

# Head Pose Estimation Based on 2D and 3D Information

Gustavo Peláez  
Intelligent Systems Laboratory  
Av. Universidad, 30  
Spain 28911, Leganés, Madrid  
gpelaez@ing.uc3m.es

Arturo de la Escalera  
Intelligent Systems Laboratory  
Av. Universidad, 30  
Spain 28911, Leganés, Madrid  
escalera@ing.uc3m.es

Jose Armingol  
Intelligent Systems Laboratory  
Av. Universidad, 30  
Spain 28911, Leganés, Madrid  
armingol@ing.uc3m.es

## ABSTRACT

A solution for driver monitoring and event detection based on 3D information from a range imaging camera is presented: The system combines 2D and 3D algorithms to provide head pose estimation and regions of interest identification. Starting with the captured cloud of 3D points from the sensor and the detection of a face in the 2D projection of it, the points that correspond to the head are determined and extracted for further analysis. Later the head pose estimation with 3 degrees of freedom (Euler angles) is estimated using ICP algorithm. As a final step, the important regions of the face are identified and used for further experimentation, e.g. gaze orientation, behavior analysis and more. The resulting application is a complete 3D driver monitoring application, based on low cost sensor; it is described how to combine both 2D and 3D computer vision algorithms for future human factors research enabling the possibility to study specific factors like driver drowsiness, gaze orientation or the head pose estimation itself.

The experimental results shown are compared with ground-truth head movements obtained using an IMU.

## Keywords

Head Pose Estimation, Face Detection, ICP, Point Clouds, Driver Monitoring.

## 1. INTRODUCTION

Driver monitoring is crucial in human factors. It is important to monitor the driver to understand its necessities, behavior or misbehaviors. In the market there is already available technologies that provide automatic driver monitoring systems, but the cost of the sensors used are generally high, and many of them require external and invasive devices that may interfere in the driving process. Modern applications, based on computer vision, take advantage of the development of information technologies and computing capacities to create applications able monitor the driver by non-invasive methods, although limitations inherent to the sensing technology are present e.g. lack of depth information or light dependent conditions. New technologies such as stereo and time of flight cameras, are able to overcome some of these difficulties. Specifically, recent gaming sensors, such as the Kinect from Microsoft, are providing a new set of previously-expensive sensors embedded in a low cost device, thus providing 3D information together with some additional features.

### Acknowledgements:

This work was supported by the Spanish Government through the CICYT projects (GRANT TRA2010-20225-C03-01) and (GRANT TRA 2011-29454-C03-02).

It is described in this paper a solution that analyzes the driver and determines certain aspects of its behavior considering the information obtained from the fusion of the color image and depth (from here, infrared information can be obtained) from the Kinect sensor bar. By combining both streams, a combination of 2D and 3D algorithms is possible and it is aimed to provide a reliable face detection solution, accurate pose estimation and trustable identification of face features, such as the eyes, the mouth and nose. The obtained results from the proposed algorithm were contrasted with a ground truth obtained from an inertial measurement unit (IMU) that was placed behind the head of the test driver. This allowed to confirm whether the combination of the algorithms (2D and 3D) delivered a good estimation. The lower the difference between the IMU angles and the solution the better the last one would be. All of the three rotation angles were tested and compared (pitch, roll and yaw).

It is worth mentioning that the hardware used for this experiment was not originally designed for this type of applications as it was intended for indoor full body tracking and gesture recognition. This presented some difficulties that will later be described but that could be easily solved by simple solutions such as raising the window.

## 2. STATE OF THE ART

The techniques used to monitor the driver can be divided according to the sensing device used. The main three different sensor sets are:

Biometric sensors, based on the measurements of biomedical signals. Although more robust, thanks to the measurement of direct input signal from the driver, they require intrusive methods [1] and [2]. These intrusive methods can lead to important drawbacks, such as a change of the behavior of the driver, or lacks of comfort, that makes them not suitable for real applications.

The second set are on-board electronics (devices, sensors, processing units etc...), mounted inside the vehicle, that analyze the behavior of the driver by extracting information from the CAN bus [3] and [4]. The availability of this information makes them useful for commercial applications but the dependence on the information released by the manufacturer and the strong dependency on the reliability of the signal, can throw back this sensors from a determined application.

