

Real-Time Robust Face Tracking for Driver Monitoring

Jesús Nuevo, Luis M. Bergasa, Miguel A. Sotelo, M. Ocaña

Department of Electronics
University of Alcalá
CAMPUS. 28805 Alcalá de Henares (Madrid), Spain.
E-Mail: jnuevo,bergasa,sotelo,mocana@depeca.uah.es

Abstract—In this paper we present an active appearance model and a fitting algorithm to track a driver's face, as a component for a driver alertness monitoring system. The proposed method represents some improvements respecting a previous prototype developed by the authors, based on computer vision using active illumination. We test the tracker under different conditions where our previous tracking system fails or exhibits poor performance, such as changing light conditions, occlusions, daylight or drivers wearing glasses. The algorithm is very efficient and able to run in real-time. Some experimental results and conclusions are presented.

I. INTRODUCTION

The increasing number of traffic accidents due to a diminished driver's vigilance level has become a serious problem for society. According to the U.S. National Highway Traffic Safety Administration (NHTSA), falling asleep while driving is responsible for at least 100.000 automobile crashes annually. An annual average of roughly 40,000 nonfatal injuries and 1,550 fatalities results from these crashes [1]. These statistics do not deal either with crashes caused by driver distraction, which is believed to be a larger problem. As car manufacturers incorporate intelligent vehicle systems in order to satisfy consumer ever increasing demand for a wired, connected world, the level of cognitive stress on drivers is being increased. That is, the more assistant systems for comfort, navigation or communication the more sources of distraction from the most basic task at hand, i.e. driving the vehicle.

In the last few years many researchers have been working on the development of systems for monitoring driver's level of vigilance using different techniques: physiological measures [2], steering wheel movements and lateral position [3] or computer vision [4]. One of the most recent research in this line is the ambitious European project AWAKE (System for Effective Assessment of Driver Vigilance and Warning According to Traffic Risk Estimation) [5]. The Consortium includes two major car manufacturers (Fiat, DaimlerChrysler), four automotive system developers (SIEMENS, ACTIA, NAVTECH and AUTOLIV) and many research institutes and universities. A multi-sensor approach is proposed in this project adapted to the driver, the vehicle, and the environment in an integrated way. This system merges, via an artificial intelligent algorithm, data from on-board driver moni-

toring sensors (such as an eyelid camera and a steering grip sensor) as well as driver behaviour data (i.e. from lane tracking sensor, gas/brake and steering wheel positioning). The conclusions of this project is that AWAKE has made a major breakthrough but with the current sensor technologies can not be used outside the well-structured highway nor can be applicable for all drivers. Accordingly, diagnosis performance needs to be improved before including a system like this in commercial vehicles.

In [6], the authors developed a real-time system for driver alertness monitoring based on computer vision using active infrared illumination. This system worked very well at night but its performance during daytime and with drivers wearing glasses decreased. In order to improve the weakness of this system we propose a real-time robust face tracker based on Active Appearance Models (AAMs)[7], [8].

Active Appearance Models are generative models closely related to Active Blobs [9] and Morphable Models [10]. Those models and their related algorithms fall in the area of "analysis through synthesis" methods. They try to parameterize the contents of an image by generating a synthetic image as close as possible to the original (or a region of it). The synthetic image is obtained from a model consisting on both appearance (or texture) and shape. This shape and appearance are learned in a training process, and thus represent a constrained range of possible appearances and deformations. The fitting algorithm iteratively minimizes the error between the instantiation of the model and the real image. AAMs are linear in both shape and appearance, but are nonlinear in terms of pixel intensities. The model fitting is thus a non-linear minimization problem.

Active Appearance Models have received considerable attention over the past few years and have a wide range of applications. Those include medical imaging, although probably the most common is face modelling [11]. This technique has been used in laboratories but not for a driver monitoring system using images obtained from a real car moving in a motorway. There is only a work [12] where the same AAM-based approach was taken, but that work did not provide any qualitative results about tracking robustness, nor driver alertness. Our aim is to use an AAM tracker as the foundations for a robust driver alertness monitoring functional at any moment of

the day and with any user in real driving conditions, as was done in our previous prototype. In this paper some preliminary results of the system are presented.

