Catadioptric Sensor for a Simultaneous Tracking of the Driver's Face and the Road Scene

Jean-Francois Layerle, Xavier Savatier, Jean-Yves Ertaud, El Mustapha Mouaddib

Catadioptric Sensor for a Simultaneous Tracking of the Driver's Face and the Road Scene

Jean-François Layerle[1,2], Xavier Savatier[1], Jean-Yves Ertaud[1] and El Mustapha Mouaddib[2]

[1] Research Institute for Embedded Systems (IRSEEM)
Technopôle du Madrillet, 76801 Saint-Etienne-du-Rouvray, France
{layerle,savatier,ertaud}@esigelec.fr
[2] MIS-University of Picardie Jules Verne
7 Rue du Moulin Neuf, 80000 Amiens, France
mouaddib@u-picardie.fr

Abstract. We propose the design of a new catadioptric sensor using two different mirror shapes for a simultaneous tracking of the driver's face and the road scene. We show how the mirror design allows the perception of relevant information in the vehicle: a panoramic view of the environment inside and outside, and a sufficient resolution of the driver's face for a gaze tracking. The design is performed in a raytracer to evaluate the resolution characteristics of the mirror. A face tracking method using Active Appearance Models is applied on real images acquired with our prototype to validate the design of this new sensor.

1 Introduction

Emergence of smart vision based systems in the vehicle allow new services for the driver by monitoring the environment inside and outside (the road scene and the driver's behavior)[1]. Among them, several devices are dedicated to the detection of the driver's hypo-vigilance, one of the major causes of road accidents. The driver's vigilance depends basically on the activity of its eyes (vision is one of the most important elements for safe driving [2]) and can be detected by observing the eyelid movements, the head orientations and the gaze direction.

Driver gaze monitoring with high performance can be achieved using head mounted devices but these systems are too intrusive in the driving task. The most popular non-intrusive system for driver monitoring is the FaceLAB[3] device. The system is composed of a stereo camera pair mounted on the dashboard to capture video images of the driver's face. The 3D head pose is determined in real time with a reliable eye gaze estimation ($\pm 3°$).

Remote sensors require that the systems are able to track all the head movements and poses. The driver's gaze estimation must be coupled with an efficient face tracking method to provide the 3D head pose. Ishikawa [3] describes an efficient driver gaze tracking system based on an Active Appearance Model (AAM)

[3] Seeing Machines: FaceLAB 4 (www.seeingmachines.com)

to track the whole head and to estimate the 3D head pose and the gaze direction with a good accuracy (5°).

However, the correlation between the gaze direction and the road scene requires additional vision systems to observe what the driver is looking at. Moreover, the use of perspective cameras whose main drawback is a reduced field of view does not allow the observation for the overall environment of the vehicle.

Catadioptric vision systems can be used in experimental vehicles to increase the field of view at different positions (on roof top, on sides, in front and back of the car)[1]. Huang [4] proposes the use of a catadioptric sensor in the vehicle to estimate the driver's head pose and the face orientation. The author highlights the interest of a catadioptric system in a car for the simultaneous capture of the driver's head pose, driving view and the surroundings of the vehicle with only one camera. Nevertheless, a limited factor of such a system is its low resolution, especially for an eye tracking on the driver's face.

An ideal vision system in the vehicle also needs to be non-intrusive and compact (to avoid distracting the driver in his task), able to observe the driver's eyes with a sufficient resolution and able to capture the road scene and the vehicle environment in real time.

In this paper, we propose the design of a new catadioptric configuration combining two different mirror shapes. One enables the capture of a panoramic road scene in the vehicle. Another allows to focus the view on the driver's face. A face tracking method using AAM is applied to real images acquired with our prototype to validate the design of this new sensor.

2 Sensor Design

The image formation of a catadioptric system is determined by the mirror shape and the associated camera. The design depends on world-to-image desired geometric properties.

Baker and Nayar [5] have described a class of mirrors that verify the theory of the single view point to preserve a central projection. The solution to this constraint induces a non-uniform resolution. Other works have proposed several methods to obtain constant resolution characteristics [6,7,8,9,10,11,12].

In both cases, the mirrors are rotationally symmetric in order to simplify their analytical model and making. Recently, Hicks [13,14] has proposed the fixed surface and vector field methods to design a mirror for a prescribed projection. The methods are based on vector fields normal to the mirror surface. Same approach was considered in [15] using spline basis function[4]. These methods enable the design of a very large class of mirrors (symmetric or asymmetric, with or without a single view point).

Among the solutions for mirror shape design, we propose to combine two methods. A classic approach [5] is used for the mirror capturing the panoramic view in the vehicle (a single view point will facilitate the projection of the driver's

[4] Matlab Toolbox REFLEX: REFlectors for FLEXible imaging and projection. In http://www1.cs.columbia.edu/CAVE/software/reflex/

gaze on the scene). The fixed surface method [13] is applied to a new prescribed projection problem in order to increase the resolution of the driver's face view with a second mirror. The position of this mirror is chosen for using a dead zone of the image plane as [16,17].