Finally, computer vision based techniques are an interesting tool for driver monitoring, that are acquiring high importance in recent times because of the nonintrusive nature for acquiring information. Besides, the high amount of information available on the images can be used to infer the state and the behavior of the driver. On the other hand, the difficulties inherent to computer vision algorithms should be addressed. There are different types of applications that use one or many of the available cameras to address the main problem.

Standard color cameras can be used during day light conditions and can use general computer vision algorithms and can take advantages of information such as color. In [5] Viola and Jones face detection algorithm is used, later condensation algorithms are used for tracking the eye state, with detection based on Gabor Filter. To detect drowsiness in the driver, SVM (Support Vector Machines) are used for open/close detection, and PERCLOSE (percentage of closed eye) to identify the state of the driver. On [6] by using two cameras, rapid 3D face modeling is performed using frontal and profile face information for accurate 2D pose synthesis. In [7] feature extraction from the camera are used together with several parameters (i.e.: percentage of eye closure over pupil over time, quantity of eyed closed over time (Micro sleep) and the Current Car Position (CCP)) to monitor the driver. More focused into head pose estimation, work in [8] provides 3D tracking based on monocular localized gradient orientation (LGO) histograms and support vector regression (SVRs). [9] Provides Eye detection and Head Pose Estimation, the latest is based on matching of specific features (eye distance) with a model and tracking the movement using optical flow.

Other approaches take advantage of the specific features of the infra-red cameras, on which, thanks to the specific illumination condition, the pupil can be easily detected, thus the detection of the eyes is easier ([10] and [11]). Finally stereo systems are very useful, because they provide 3D information but with the disadvantage of the strong processing requirements [12], some of the available commercial systems includes these stereo system [13] and [14].

Not only in the field of intelligent transport system is possible to find face detection and head pose estimation, [15] provides a complete survey of the topic, giving full information of the results and categorizing the different available schemes according to the technique used for head pose estimation: Appearance template methods [16], Detector arrays methods [17], NonLinear regressions methods [18] and [19], Manifold embedding methods [20] and [21], flexible models methods [22], Geometric methods [23], Tracking methods [24] and [25] and Hybrid methods [25], [20] and [26]. Some of the results summarized on [15] are used to compare with the presented work, as it is depicted in tests section.

Not so many of these works take the depth information into account, due to hardware restrictions, price of the devices or strong processing requirements. With the release of the Kinect, it is available to obtain color images, 3D data and much more information based on a low cost sensor. Thus the present algorithm enhances the classical vision detection with depth information, allowing to use accurate point clouds in order to provide precise results in real time.

## 3. SYSTEM DESCRIPTION

The proposed solution is one of the many applications in the IVVI 2.0 which is the second platform of the Intelligent System Laboratory. It is a commercial vehicle, equipped with a large variety of sensors (mostly advanced cameras) to test and develop different technologies to assist the driver (Figure 2).

From this set of cameras, the Kinect is one them. It was attached to the dashboard, as shown in Figures 1 and 2. The location of the Kinect sensor was carefully chosen, taking into account several factors:

- The minimum distance required for the Kinect range detection (0.5 meters).

- Avoid interference with the driver's field of view, therefore the positioning had to be the closest to the dashboard in order to avoid this issue.

- The final position of the Kinect had to take into account the steering wheel and the driver's maneuvers that may block the cameras. To avoid this, a balance between the steering wheel presence and the driver's field of view had to be achieved.

Other locations of the sensor were discarded since they did not fulfill the fore mentioned constraints.