In section II, we will present a short description of our previous system. Our tracker algorithm based on Active Appearance Models is introduced in section III. Test results can be found in section IV. Conclusions and future works are presented in section V.

II. BACKGROUND: DRIVER ALERTNESS MONITORING

Our real-time system for driver alertness monitoring was based on previous studies [13], [14], [2] that identified drowsy and inattentive driver behaviours. We explored the opportunities of parameters such as PERCLOS (percent of eye closure, [15]), blinking frequency and duration, presence of nodding, and the *fixed-gaze* parameter. We showed that a multi-parameter fusion using a fuzzy system yields accurate results. See [6] for further details.

The image capture system was based on a single camera with near-infrared (near-IR) sensitivity placed on the car dashboard in the center of two concentric rings of near-IR LEDs. The IR light from the inner ring reflected in the driver’s retina, and the *bright pupil* effect appeared clearly in the images, easing the segmentation and detection of the driver’s eyes. Light from the outer ring was used for ambient illumination. An example of two images obtained with this method can be seen in figure 1. This illumination system worked well at nights, and showed to be robust in most situations. However, it could not be used with drivers wearing glasses that reflect IR light, or during daytime, as sunlight hid the LED emissions almost completely. Face tracking was based on two Kalman filters, one for each pupil. When, due to blinks or occlusions, the pupils disappeared from the image, an adaptative search algorithm would locate the pupils when they appeared again. Although this algorithm is effective in most situations, the uncertainty of the process increases with time while the pupils are hidden.

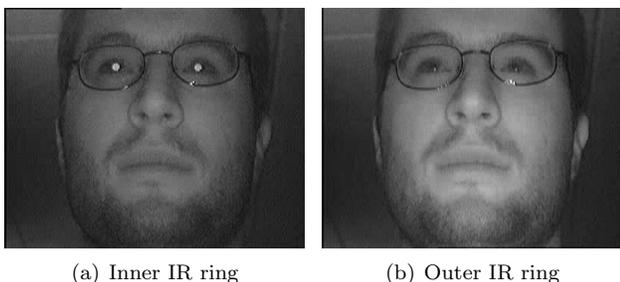


Fig. 1. Bright pupil effect

We seek to improve our system to make it functional at any moment of the day. During nights, both tracking systems, AAM and IR-based, are expected to collaborate, and AAM to work on its own during daytime. Additionally, as the AAM tracker is based on the whole

face, the tracker will be more robust than the IR-based, and continue working at any moment, making new data available to the monitoring system.

III. ACTIVE APPEARANCE MODELS

In this section, we will present our tracking algorithm based on AAMs. A brief description of the AAM features, the fitting algorithm and some extensions to it. We refer to the extensive AAM literature for further details.

The AAMs that model shape and appearance separately are known as *independent* AAMs. The shape of the AAM is defined as the coordinates of the v vertices of the shape

$$\mathbf{s} = (x_1, y_1, x_2, y_2, \dots, x_v, y_v)^t \quad (1)$$

This shape is instantiated from a set of *shape vectors* and a *base shape*

$$\mathbf{s} = \mathbf{s}_0 + \sum_{i=1}^n p_i \cdot \mathbf{s}_i \quad (2)$$

Without loss of generality, the vectors are assumed to be orthonormal. The shape vertices are usually triangulated in a mesh. The gray-level (or colors) of the pixels that fall in the triangles of the base mesh \mathbf{s}_0 represent the appearance of the AAM. Using a convenient abuse of notation, let $\mathbf{x} = (x, y)^t$ also denote the pixels inside the triangles. An appearance $A(\mathbf{x})$ is instantiated then as

$$A(\mathbf{x}) = A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i \cdot A_i(\mathbf{x}) \quad (3)$$

where $A(\mathbf{x})$ is the *base appearance* and the $A_i(\mathbf{x})$ are the *appearance vectors*. As with the shapes, the $A_i(\mathbf{x})$ are orthonormal.

