Human Facial Behavior Analysis for Human Computer Interaction

Zhiwei Zhu, Yan Tong, Wenhui Liao and Qiang Ji
Intelligent Systems Laboratory

Introduction

➢ **Goal:**
The human face is a rich and powerful source of communicative information about human behavior. Using a video camera together with a set of computer vision techniques, vision-based human sensing technologies have been trying to equip the computers with the capability of perception, seeing, and sensing the user's affective states, intentions and needs from a set of non-verbal cues.

➢ **Advantage:**
Non-intrusiveness and naturalness

➢ **Vision-based human sensing technologies developed in this project:**
- ❖ eye detection and tracking
- ❖ eye gaze tracking
- ❖ facial feature tracking
- ❖ head gesture analysis
- ❖ facial expression analysis
- ❖ facial component movement recognition

1

Real-time Eye Detection and Tracking

(a) The IR LED Illuminators and the CCD Video Camera

(b) Face image with bright pupils

Existing active IR based Eye Tracking are simple and effective, but requires distinctive bright pupil effect to work well and fails when:
- ❖ eye closure
- ❖ eye occlusion due to face rotation
- ❖ illumination interference
- ❖ glasses glare and etc.

2

Proposed Eye Detection and Tracking Algorithm

➢ **Eye detection:**
--- SVM eye classifier is employed to verify the eye images with bright pupil.

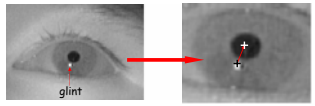
➢ **Eye tracking:**
--- Kalman filtering and mean-shift alternate. Mean shift eye tracking activates when kalman filtering fails.
--- Kalman filtering based pupil tracking provides accurate eye model for the mean-shift tracking, therefore avoids drifting away.

An image sequence demonstrates the combined eye tracker: white rectangles show the tracking results by Kalman filter, and black rectangles show the results by mean-shift.

3

Eye Gaze Tracking – 2D Mapping-based Techniques

Eye gaze is the line of sight of a person, representing a person's focus of attention. It has the applications in Human Computer Interaction, Virtual Reality, Eye disease diagnosis, Human Behavior Studies, etc.



✓ A 2D mapping function between the pupil-glint vector v and the gaze point S_{gaze} on the screen is used to obtain the eye gaze point directly:

$$S_{gaze} = f(v)$$

Challenges:
❖ Require a static head to work well.
❖ Require a calibration procedure for each person each time, as well as recalibration whenever the head moves.

4

Eye Gaze Tracking – Head Movement Compensate

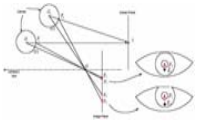
➢ **Dynamic head movement compensation model**:

$$v_2' = g(v_2, O_1, O_2)$$

v_2' is the corrected pupil-glint vector by eliminating the head movement

Then, if the gaze mapping function f_{O1} at position O_1 is known, the gaze of the eye at position O_2 can be estimated as:

$$S_{gaze} = f_{O1}(v_2') = f_{O1}(g(v_2, O_1, O_2)) = F(v_2, O_2)$$



Top view of cornea movement

➢ **Advantages:**
- ❖ Accurate, comfortable, reliable, with free head movement;
- ❖Only one-time calibration for each person.

Methods	Head movement volume (Z)	Average accuracy	Features
Hutt06	< 20 mm	0.8°	1 stereo cameras, eye tracking
Beymer03	N/A, but > 20 mm	0.6° (only one person)	2 stereo cameras, face & eye tracking
Ours	around 200 mm	1.6°	1 stereo cameras, eye tracking
Ebisa08	around 300 mm	3°	single camera, eye tracking

Table 1. Comparison on eye gaze estimation accuracy with other systems

5

Facial Feature Detection

➢ An eye-guided facial feature detection algorithm is developed:

o Twenty-eight facial features are selected.

o The eye positions provide the most reliable information about the positions for other facial features.

o Each facial feature is located via the fast phase-based Gabor Wavelet matching.