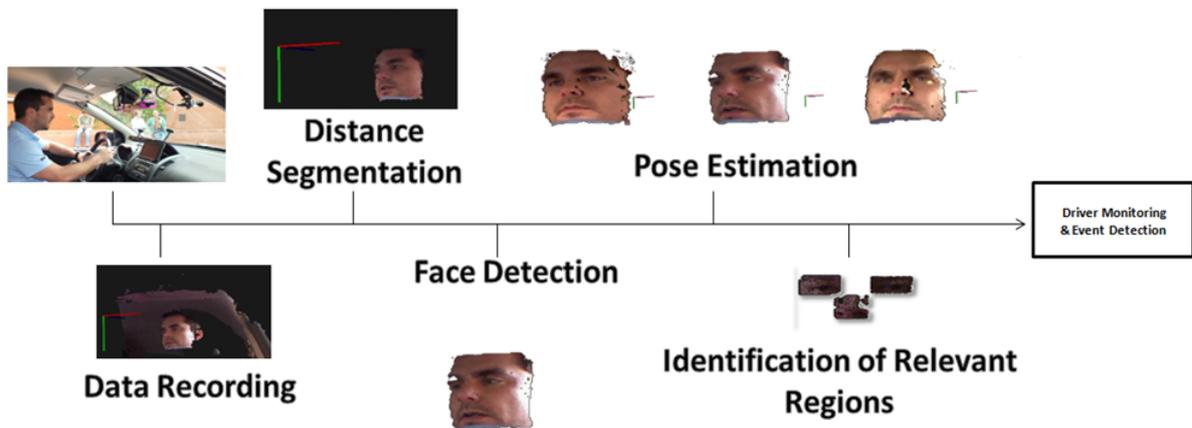


Figure 1: Overview of the system.



Figure 2: Test vehicle IVVI 2.0.

Figure 1 shows the diagram that gives an overall idea of the application. The proposed solution is a new approach to deal with the detection of the driver, fast estimation for the position of the face, determination of its correspondent motion angles and to identify relevant regions of the face for further analysis.

The solution begins by capturing the cloud of points from the Kinect sensor and applying a distance filter to remove the background. Later a color image is built from the cloud, where a face will be searched by applying a fast two dimensional based algorithm. In case a face is detected, a point cloud obtained from the original one is created but only with the points corresponding to the face. The solution computes the head pose constantly by comparing the actual point cloud of the face with the first face detected. The rotation matrix is later obtained and therefore, the Euler angles (3DoF) corresponding to the rotation between the two. Finally, the regions of interest for human factors are searched on the actual face, this is, the eyes and nose.

### Face detection and segmentation

The cloud of points retrieved from the Kinect sensor is filtered, in order to reduce the points to analyze and thus reducing both the search space and the processing time for further algorithms. By taking into account the restrictions of space in the vehicles' cockpit, it is possible to model the environment and reduce the space search to half of the original size by ignoring

those points that are more than 2 meters away from the sensor and the clouds in the lateral extremes as it is more probable that the driver will be seated properly on the seat therefore appearing in the center of the screen.

Once the cloud is filtered and remodeled, a color image is obtained from it by projecting the points to the 2D world and extracting the information of the 3 channels (Red, Green and Blue). As a result, a 2D representation of the cloud (or color image) is obtained and the driver's face is searched using Viola-Jones algorithm [27].

The Viola-Jones algorithm takes less time to accomplish a full search of the face thanks to the space reduction previously achieved and proper configuration for the size of the face to be located in the image.

With the face detected, the next step is to build a 3D object (cloud of points) with the corresponding points of the face. This is done by extracting only the points that correspond to the rectangle that contains the detected face previously found with the Viola-Jones algorithm (Figure 3). Because both, the cloud and the color image, have the same dimensions and are calibrated to match their center and borders (the corners of the image match the corners of the cloud) the position of the pixel of the image has an associated point in the cloud and the cloud of the face is built.



Figure 3: 3D Segmentation of a detected face without cropping

The obtained cloud is later reduced in size by cropping 15% from each side (top and bottom of the rectangle) in order to focus on the space of the face that would allow to estimate the head pose. By cropping the rectangle, information not relevant that may lead to misinterpretations is eliminated i.e. hair in the top, beard, pendants and more. The main reason is that it can affect the cloud due to the inaccuracies of the depth sensor that may capture the points at one moment but ignore them in the next. The second reason is that some of the excluded areas represent parts of the body that do not rotate with the head thus they should be removed for the head pose estimation as not only they do not contribute (may affect negatively) to the rotation estimation.

Although the capturing process is critical, it has an important role at the beginning of the solution, since once a face is detected, the cloud obtained is used as a reference for the comparison with the rest of the next-to-be-found clouds. Cropping is also applied to the reference cloud with the same restrictions and dimensions as previously mentioned.

**Calculate the rotation**

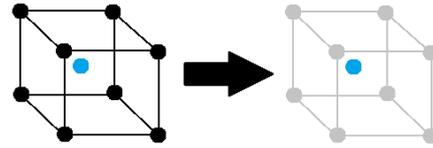
Knowing that the computational time is paramount, the cloud of points is down sampled i.e.: the amount of points is reduced with a voxel filter (Figure 4). This will remove the points from the cloud in a predetermined area and replace them with one point only (its centroid). The filtering is applied to both clouds (the reference and the current). It is important to determine the optimal value of the parameter of the voxel filter. Too small and the amount of points will remain almost the same and the computational time reduction won't be appreciated. Too big and too few points will be available for the next algorithms thus compromising the accuracy of the whole solution.