A. Model construction

One of the main drawbacks of AAMs has been the computation of the mean shape and appearance, and the variation bases. The first step of this process was to manually annotate the vertices of the shape in every image of the training set, which could be very large (over hundreds of samples), thus requiring important amounts of time. Lately, automatic model construction has received attention from researchers [16], [17].

While we plan to use these techniques for a future improvement of the system, the preliminary results presented in this paper are obtained with models constructed *by hand*. The outline of the process is as follows. Once the vertices have been marked, the resulting shapes are aligned using the Procrustes algorithm [7]. The shape \mathbf{s}_0 is set to be the mean of those shapes. Principal Components Analysis (PCA) is then performed on the aligned shapes, and most commonly, the variation vectors that explain a 95% of the samples are chosen.

The appearance in the hand marked shapes is warped to a properly scaled \mathbf{s}_0 . The mean of the appearance values is chosen to be \mathbf{A}_0 . PCA is used again to obtain the

vectors that explain a given percentage of the samples, usually 95%.

B. AAM Fitting

The purpose of the fitting process of an AAM is to obtain the parameters that minimize the error between the input image I and the model instance $A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x})$:

$$\sum_{\mathbf{x} \in \mathbf{s}_0} \left[A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right]^2 \quad (4)$$

The input image is warped back using a piecewise affine warp, $W(\mathbf{x}; \mathbf{p})$, that relates the triangles of \mathbf{s} and those of \mathbf{s}_0 .

We use the approach of Baker and Matthews to minimization, and their *inverse compositional algorithm* (IC) [8], [18] for an efficient computation, that allows the algorithm to run in real-time. We also implement some of the extensions they proposed [19], with some minor refinements. Here we will just outline the *project-out inverse compositional algorithm*, for the sake of completeness.

Solving for the parameters of the AAM requires obtaining the values of the $(m+n)$ coefficients simultaneously. The *project-out* algorithm reduces the dimension of the problem by solving first for the n shape parameters and then the m appearance parameters. If we divide the space of appearances in the subspace defined by the appearance vectors ($sub(A_i)$) and its complementary, equation 4 can be rewritten as:

$$\begin{aligned} & \left\| A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right\|^2 = \\ & \left\| A_0(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right\|_{sub(A_i)^\perp}^2 + \\ & \left\| A_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i A_i(\mathbf{x}) - I(\mathbf{W}(\mathbf{x}; \mathbf{p})) \right\|_{sub(A_i)}^2 \end{aligned} \quad (5)$$

We can then solve for the first half of the right hand side, and then obtain the appearance parameters solving the other half

$$\lambda_i = \sum_{\mathbf{x} \in \mathbf{s}_0} A_i(\mathbf{x}) \cdot \left[I(\mathbf{W}(\mathbf{x}; \mathbf{p})) - A_0(\mathbf{x}) \right] \quad (6)$$

The first half of the equation is solved with the *inverse compositional algorithm*. The update to the shape parameters is obtained from the error image, $Err(\mathbf{x}) = [I(\mathbf{W}(\mathbf{x}; \mathbf{p})) - A_0(\mathbf{x})]$, using the Gauss-Newton approximation to the Hessian:

$$\Delta \mathbf{p} = \mathbf{H}^{-1} \sum_{\mathbf{x} \in \mathbf{s}_0} \left[\nabla A_0(\mathbf{x}) \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \right]_{sub(A(\mathbf{x}))^\perp} Err(\mathbf{x}) \quad (7)$$

with:

$$\mathbf{H} = \sum_{\mathbf{x} \in \mathbf{s}_0} \left[\nabla A_0(\mathbf{x}) \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \right]_{sub(A_i)^\perp}^T \left[\nabla A_0(\mathbf{x}) \frac{\partial \mathbf{W}}{\partial \mathbf{p}} \right]_{sub(A_i)^\perp} \quad (8)$$

The steepest descent images are constant, as they are obtained from the base appearance, and the warp is evaluated at $\mathbf{p} = 0$. Thus the Hessian is constant too, making the overall algorithm very fast.

IV. TEST RESULTS

We want to validate the AAM-based tracking system in all driving situations, but specially in those where the existing IR-based system lose track of the eye position due to occlusions, or did not work properly. The main ones were blinks, driver wearing glasses, daylight and fast changes in illumination.