Mean Face Mesh

6

Real-time Facial Feature Tracking

➢ **Tracking stage:**
(1) kalman Filtering for modeling the dynamics of each facial feature.
(2) fast phase-based displacement estimation for locating each facial feature.
(3) updating facial feature model in each frame.
- o **Advantage:** adapting to the appearance changes rapidly.
- o **Issues:**
  - ▪ It will drift away due to the accumulated error under the significant appearance changes.
  - ▪ No effective measurement to identify the tracking failure.

➢ **Correction stage:**
(1) Obtain the most similar model from the training set for each facial feature.
(2) A new position is estimated by using the obtained training model for each facial feature.

➢ **Fusing stage:**
Probabilistically combine the results from tracking and correction information together.
- o **Advantages:**
  - ▪ solving the drifting away issue;
  - ▪ obtaining a similarity measurement.
  - ▪ robust under significant appearance changes.

7

Efficient Recovery of Rigid and Non-rigid Facial Motions

➢ **The motion of the face is the sum of the rigid motion (face pose) and the non-rigid motion (facial expression).**
- ❖ **Issue:** Both motions are nonlinearly coupled in the face image, and it is difficult to separate them from a single face image.

➢ **Proposed technique for motion decomposition:**
The face pose and facial expression parameters is recovered simultaneously as the following minimization process:

$$\min_{M,\alpha_i} \sum_{j=1}^{k} \left\| \begin{pmatrix} u_j \\ v_j \end{pmatrix} - M \begin{pmatrix} x_j^N \\ y_j^N \\ z_j^N \end{pmatrix} - \sum_{i=1}^{k} M\alpha_i \begin{pmatrix} \Delta x_j^i \\ \Delta y_j^i \\ \Delta z_j^i \end{pmatrix} \right\|^2$$

➢ Once the parameters are recovered, face pose (M) and the facial deformation (ΔX) can be derived, based on which we can perform facial expression analysis.

8

Probabilistic Facial Expression Analysis

➢ **Based on the recovered non-rigid facial motions, the six basic facial expressions can be recognized under arbitrary face orientations via Dynamic Bayesian Networks.**

➢ **Other applications**

Monitoring Human Fatigue

Face animation

9

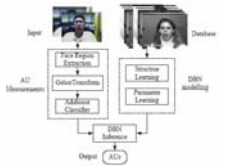
Facial Action Units Recognition

Table 2. A list of action units and their interpretations

➢ **Facial Action Units (AUs)**
represent the muscular activity that produces facial appearance changes (defined in Facial Coding System).

➢ **Objective:** real time recognize action units under real-world circumstance for subtle facial component movement analysis.

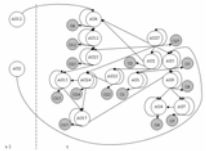
➢ **System modeling:**
Modeling the relationships among AUs as well as their temporal evolutions
- ❖ Face region extraction and feature extraction;
- ❖ Action Units Measurement extraction through AdaBoost;
- ❖ Dynamic Bayesian network inference.



The flowchart for real-time AU recognition system

10

AUs Recognition – Modeling and Experimental Results

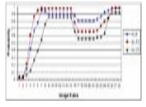


The Dynamic BN for AU modeling.

The DBN model is made up of two interconnected time slices of static BNs.

Advantage:
➢ accounting for uncertainties in AU measurement;
➢ representing probabilistic relationships among different AUs;
➢ modeling the dynamics in AU development;
➢ providing principled inference solutions;
➢ solving the missing measurement using available measurements of other AUs.

(a) An image sequence where the face undergoes an out-of-image plane rotation and the mouth is occluded in some frames



(b) The probabilities of involved AUs in the corresponding frames. (Occlusion occurs from frame 15 to frame 25.)

11

Summary

This project addresses the problems of real-time and non-intrusive human facial behavior understandings for Human Computer Interaction.

The main contributions of this project are summarized as below:

➢ real-time eye detection and tracking under variable illumination and face orientations;

➢ improved eye gaze tracking allowing natural head movement, with minimum personal calibration;

➢ robust facial feature tracking under significant changes in lighting, facial expression and head movement;

➢ rigid and non-rigid facial motions decomposition;

➢ facial expression analysis for six basic expressions;

➢ subtle facial component movement analysis (face action units recognition).

12