# Real Time 3D Face Pose Discrimination Based On Active IR Illumination

Qiang Ji
Dept. of Electrical, Computer, and Systems Eng.
Rensselaer Polytechnic Institute

Xiaojie Yang
Dept. of Computer Science
University of Nevada at Reno

## 1 Abstract

In this paper, we introduce a new approach for real-time 3D face pose discrimination based on active IR illumination from a monocular view of the camera. Under the IR illumination, the pupils appear bright. We develop algorithms for efficient and robust detection and tracking pupils in real time. Based on the geometric distortions of pupils under different face orientations, an eigen eye feature space is built based on training data that captures the relationship between 3D face orientation and the geometric features of the pupils. The 3D face pose for an input query image is subsequently classified using the eigen eye feature space.

## 2 Introduction

Face pose determination is concerned with computation of the 3D face orientation and position. It represents an important area of research in human computer interaction (HCI) since face pose contains critical information about one's attention, needs, gaze, and level of fatigue. Methods for face pose estimation can be classified into three main categories: *model-based*, *appearance-based*, and feature-based. Model-based approaches typically recover the face pose by establishing the relationship between 3D face model and its 2D projection ([8], [2], [4]). Appearance-based approaches are based on view interpolation and their goal is to construct an association between appearance and face orientation ([3], [6], [7]). Although appearance-based methods are simpler, they are expected to be less accurate than model-based approaches and are mainly used for pose discrimination. Feature-based approach determines face pose using some facial features and their image, and then determines the pose using the conventional point-based pose estimation method. The major challenge with feature-based approach is to robustly detect and track the required facial features from frame to frame under varying illumination conditions, facial expressions, and different head orientations.

In this paper, we present a new model-based approach. Our approach determines 3D face pose from a monocular view of the face. Our study shows that there exists a direct correlation between 3D face pose and the geometric properties of pupils such as pupils size, inter-pupil distance, and pupils shape. These relationships are exploited to classify 3D face poses.

## 3 Image Acquisition

To assist face pose discrimination based on geometric properties of pupils, we built an active infrared (IR) illuminator to illuminate the person's face and use an IR-sensitive CCD camera for image acquisition. According to the original patent from Hutchinson [5], the use of IR illuminator allows to produce the bright-pupil effect, which will be used for pupil detection and tracking.

Our IR illuminator consists of two rings of IR LEDs. The IR light source illuminates the user's eye and generates two kinds of pupil images: bright and dark pupil images. The bright pupil image is produced when the inner ring of IR LEDs is turned on and the dark image is produced when the outer ring is turned on. Figure 2 (a) and (b) presents examples of the acquired images using the image acquisition system described above.

## 4 Pupil Detection and Tracking

The goal of pupil detection and tracking is for subsequent face orientation estimation. A robust, accurate, and real-time pupil detection is therefore crucial. Pupil detection and tracking starts with pupil detection. Figure 1 gives an overview of our pupil tracking system. Pupil tracking can be divided into two stages: pupil detection and pupil tracking as discussed below.

Pupil detection involves locating pupils in the image. A preprocessing is applied to minimize interference from illumination sources other than IR illuminator. This includes Sun light and ambient light inter-