A. Test setup

The situations mentioned above were tested for different users and in at least two sequences for each driver. The sequences duration and events on them are shown in table I. The testing sequences are selected moments from the longer video sequences used in [6].

Models used are person specific. While this is a simpler case, the extension to generic models is well studied [20]. For each user, up to 8 frames were hand marked to train the model. Those frames were not present in the testing sequences. The mesh contains 62 vertices and 101 triangles (see figure 2). The template size is 110x110 pixels. Most models contained four shape vectors and six appearance ones.

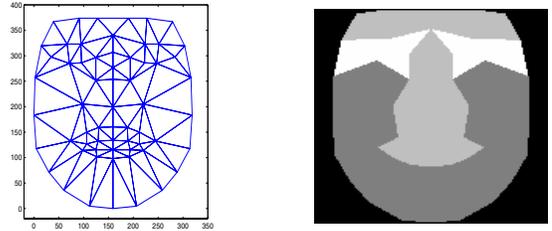


Fig. 2. Model mesh and weighting mask

From the various fitting algorithms based on the *inverse compositional* approach, we chose the *reweighted normalising IC* with the *spatial coherence of outliers* approximation [21], [22], [19]. This algorithm is robust to occlusions, and requires the estimation of the *scale* parameter of the robust function it uses [23], [24], [25]. We tested several of the most common robust functions, including Huber, Cauchy, Geman-McLure and Tukey functions, among others [26].

The value of the scale parameter σ is the same for all pixels in the template. Following [23] and [26], we estimate σ as shown in equations 9 to 12.

TABLE I
TEST SEQUENCES

Nr.	Event	Frame count	Driver	# of training frames	# of Shape / App. vectors
1	Closed eyes	80	Driver 1	5	4 / 6
2	Short blinks	100	Driver 2	7	4 / 5
3	Head turns, up to 30% occ.	70	Driver 2	7	4 / 5
4	Illumination changes	100	Driver 1	6	4 / 6
5	Daylight, short blinks	80	Driver 1	5	3 / 4
6	Driver wearing glasses	100	Driver 3	8	4 / 6
7	Severe and total occlusion	330	Driver 1	5	4 / 6
8	Strong lights, saturated image	250	Driver 4	5	4 / 5

$$\sigma_{Huber} = 1.345 \text{ median}(|Err(x)|) \quad (9)$$

$$\sigma_{GM} = 1.483 \text{ median}(|Err(x)|) \quad (10)$$

$$\sigma_{Tukey} = 4.685 \text{ median}(|Err(x)|) \quad (11)$$

$$\sigma_{Cauchy} = 2.385 \text{ median}(|Err(x)|) \quad (12)$$

The value is estimated in each iteration of the algorithm. While it does not provide information about the behaviour of each pixel, it does not require any previous knowledge of the error distribution. The best performing, in terms of convergence and error, were Huber and Geman-McLure functions.

$$\psi_{Huber}(Err(x); \sigma) = \frac{\sigma^2}{\sigma^2 + Err(x)^2} \quad (13)$$

$$\psi_{GM}(Err(x); \sigma) = \frac{\sigma^2}{(\sigma^2 + Err(x)^2)^2} \quad (14)$$

In addition to the existing algorithm, we include a mask to reweight the points inside the mesh, as shown in figure 2. We give extra importance to points corresponding to the eyes, nose and mouth. These are the most interesting zones of the face as they contain information valuable to our system, and they are also less sensitive to changes in illumination. On the other hand, these changes are usually localized to one side of the face. To reduce its influence, we include local offset images in the building process of the appearance vector base. Each image has an area filled with 1s, and the rest is set to zero. We include six of these images, that added result in a global offset covering the whole face.

Limiting the range of values the shape parameters (λ_i) could take is a simple way to avoid excessive deformation of the model. The limits are calculated as 3 times the standard deviation of the projections of the shapes in the training set over the vectors that form the shape base.