The point  $p'$  of the obtained cloud from the down sampling is defined by Equation 1 where  $N$  (0.015 for this application) is the parameter that determines the size of the region to be considered when replacing it for a single point.

$$p'(x', y', z') = \left( \frac{\sum_{i=0}^N x_i}{N}, \frac{\sum_{i=0}^N y_i}{N}, \frac{\sum_{i=0}^N z_i}{N} \right) \quad (1)$$

Once the reference and the current cloud are properly filtered, the Iterative Closest Points algorithm (ICP) is applied. This algorithm compare the two clouds, and returns the transformation matrix, composed by the rotation matrix and the translation vector, by minimizing the sum of the square error of the distance between points. The ICP algorithm is applied in order to obtain the transformation matrix  $Mt$  (2) determining the new coordinates of the target cloud.

$$Mt = R \left( \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + T \right) \quad (2)$$



**Figure 4: Example of the voxel filter.**

Where  $T$  is defined as the translation vector for the 3 axis  $\begin{bmatrix} x_t \\ y_t \\ z_t \end{bmatrix}$

From ICP, the matrix  $Mt$  (3x3) is obtained and from where the Euler's rotation angles are obtained according to the equations (3) (4) and (5). The effect of the translation vector is eliminated by using the centroid of the point of clouds corresponding to the face obtained.

$$\Delta\theta = \sin^{-1}(R_{2,1}) \quad (3)$$

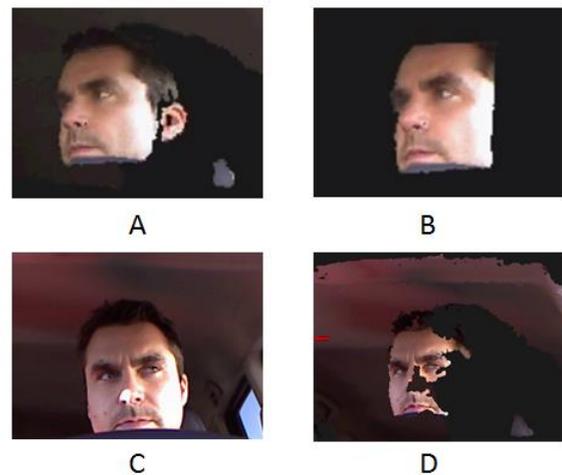
$$\Delta\delta = \tan^{-1} \left( \frac{-R_{3,1}}{R_{1,1}} \right) \quad (4)$$

$$\Delta\varphi = \tan^{-1} \left( \frac{-R_{2,3}}{R_{2,2}} \right) \quad (5)$$

Where  $\Delta\theta$ ,  $\Delta\delta$  and  $\Delta\varphi$  are the rotation angles for Pitch, Roll and Yaw respectively.

It has to be mentioned that ICP needs to be properly configured in order to deliver reliable results. After a variety of tests, it was noticed that the slight variations of the Viola-Jones algorithm when delivering the rectangle containing the face (therefore the cloud of points of the face) altered significantly the results obtained between measurements. In order to avoid these errors, a constant size is determined for the two clouds to be analyzed by ICP.

Finally, at the last part of the solution, the eyes and the nose are searched inside the area defined as face, based on physical constraints and Haar-Like features.



**Figure 5: Examples of normal face detection (A and B) and error due to extreme light conditions (C and D).**

Facial drawing techniques were used in order to accelerate the process of searching the desired facial features. One of the drawing techniques states that the eyes should be in the center of the whole face and the nose slightly lower. These restrictions were implemented with flexible parameters in order to avoid false positives.

These facial features, combined with the use of different techniques, such as PERCLOSE analysis or gaze orientation according to the pupil's, are suitable for further human factors and driver behavior analysis.

#### 4. RESULTS

In order to determine the reliability of the solution, a series of different test were done with a total of 20 individuals performing head movements in the three rotation axels: lateral movements (left to right) vertical movements (up and down) and roll movements. An IMU that registered the rotation angles was attached to the back of the head of the individuals for ground truth measurements. Results showed in Table I depict the Mean Absolute Error (MAE) and figures 6 to 8 represent the results corresponding to one of the test performed.