1051-4651/02 $17.00 (c) 2002 IEEE

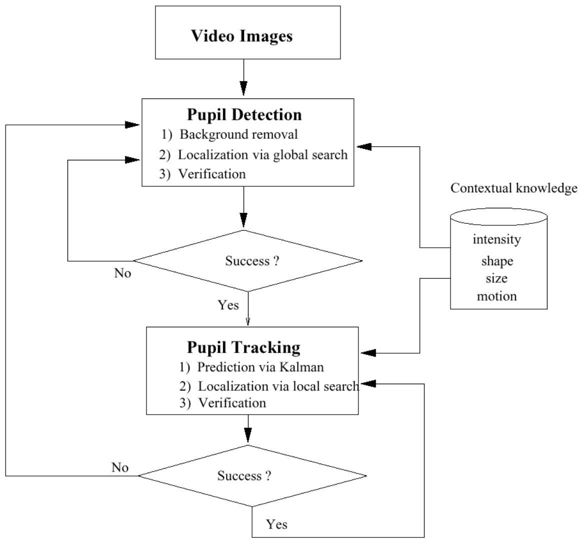



**Figure 1.** Pupil detection and tracking system flowchart

ference. Figure 2 shows an image where parts of the background look very bright, almost as bright as the pupil. This must be eliminated or they may adversely affect pupil detection. Their removal is accomplished by subtracting the image with only ambient light from the one illuminated by both the infrared illuminator and the ambient light. The resultant image, as shown in Figure 2 (c) contains the illumination effect from only the infrared illuminator. A micro-controller with a video decoder has been built to perform real image subtraction. The controller separates each incoming interlaced image frame from the camera into even and odd fields and alternately turns the inner ring on for the even field and off for the odd field. The difference image is then produced by subtracting the odd field from the even field. The images shown in Figure 2 are such produced.

(a) (b) (c)

**Figure 2.** Background Removal via Image Subtraction: (a) the image obtained with both ambient and IR light (even field); (b) the image obtained with only ambient light (odd field); and (c) the image resulted from subtraction (b) from (a)

Given the image resulted from the background removal procedure, pupils may be detected by searching the entire image to locate two bright regions that satisfy certain size, shape, and distance constraints. Given the detected pupils in the initial frames, pupils can then be tracked from frame to frame in real time. Tracking can be done more efficiently by using the location of the pupil in previous frames to predict the location of the face in future frames based on Kalman filtering [1], assuming that the person's pupil will not undergo significant locational change in two consecutive frames. The Kalman filter tracker has been implemented in near real time (25 frames/second) and is found to be rather robust under different face orientations, distances, and illuminations. Sample tracking results are shown in figure 3 , which shows pupils tracking for people with glasses.

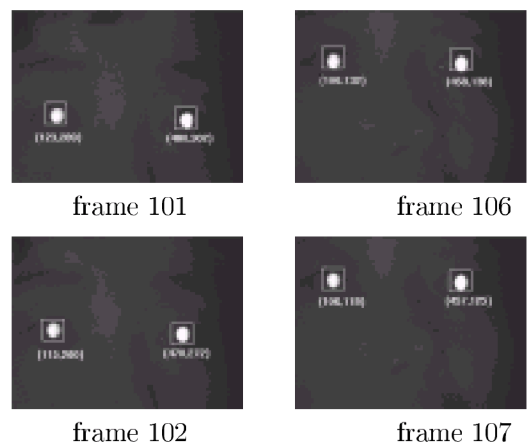



**Figure 3.** Tracking result in consecutive frames with Kalman filtering track scheme.

## 5 3D Face Orientation Discrimination

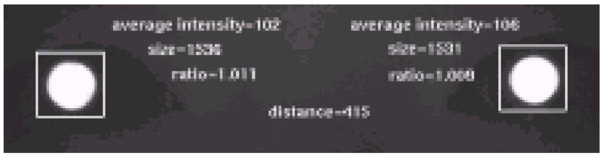
We propose a new model-based approach to classify 3D face pose from a monocular view of the face with full perspective projection. Our study shows that there exists a direct correlation between 3D face pose and the properties of pupils such as pupils size, inter-pupil distance, and pupils shape. Figure 4 shows pupil measurements under different head orientations. It is apparent from these images that

- the inter-pupil distance decreases as the face rotates away from the frontal orientation.

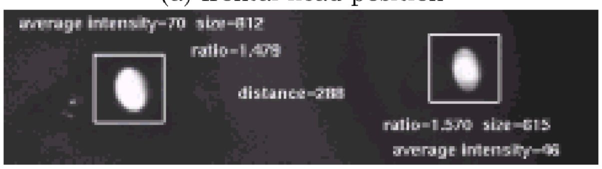
- the ratio between the average intensity of two pupils either increases to over one or decreases to less than one as face rotates away or rotates up/down.

- the shapes of two pupils become more elliptical as the face rotates away or rotates up/down
- the sizes of the pupils also decrease as the face rotates away or rotates up/down.

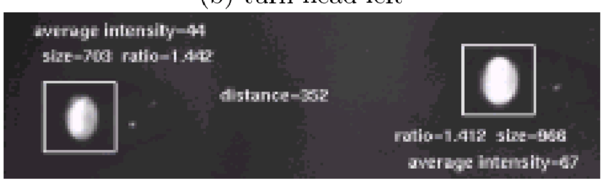
The above observations serve as the basis for estimating face orientation from pupils.