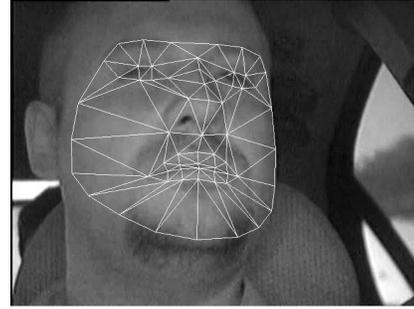
B. Test cases and results

The first objective was to test if the tracking was robust in a regular driving situation, i.e. frontal face position, short blinks and no severe illumination changes. The fitting is correct, fast and stable with blinks and minor light changes and shadows.

During longer blinks, the driver's face usually remains in frontal position. In this case, the main changes may

come from illumination variation, or a change in scale as the head may come closer to the camera. Figure 4 shows frames during a long blink. Again the fitting is correct.

The AAM fitting algorithm was tested in presence of daylight and worked properly. In this same sequence, short blinks were present. Although the IR illumination was also active, its influence on the appearance is negligible. Fitting results can be seen in figure 3.



(a) Frame 43

Fig. 3. Fitting results with daylight

For the IR-based system, fast changes in illumination represented a problem, as one of the parameters of the tracker was the pupil pixel intensities. If the changes in illumination affected the eye region, tracking could be lost. If the mean gray level of the image increased or decreased, the camera could adjust its gain control, rapidly modifying pupil pixel intensities. We tested the AAM-based system in a number of sequences under fast illumination changes, both local and global. Figure 4 shows some frames of one of those sequences. It should be noted that in this case, the influence of the illumination in the tracker is not restricted to the eye region.

Head turns that provoke occlusion over a certain degree hide one of the pupils to the IR-based system. As can be seen in figure 5, the AAM-based tracker can handle remarkable head turns, that in the sequences are up to 45°. We also tested the system with sequences with the face severely occluded by the steering wheel (up to 100% occlusion). In our tests, for a model correctly fit in advance, tracking is maintained for occlusions of up to 35%. In figure 8, triangles with a very low weight are

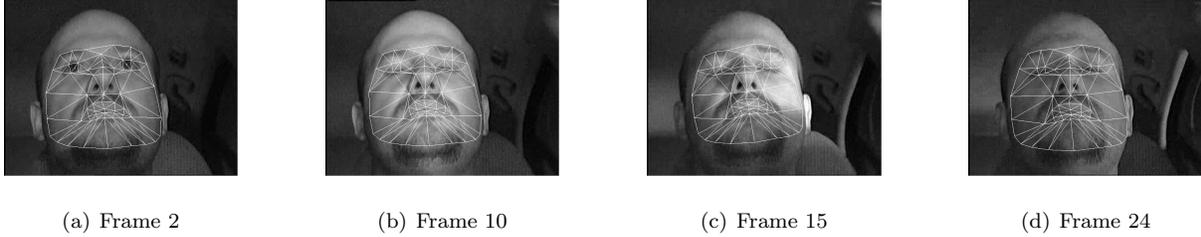


Fig. 4. Fitting results during blink and fast illumination change

Painted in black. The algorithm, in this case, does not assign the lowest weights to all the triangles over the wheel. This is so because the gray level of those pixels is not very far from what can be instantiated from the model, thus yielding moderate error values. This could be solved by using color images instead of gray images.

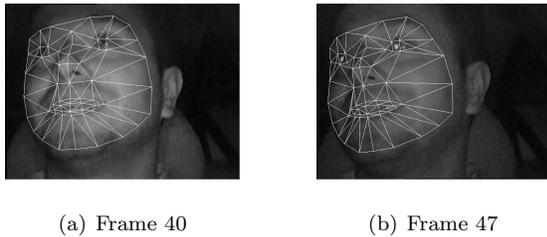


Fig. 5. Fitting results during a right head turn

As happened for blinks, the tracker can handle drivers wearing glasses. However, this model required more training images, and the fitting in the eye region was less precise, due to the reflections on the glasses from the IR LEDs and also from car and street lights (see figure 6).