TABLE I

MEAN ANGLE ERROR IN THE OVERALL TEST BETWEEN THE IMU AND THE

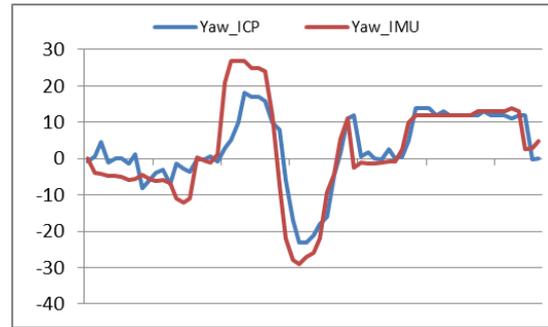
Angle	Error
<i>Pitch</i>	2.5
<i>Yaw</i>	3.8°
<i>Roll</i>	2.7°

**Results of the errors per angle for the overall test.**

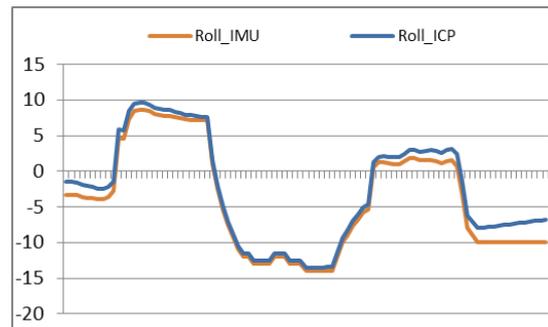
Measuring Mean Absolute Error (MAE) as:

$$\frac{1}{n} \sum_{i=1}^n |\text{Angle}_{\text{IMU}} - \text{Angle}_{\text{ICP}}|$$

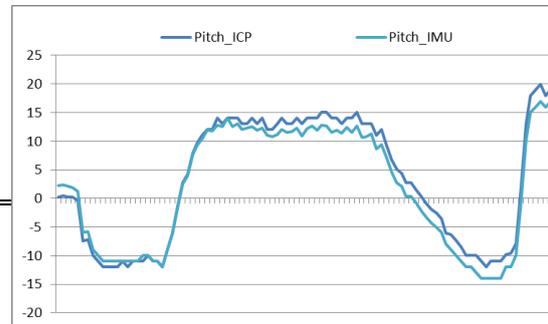
The system was tested with different lighting conditions and positions of the face, in a controlled scenario and later during a ride in the streets. The speed of the solution, running on a PC Intel core i7 is up to 10 frames per second. Thanks to the continuity of the sequences, more than 30 tests were done for each angle in every subject, focusing (but not only) on Yaw and Pitch as these are the two main rotation angles that the driver performs while driving, thus the information this two angles can reveal is more important on this context. As an application example, Yaw could be used to determine if a driver was distracted or not at a determined moment. Pitch could be part of a drowsiness detection, combined with other techniques such as the PERCLOSE analysis of the driver (Figure 9). As seen on (Figs 6-8.) The system



**Figure 6: Comparison of the angles obtained for Yaw from ICP and the IMU. The Y axis indicates the angle or rotation in degrees and the X axis indicates the time stamp in seconds.**



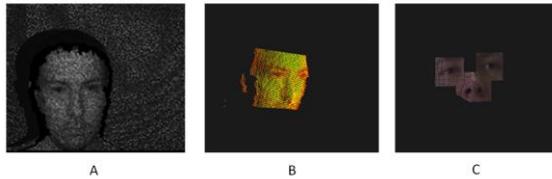
**Figure 7: Comparison of the angles obtained for Roll from ICP and the IMU. The Y axis indicates the angle or rotation in degrees and the X axis indicates the time stamp in seconds.**



**Figure 8: Comparison of the angles obtained for Pitch from ICP and the IMU. The Y axis indicates the angle or rotation in degrees and the X axis indicates the time stamp in seconds.**

was able to track the movement of the head for its 3 rotation angles that with a very limited error.