(a) frontal head position




(b) turn head left




(c) turn head right

**Figure 4.** Pupil images for different head orientations. It is clear that pupil properties such as size, intensity, and shape vary with face orientations.

Based on the above observations, we can develop a face pose classification algorithm by exploiting the relationships between face orientation and these pupil parameters. We build a so-called *Pupil Feature Space* (PFS) which is constructed by seven pupil features: inter-pupil distance, sizes of left and right pupils, intensities of left and right pupils, and ellipse ratios of left and right pupils. To make those features scale invariant, we further normalize those parameters by dividing over corresponding values of the front view. Figure 5 shows sample data projections in 3D PFS, from which we can see clearly that there are five distinctive clusters corresponding to five face orientations (5 yaw angles). Note that although we can only plot 3-dimensional space here, PFS is constructed by seven features, in which the clusters will be more distinctive. So a pose can be determined by the projection of pupil properties in PFS.

To maximize the separation of different clusters, we need to find a representation of the PFS by which different pose classes are most apart from each other. A well known method to achieve this goal is principal component algorithm (PCA), or eigen space algorithm, which is to find the principal components of the distribution of poses. The eigenvectors are ordered, each one accounting for a different amount of the variation among the poses, and each individual pose can be represented exactly in terms of a linear combination of the eigenvectors. Training data are collected to build the eigen PFS, and store several models representing typical poses, which are, in our experiments, vary between $-45°$ and $45°$. Figure 6 shows the distribution of the models in eigen PFS, where again a 3-dimensional projection is used while the actual dimensions of the eigen PFS is 7. The face orientation of an input face can then be mapped to one of the clusters based on its Euclidean distance to the center of each cluster.

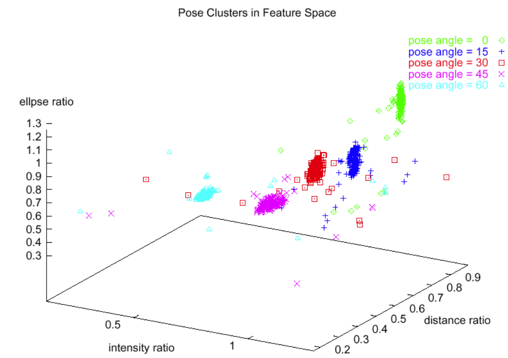



**Figure 5.** Face pose clusters in pupil feature space

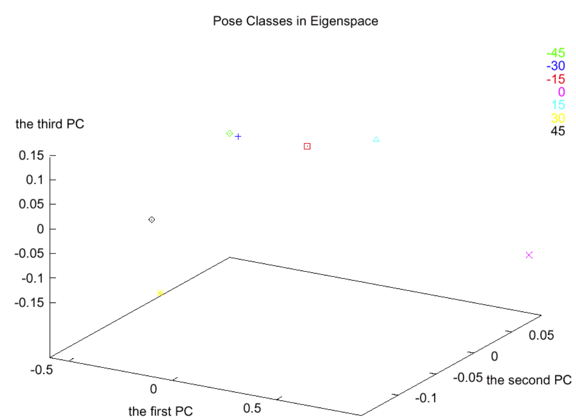



**Figure 6.** Projection of pose classes in eigen PFS

The proposed method may not work well when pupils are occluded by eyelids since this also lead to geometric distortion of the pupils. To overcome this limitation, the face orientation for a particular time instant is computed as the average of the face orientation estimates for a sequence of neighboring frames. And also heuristic constraint such as the estimated face poses for two neighboring time instants should be close, assuming smooth head movement.

