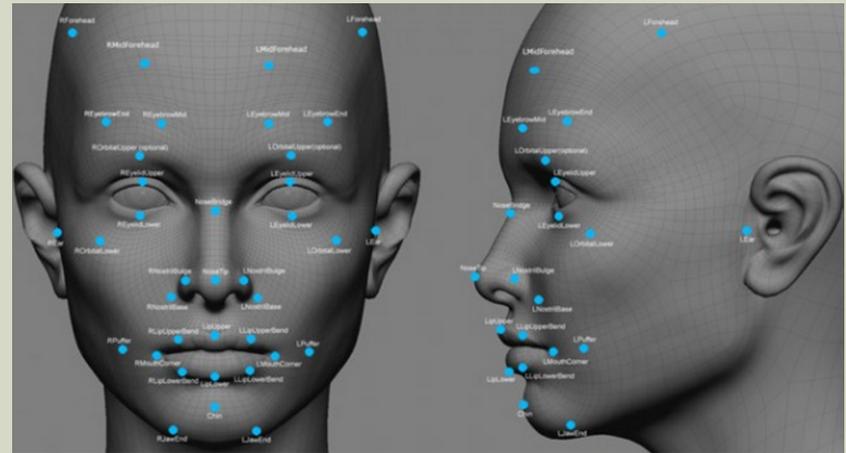


Figure 1. Facebook's Deep face Project

INTRODUCTION

- Brain activity studies revealed that for face perception, not only face-specific brain areas were involved, but a coherent neural activity was devoted to motion perception and gaze control.
- In addition to basic static features related to shape and color, a “dynamic signature” can also be utilized to augment face recognition.
- We believe that dynamic features represent facial behaviors which make them impossible to be replicated and thus more distinctive.

INTRODUCTION

- Since most face recognition algorithms only analyze static features, they cannot distinguish between a real face and a counterfeited face.
- Researchers had shown that the face recognition software offered on laptops from Lenovo, Toshiba and Asus can be fooled by manipulating a digital image.
- This is known as the playback attack which can be avoided if 3D dynamic features were employed.



Figure 2. Face recognition software detecting a face displayed on a computer screen.

LITERATURE REVIEW

- **Dynamic face recognition**
 - Clusters facial expressions into Hidden Markov Models (HMMs) and estimates the emission probabilities of each state
 - Not applicable to our approach since we are interested in consecutive facial movements
- **Face liveness detection against playback attack**
 - Considers movements of eyes, eyebrows, nose, mouth, and lip
 - 2D-based method will still suffer from video-based spoofing attack
 - Our approach utilizes “universal” facial movements, thus eliminating the need of specialized liveness detection unit.

METHODOLOGY

- A 3D face model (CANDIDE-3) consisting of shape units (SUs) is constructed to characterize facial features (Figure 3.)
- A facial expression is decomposed into action units (AUs) as defined by the Facial Action Coding System (FACS) which encodes individual facial muscle movement from changes in facial appearance
- Each AU lies within the range of -1 to +1, thus face recognition is reduced to matching two time series.

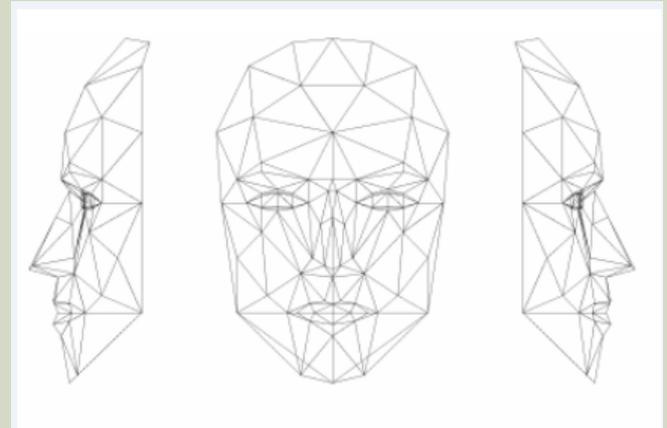


Figure 3. The CANDIDE-3 3D face model.