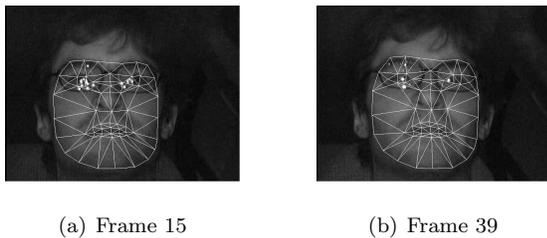


Fig. 6. Fitting results with glasses

C. Fitting accuracy

Fitting accuracy is a critical point of the tracking system, as location of the eyes have to be precise enough to not interfere with the rest of the system, adding an extra error component. Fitting error is somehow difficult to evaluate quantitatively, as the model can not generate the whole space of shape and appearances where the testing images fall. One possibility is to obtain the mean square error (MSE) between the target image and the model instantiation (although, as a non-linear problem,

there are local minima that could result in low MSE while being an incorrect solution). In figure 7 the MSE values for the sequence #4 (fast illumination changes) for different robust error functions are shown. These error image values are multiplied by the weighting mask of figure 2. We run the tests with different robust error functions (see section IV-A), and, as reported in [27], all functions yielded similar fitting errors when they converge.

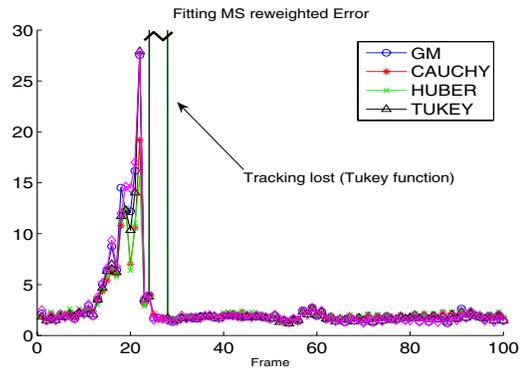


Fig. 7. Fitting error for robust error functions

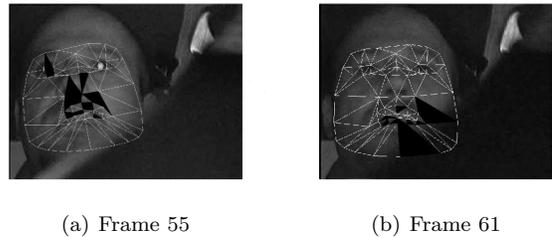


Fig. 8. Fitting to occluded face

Tracking losses can be detected by monitoring MSE values, and fast, incoherent changes in the instantiated shape (for example, if the shape is upside down, or with too many triangles outside of the image bounds). In figure 7, the test using Tukey robust function diverges, losing track for 4 frames. The model was able to recover by itself, as the shape stayed close to the face. When this is not possible, the model can be reinitialized with safe values for the warp parameters, and reasonable scale,

position and orientation. It should be noted that, as the eyes are the highest weighted areas of the image, tracking can be lost in other parts of the face while keeping the eyes correctly located, as shown in figure 8(a).

D. Computational efficiency

We have tested the computational efficiency of the tracker both in MATLAB and in a custom C/C++ implementation. As reported in the literature, the algorithm is very fast, thanks to the *inverse compositional* approximation. The template size is 110x110 pixels, and the image size is 320x240. Running on a Pentium-4 3.06GHz, the MATLAB code averages 2.4 frames per second. The C/C++ implementation is able to run in real-time (25 fps) under the same hardware configuration.

V. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a system based on Active Appearance Models for real-time face tracking of drivers. This system is an improvement over previous works for driver vigilance monitoring. The Active Appearance Model approach has shown to be very robust in many different situations, and overcomes several drawbacks of the previous system.

The results are encouraging, and we plan to include several improvements to the current implementation. In the near future, we will extend the construction system, to eliminate the need of handmarking, and to use generic models. We will also explore similar 2.5D and 3D techniques.

ACKNOWLEDGMENT

This work has been funded by grant TRA2005-08529-C02-01 (MOVICOM Project) from the Spanish Ministry of Science and Technology (MCyT), and S-0505/DPI/000176 (Robocity2030 Project) from the Science Department of the Comunidad de Madrid. J. Nuevo is also working under a researcher training grant from the Education Department of the Comunidad de Madrid and the European Social Fund.

