**Introduction**

In this work, we introduce a computer vision system that performs automatic facial and body behavior analysis and recognition using RGB-D camera. From RGB video, the system performs facial landmark tracking, which automatically determines the locations of fiducial facial points near major facial components. From depth video, the system extracts skeleton joint positions. Given tracking results, the system performs human eye gaze tracking, head pose estimation, facial expression recognition and gesture recognition.

![Facial Video](image1) ![Depth Video](image2) ![Facial Landmark Detection and Tracking](image3) ![Body Skeleton Tracking](image4) ![Eye Gaze Estimation Method](image5) ![Head Pose Estimation](image6) ![Gesture Recognition](image7)

**Facial Landmark Tracking**

**Approach [1]:**

We proposed the shape augmented regression based facial landmark tracking algorithm that iteratively predicts the face shape updates and estimates the landmark locations. In particular the regression function would automatically change for different face shapes.

![Fig. 2 Shape augmented regression based facial landmark detection method](image8)

**Algorithm 1: Shape augmented regression method**

```plaintext
for t = 1, 2, ..., T or convergence do
    Predict the landmark location update
    \[ \Delta x^t = R^t \Phi(x^{t-1}) + b^t = Q^t \Phi(x^{t-1}) \]
    Update the face shape
    \[ x^t = x^{t-1} + \Delta x^t \]
end
```

Output the estimated landmark locations \( \hat{x} \)

![Fig. 5 Screenshots of gaze estimation demo. The red circles represent the current positions the subject looks at. Top right shows the detected facial landmarks and pupil positions.](image9)

**Facial Landmark based Eye Gaze Estimation**

**Approach:**

1. Relate gaze features (eyeball center) with rigid facial landmarks, and obtain their relative position information from offline training.
2. During online gaze tracking, given the detected 2D facial landmark positions, recover the 3D eyeball center and 3D pupil center.
3. Obtain gaze direction or PoS given eyeball center, pupil center and personal eye parameters.

![Fig. 3 Facial landmark tracking results](image10) ![Fig. 4 Relationship between rigid facial landmarks and eye ball center and extracted relative position information. Images are from [2].](image11) ![Fig. 6 Pose estimation results](image12)

**Facial Landmark based Head Pose Estimation**

**Approach:**

1. Estimate the head pose based on the predicted 2D landmark locations (51 landmarks) and a general 3D deformable face model.
2. The method estimates the projection matrix (depending on the head pose angles) by minimizing the projection error that measures the differences between the projects 2D points (based on 3D deformable model and head pose) and the detected 2D landmarks.

**References**


**Contact**

Intelligent Systems Lab, Dept. of ECSE, Rensselaer Polytechnic Institute
JEC 6219, ECSE Department, Troy, NY, 12180, USA
Prof. Qiang Ji
Tel: +1-518-276-6440
Email: qj@rpi.edu
Website: [http://www.ecse.rpi.edu/~qsj/index.html](http://www.ecse.rpi.edu/~qsj/index.html)