METHODOLOGY

- **Longest Common Subsequence (LCSS)** is elected for matching time series
 - Compared to other methods (Euclidian distance and dynamic time warping) which are sensitive to noise and time variance, LCSS is able to achieve a more robust result
 - LCSS offers flexibility in matching by setting the time and space thresholds
- **The authentication scheme can be described as:**
 - **Enrollment:**
 - Based on LCSS similarity measurements, a model is computed from a certain number of AU trajectories to represent the identity model.
 - Time and space thresholds are adaptively computed and stored along with the model to guarantee robustness.
 - **Validation:**
 - The incoming AU trajectories are collected and compared with existing legitimate models, similarities resulted by each AU trajectory matching are combined to determine the final authentication decision

EXPERIMENT

- Five subjects were invited to participate in our experiment. Each subject was instructed to pronounce “WE” 30 times in front of Kinect camera within 4 seconds
- For each sample, six action units were extracted where each of them forms a time series.
 - 20 samples were used for enrollment
 - 10 samples were used for cross-validation
- For each AU, a time series model was constructed and stored.
 - For authentication, six LCSS similarities obtained from each AU were measured and fused together to give a combined result

EXPERIMENT

■ Preliminary experiment results:

- 30 samples of six AUs were acquired from each subject. Figure 4 shows the Lip Stretcher(LS) obtained from one subject.
- Figure 5 shows the combined similarity result for each subject. Subject 1 was chosen to be the probe model, and was compared with itself with the 10 samples used for cross-validation (Figure 5.)

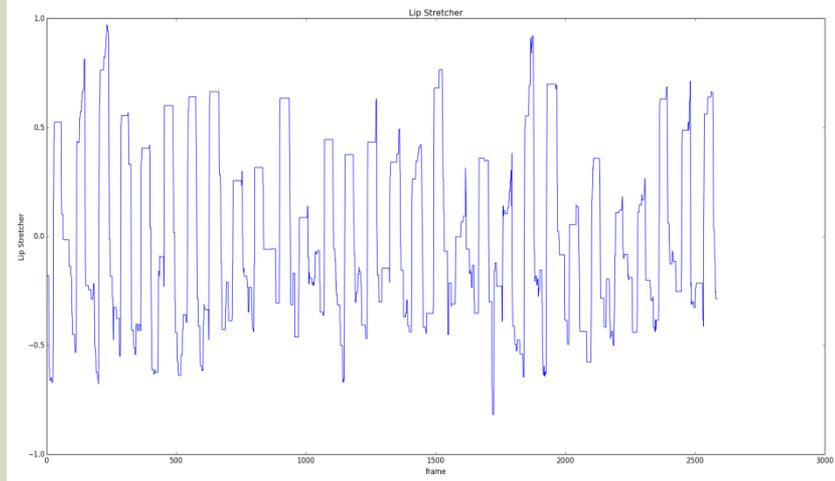


Figure 4. 30 samples of LS obtained from one subject.

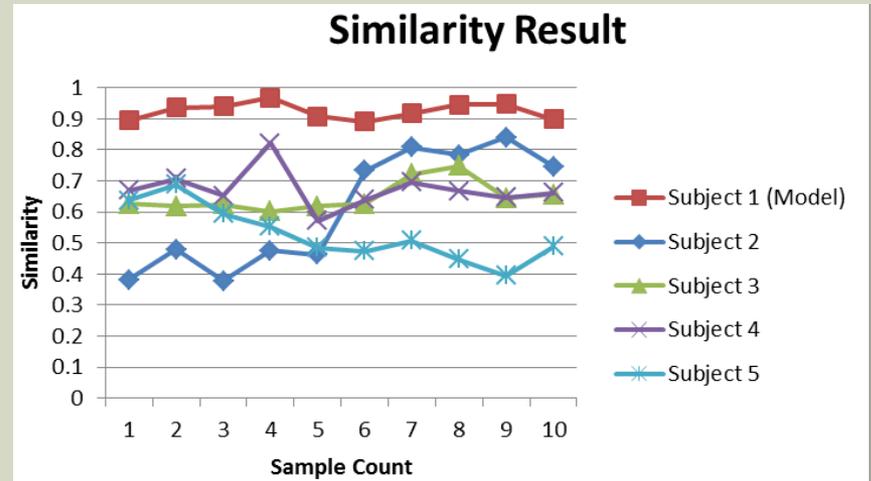


Figure 5. 10 combined similarity measurement for each subjects.

CONCLUSION AND FUTURE WORK

- We proposed an authentication method using spatial temporal facial movements. We had seen how it could prevent playback attack by employing 3D dynamic features.
- Our preliminary experiment had shown promising results.
- Since the facial expression in our experiment is carried out artificially, it is believed that up slopes (neutral to maximal) facial movements may be similar for all subjects whereas down slopes (maximal to neutral) are distinctive.
- Our future work involves developing an automated facial authentication system to be used in most security systems.

QUESTIONS?