### 5.1 Experimental Results

Here we present results of experiments with real data. The experiment involve a human subject rotating her face left and right to produce difference face orientations. Our face pose estimation technique is used to estimate each pose. The face orientation is quantized into 7 angles -45, -30, -15, 0, 15, 30, and 45. A total of 300 data points were collected for each face orientation, representing 7 face orientations. Each data point is comprised of seven raw features (Inter-pupil distance, sizes of left and right pupils, intensities of left and right pupils, and ellipse ratios of left and right pupils.) are collected and pre-processed through a median filtering to remove outliers. The pre-processed data are subsequently normalized by dividing each feature data by the according data of the front view. The leave-N-out technique was used to train and test the algorithm since there are not enough training data available. It works by repeatedly using part of the data for training and the other part for test. This repeats until all data have been training and testing data. In our case, 95% data are used for training and 5% for test in each iteration. The experimental results are summarized in Table 1.

| ground truth | total # of data | # of test data | overall correct estimation | correctness |
|---|---|---|---|---|
| -45 | 300 | | 15 | 100.00 |
| -30 | 300 | | 14.95 | 99.67 |
| -15 | 300 | | 14.9 | 99.33 |
| 0 | 300 | 15 | 14.8 | 98.67 |
| 15 | 300 | | 14.45 | 96.33 |
| 30 | 300 | | 15 | 100.00 |
| 45 | 300 | | 14.2 | 94.67 |

**Table 1.** Overall performance: statistics of pose estimation output.

To further validate the face pose estimation method, we apply it to monitor the face orientation of a driver to detect inattentive driving as measured by the percentage of side view while driving. Figure 7 shows the running average face pose estimation for a period of 6 minutes. As can be seen, most times during this period, face pose is frontal. But there are times when an extended period of time is spent on other directions (left or right), representing inattention. Similar results were obtained for head rotating up and down (head tilt).

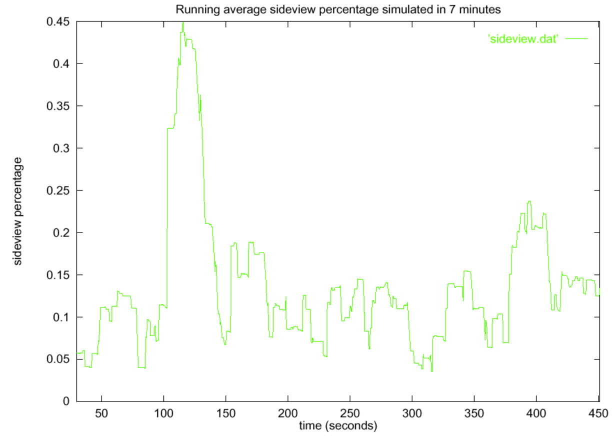



**Figure 7.** Face Orientation Monitoring Over Time. The vertical axis represents the time percentage of the driver's face facing direction other than frontal.

## 6 Conclusions

In this paper, we briefly present a new approach for 3D face pose classification based on geometric properties of pupils under active IR illumination. Using training data, we build an eigen eye feature space that captures relationship between 3D face pose and the geometric properties of the pupils. The eigen eye space is then used for 3D face pose classification. Experiments show the technique can estimate face pose in real time and produce rather robust results for subjects close to the camera. Additional experiments will be conducted to further validate the performance of our algorithm and to study its capability to generalize for people whose data are not involved in training.

## References


[1] A. Blake, R. Curwen, and A. Zisserman. A framework for spatio-temporal control in the tracking of visual contours. *Int. Journal of Computer Vision*, 11(2):127–145, 1993.

[2] A. Gee and R. Cipoll. Determining the gaze of faces in images. *Image and Vision Computing*, 30:639–647, 1994.

[3] S. Gong, E. Ong, and S. McKenna. Learning to associate faces across views in vector space of similarities of prototypes. *Proc of British Machine Vision*, 1998.

[4] T. Horprasert, Y. Yacoob, and L. Davis. Computing 3d head orientation from a monocular image. *Proc. of Int. Conf. on automatic face and gesture recognition*, pages 242–247, 1996.

[5] T. E. Hutchinson. Eye movement detection with improved calibration and speed. *U.S. patent 4950069*, April, 1990.

[6] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. *CVPR-94*, pages 84–91, 1994.

[7] R. Rae and H. Ritter. Recognition of human head orientation based on artificial neural networks. *IEEE Transactions on Neural Networks*, 9(2):257–265, 1998.

[8] A. Tsukamoto, C. Lee, and S. Tsuji. Detection and pose estimation of human face with synthesized image models. *ICPR 94*, pages 754–757, 1994.