As shown in Table I and (Figures 6-8.), the face pose estimation algorithm determined with success the position of the head; with error rates always below 4°. Results comparison provided by [15] shows a compendium of 38 different works with results regarding the fine pose estimation. The work presented here provides considerably better results (in terms of MAE) than 33 of them (approx. 87%), and very similar to the other 5. Furthermore, the results



**Figure 9: Application results with the original IR information from the sensor (a), 3D face detection (b) and relevant regions identification (eyes and nose) (c).**

shown on this paper are similar to other more recent approaches [8] and [9]. Although a comparison of overall results is statistically insignificant, due to the diversity of the databases used, it has been proved that the performance of the system, based on a low cost sensor, is close to other state of the art systems, even with the lack of a tracking stage that would allow smoother error rates. To allow a real and statistically significant results comparison, Intelligent System Lab [28] provides on its web page the database, with both 2D and 3D information from the Kinect sensor, as well as the IMU data for ground truth.

During the development of the aforementioned tests, it has been proved that, although the Kinect device was designed for analyzing objects the size of a human body, it delivered good results for head pose estimation. Another advantage obtained from this system is the possibility of the reduction of false positives when detecting faces inside the extracted RGB image. Thanks to a simple mask applied to the 2D image, based on the distance discrimination from the depth information, the resulting image would deliver a much reduced search space for the face and therefore eliminating false candidates, that have geometrical and chromatic similarities with a face.

## 5. CONCLUSIONS AND FUTURE WORK

With the described solution, the orientation of the head was determined with the ICP algorithm and Viola-Jones' face detection. This was possible due to the collaboration between the analysis in 2D from the image and the corresponding 3D point in the cloud both coming from the same sensor. Besides, the solution is able to fast locate relevant regions of the face, such as eyes and nose thanks to geometrical techniques borrowed from drawing portraits

It is important to notice that the technology used for this solution showed a remarkable performance considering its low cost, although it presents some limitations that should be mentioned. Since Kinect was designed for indoor applications, its usability for outdoors applications is limited, specifically it is very sensitive to strong illumination (Figure 5 C and D). However, this is a common issue on most of the video based systems. On the other hand, although this is one

of the limitations to take into account, the system performance was highly reliable under other circumstances i.e. not extremely strong illumination during daylight conditions, nightlight conditions, or indoor. This makes this algorithm suitable for a wide variety of applications, some of them are: An application where the system performance would be useful is night-fatigue monitoring as most applications use normal color cameras that won't work in the night whereas the Kinect can use the infrared intensity stream to get a clear picture of what is in front. Future work lines will focus on the analysis of the eyes, more specifically, the amount of time they are closed in order to determine if the driver is falling asleep. This could be combined with the analysis of the orientation of the head to give a more advance driver monitoring system.

## 6. References

- [1] C. Papadelis, Z. Chen, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, E. Bekiaris, and N. Maglaveras, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents.," *Clin. Neurophysiol.*, vol. 118, no. 9, pp. 1906–1922, 2007.
- [2] C. Papadelis, C. Lithari, C. Kourtidou-Papadeli, P. D. Bamidis, E. Portouli, and E. Bekiaris, "Monitoring driver's sleepiness on-board for preventing road accidents.," *Stud. Health Technol. Inform.*, vol. 150, pp. 485–489, 2009.
- [3] T. Wakita, K. Ozawa, C. Miyajima, K. Igarashi, K. Itou, K. Takeda, and F. Itakura, "Driver identification using driving behavior signals," *Proceedings. 2005 IEEE Intell. Transp. Syst. 2005.*, 2005.
- [4] Y. Takei and Y. Furukawa, *Estimate of driver's fatigue through steering motion*, vol. 2, no. 1. Ieee, 2005, pp. 1765–1770.
- [5] M. J. Flores, J. M. Armingol, and A. Escalera, "Real-Time Warning System for Driver Drowsiness Detection Using Visual Information," *J. Intell. Robot. Syst.*, vol. 59, no. 2, pp. 103–125, 2009.
- [6] J. Heo and M. Savvides, *Rapid 3D face modeling using a frontal face and a profile face for accurate 2D pose synthesis*. IEEE, 2011, pp. 632–638.
- [7] X. Li, E. Seignez, and P. Loonis, "Vision-based estimation of driver drowsiness with ORD model using evidence theory," in *Intelligent Vehicles Symposium (IV), 2013 IEEE*, 2013, pp. 666–671.
- [8] E. Murphy-Chutorian and M. M. Trivedi, "Head Pose Estimation and Augmented Reality Tracking: An Integrated System and Evaluation for Monitoring Driver Awareness," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 2, 2010.