REFERENCES

- [1] D. Royal, "Volume I - Findings; National Survey on Distracted and Driving Attitudes and Behaviours, 2002," The Gallup Organization, Tech. Rep., Mar. 2003.
- [2] A. Kircher, M. Uddman, and J. Sandin, "Vehicle control and drowsiness," Swedish National Road and Transport Research Institute, Tech. Rep. VTI-922A, 2002.
- [3] DaimlerChryslerAG. (2001, June) The electronic drawbar. [Online]. Available: <http://www.daimlerchrysler.com>
- [4] Seeing Machines. (2004, Aug.) Facelab transport. [Online]. Available: <http://www.seeingmachines.com/transport.htm>
- [5] A. C. (IST 2000-28062), "Awake-system for effective assessment of driver vigilance and warning according to traffic risk estimation," sep 2001-2004.
- [6] L. M. Bergasa, J. Nuevo, M. A. Sotelo, R. Barea, and E. López, "Real-time system for monitoring driver vigilance," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 1524–1538, March 2006.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 681–685, Jan. 2001.

- [8] I. Matthews and S. Baker, "Active appearance models revisited," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 135–164, November 2004.
- [9] S. Sclaroff and J. Isidoro, "Active blobs," in *Proceedings of the Sixth IEEE International Conference on Computer Vision*, 1998, pp. 1146–1153.
- [10] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *SIGGRAPH*, 1999, pp. 187–194.
- [11] A. Lanitis, C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 742–752, 1997.
- [12] S. Baker, I. Matthews, J. Xiao, R. Gross, T. Kanade, and T. Ishikawa, "Real-time non-rigid driver head tracking for driver mental state estimation," in *11th World Congress on Intelligent Transportation Systems*, October 2004.
- [13] D. Dinges, "F. perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance," Federal Highway Administration. Office of motor carriers, Tech. Rep. MCRT-98-006, 1998.
- [14] D. Royal, "Volume I - Findings; National Survey on Distracted and Driving Attitudes and Behaviours, 2002," The Gallup Organization, Tech. Rep. DOT HS 809 566, Mar. 2003.
- [15] Q. Ji and X. Yang, "Real-time eye, gaze and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, pp. 357–377, Oct 2002.
- [16] T. Cootes, C. Twining, V. Petrovic, R. Schestowitz, and C. Taylor, "Groupwise construction of appearance models using piece-wise affine deformations," in *Proc. British Machine Vision Conference*, 2005, pp. 879–888.
- [17] S. Baker, I. Matthews, and J. Schneider, "Automatic construction of active appearance models as an image coding problem," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 26, no. 10, pp. 1380–1384, October 2004.
- [18] S. Baker and I. Matthews, "Lucas-kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221–255, March 2004.
- [19] R. Gross, I. Matthews, and S. Baker, "Constructing and fitting active appearance models with occlusion," in *Proceedings of the IEEE Workshop on Face Processing in Video*, June 2004.
- [20] —, "Generic vs. person specific active appearance models," *Image and Vision Computing*, vol. 23, no. 11, pp. 1080–1093, November 2005.
- [21] S. Baker, R. Gross, I. Matthews, and T. Ishikawa, "Lucas-kanade 20 years on: A unifying framework: Part 2," Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-03-01, February 2003.
- [22] S. Baker, R. Gross, and I. Matthews, "Lucas-kanade 20 years on: A unifying framework: Part 3," Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-03-35, November 2003.
- [23] H. S. Sawhney and S. Ayer, "Compact representations of videos through dominant and multiple motion estimation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 8, pp. 814–830, 1996.
- [24] P. J. Huber, *Robust Statistics*, ser. Wiley Series in Probability and Mathematical Statistics. Wiley-Interscience, 1981.
- [25] V. Barnett, *Outliers in Statistical Data*, ser. Wiley series in probability and mathematical statistics. John Wiley & Sons, 1978.
- [26] Z. Zhang, "Parameter estimation techniques: A tutorial with application to conic fitting," *Image and Vision Computing Journal*, 1997.
- [27] B.-J. Theobald, I. Matthews, and S. Baker, "Evaluating error functions for robust active appearance models," in *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, April 2006, pp. 149 – 154.