2.1 Single View Point to Capture the Vehicle Environment

The position of the sensor in the vehicle strongly affects the design of the mirror. We should be able to capture the driver's eyes and the road scene (vehicles, obstacles, pedestrians, road signs) without interfering with his vision field. Moreover, the position depends on vehicle features (dashboard size and equipments, windshield slope). As the estimation of the driver's gaze projection in the road scene is only required in this application, no constraint is set for the object resolution of the road. We choose to place the sensor at the center of the dashboard with the camera looking at the mirror upwards.

In practice, to capture a panoramic view, parabolic and hyperbolic mirrors are mainly used due to the position of the view point. Their panoramic field of view is specified by the upper-limit angle θ above the horizontal plane (Fig. 1).

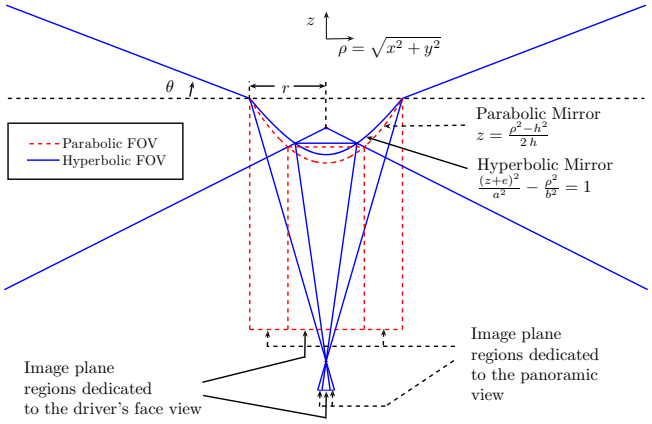


Fig. 1. Geometry of parabolic and hyperbolic mirrors to have a same field of view. Starting from the panoramic mirror renders in the vehicle (Fig. 3), we can specify the image plane regions dedicated to the panoramic road scene view and the focus view on the driver's face. We deduce the size of the second mirror for each projection.

From θ and r, we can describe the mirror parameters (Eq. 1 for the parabolic mirror and Eq. 2 for the hyperbolic mirror).

$$h = r\left(\sqrt{\tan^2\theta + 1} - \tan\theta\right) \tag{1}$$

$$a = \frac{b^2}{r}\left(\sqrt{\tan^2\theta + 1} + \tan\theta\sqrt{1 + \frac{r^2}{b^2}}\right) \tag{2}$$

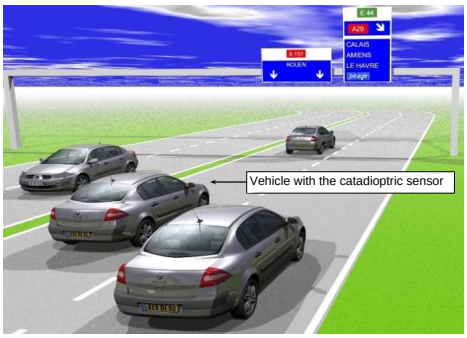


Fig. 2. A portion of motorway in which one vehicle contains the catadioptric sensor. The sensor is positioned at the center of the vehicle dashboard.

A 3D virtual environment[5] of a driving situation (Fig. 2) is used to simulate the images formed by these panoramic mirrors (Fig. 3). A panoramic scene view allows the capture of whole environment in and out the vehicle (driver, occupants, road and blind spot). Parabolic and hyperbolic mirrors have a similar render for a same θ and r. The choice of the best configuration depends only on our second mirror dedicated to the observation of the driver. It can be noted that the center of the image can be discarded (vehicle dashboard and below, reflection of the camera itself). Thus, to maintain a compact system, this dead zone can be used to set the second mirror. By fixing this image plane area for the driver's face, we can deduce the size of the second mirror (Fig. 1).

(a) Parabolic Mirror with $h = 2.2199$

(b) Hyperbolic Mirror with $a = 2.8186$ and $b = 2.3413$

Fig. 3. Panoramic scene views in the vehicle for the 2 projections with $r = 3$ cm, a panoramic field of view of $360° \times 214°$ ($\theta = 17°$) and a resolution of 640×480 pixels.

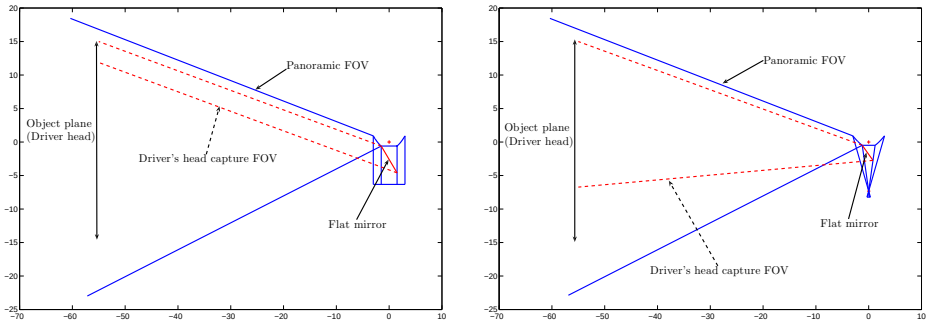
[5] 3D models of Renault Megane and driver by F. Heillouis, V. Gus'ev, A. Kator and J. Legaz. (http://www.dmi3d.com and http://www.katorlegaz.com)