- [9] R. Oyini Mbouna, S. G. Kong, and M.-G. Chun, "Visual Analysis of Eye State and Head Pose for Driver Alertness Monitoring," *Intell. Transp. Syst. IEEE Trans.*, vol. 14, no. 3, pp. 1462–1469, 2013.
- [10] I. Garcia, S. Bronte, L. M. Bergasa, J. Almazan, and J. Yebes, "Vision-based drowsiness detector for real driving conditions," in *Intelligent Vehicles Symposium (IV), 2012 IEEE*, 2012, pp. 618–623.
- [11] M. J. Flores, J. M. Armingol, and A. de la Escalera, "Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions," *EURASIP J. Adv. Signal Process.*, vol. 2010, p. 3, 2010.
- [12] H. Eren, U. Celik, and M. Poyraz, "Stereo Vision and Statistical Based Behaviour Prediction of Driver," *2007 IEEE Intell. Veh. Symp.*, 2007.
- [13] "Seeingmachines." [Online]. Available: <http://www.seeingmachines.com/>. [Accessed: 01-Oct-2013].
- [14] "Smarteye." [Online]. Available: <http://www.smarteye.se/>. [Accessed: 01-Oct-2013].
- [15] E. Murphy-Chutorian and M. M. Trivedi, "Head pose estimation in computer vision: a survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 4, pp. 607–626, 2009.
- [16] J. Sherrah, S. Gong, and E. J. Ong, "Face distributions in similarity space under varying head pose," *Image and Vision Computing*, vol. 19, no. 12, pp. 807–819, 2001.
- [17] H. A. Rowley, S. Baluja, and T. Kanade, "Rotation invariant neural network-based face detection," *Proceedings. 1998 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (Cat. No.98CB36231)*, 1998.
- [18] R. Stiefelhagen, J. Yang, and A. Waibel, "Modeling focus of attention for meeting indexing based on multiple cues," *IEEE Trans. Neural Netw.*, vol. 13, no. 4, pp. 928–938, 2002.
- [19] M. Voit, K. Nickel, and R. Stiefelhagen, "Head Pose Estimation in Single and Multi-view Environments - Results on the CLEAR'07 Benchmarks," in *Proc. Int'l Workshop Classification of Events, Activities and Relationships*, 2007.
- [20] J. Wu and M. Trivedi, "A two-stage head pose estimation framework and evaluation," in *Pattern Recognition*, 2008, vol. 41, no. 3, pp. 1138–1158.
- [21] S. Yan, Z. Zhang, Y. Fu, Y. Hu, J. Tu, and T. Huang, "Learning a Person-Independent Representation for Precise 3D Pose Estimation," in *CLEAR 2007*, 2009, pp. 297–306.
- [22] J. X. J. Xiao, S. Baker, I. Matthews, and T. Kanade, "Real-time combined 2D+3D active appearance models," *Proc. 2004 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, 2004. CVPR 2004.*, vol. 2, 2004.
- [23] J.-G. Wang and E. Sung, "EM enhancement of 3D head pose estimated by point at infinity," *Image Vis. Comput.*, vol. 25, no. 12, pp. 1864–1874, Dec. 2007.
- [24] G. Zhao, L. Chen, J. Song, and G. Chen, "Large Head Movement Tracking Using SIFT-Based Registration," in *Proc. ACM Int'l Conf. Multimedia*, 2007, pp. 807–810.
- [25] E. Murphy-Chutorian and M. M. Trivedi, "HyHOPE: Hybrid Head Orientation and Position Estimation for vision-based driver head tracking," *2008 IEEE Intell. Veh. Symp.*, 2008.
- [26] S. O. Ba and J. M. Odobez, "A probabilistic framework for joint head tracking and pose estimation," *Proc. 17th Int. Conf. Pattern Recognition, 2004. ICPR 2004.*, vol. 4, 2004.
- [27] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, 2004.
- [28] "Intelligent System Lab." [Online]. Available: [http://www.uc3m.es/portal/page/portal/dpto\\_ing\\_sist\\_emas\\_automatica/home/research\\_activities/isl](http://www.uc3m.es/portal/page/portal/dpto_ing_sist_emas_automatica/home/research_activities/isl). [Accessed: 01-Apr-2012].