In these images (without noise), the resolution of the driver's face in the panoramic view is 70×70 pixels and the eyes is 3×3 pixels. The accuracy of gaze tracking depends upon the number of pixels per eye. Knowing that the eyeball rotation is about $70°$ in both azimuth and elevation angles, the possible accuracy could be $23°$ for an error of 1 pixel. In real conditions, it is necessary to increase the resolution of the driver's face for a greater accuracy as the one obtained in [3] (an accuracy of $5°$ requires a resolution of 15×15 pixels per eye).

2.2 A Second Mirror to Capture the Driver's Face

A high accuracy for gaze tracking involves a narrow field of view on the driver's eye (a high magnification) while the ability to track a maximum of head movements requires a wide field of view (face tracking is necessary to preserve driver's freedom of movement, estimate the driver's head pose and extract the eye regions).

The position of the catadioptric sensor induces the position of the driver's head that we need to capture. In our case, we want to project an object plane (around the driver's head) on the center of the image plane (the previous dead zone). This object plane is positioned at 55 cm from the sensor in the vehicle with a size of 30×30 cm (to track all head movements). The projection can be obtained by placing a second mirror under the first mirror. One simple way could be the use of a flat mirror. Figure 4 illustrates the geometry of the sensor combining conic and plane surfaces.



(a) Parabolic mirror with a flat mirror. We can only capture a plane of 5×5 cm.

(b) Hyperbolic mirror with a flat mirror. We can capture a larger plane (20×20 cm) but not the complete desired region.

Fig. 4. Use of a flat mirror to focus on the driver's face

In both projection cases (telecentric and perspective), a flat mirror is not sufficient to capture the expected object plane (the field of view available with this second mirror depends only on the size of the plane surface). We need to design a specific curved mirror shape to increase this field of view.

2.3 Fixed Surface Method to Increase the FOV of the Face Capture

Hicks [18] has introduced the problem of designing a mirror for imaging prescribed portion of space. In general, there is no reflective surface to achieve the desired projection but an optimal approximate solution can be obtained. We propose to design this mirror in the orthographic case using the fixed surface method [13] which reduces the problem to a system of two differential equations. The results are compared with the images formed by mirror shapes designed with the REFLEX Toolbox.

Prescribed Projection The prescribed projection problem of our system is illustrated in Fig. 5. The size and position of the object plane are defined with β, k and θ (driver head center position is not aligned with y-axis and requires to take into considerations the angle θ). The size and the position of the image plane are defined with α and c. We search for a reflective surface $z = f(x, y)$ such that we can capture the object plane by a scaling transformation. A ray intersecting the image plane at a point $(x, y, -c)$, starts from the object plane at a point $(\frac{\beta}{\alpha}x, -k\cos\theta - \frac{\beta}{\alpha}y\sin\theta, -k\sin\theta + \frac{\beta}{\alpha}y\cos\theta)$ and strikes the reflective surface at the point $m(x, y, f(x, y))$.

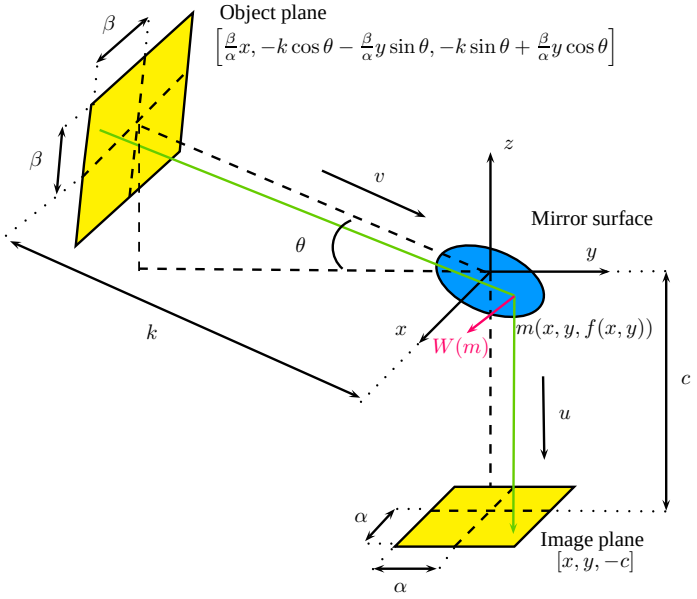


Fig. 5. Prescribed projection for the second mirror. The reflective surface is defined such that we can capture the object plane (driver's head) by a scaling transformation.

Fixed Surface Method Assuming an orthographic projection (Fig. 5), the ray between the mirror surface and the camera image plane is parallel to the vector $u = [0, 0, -1]$. The direction vector of the ray between the object plane and the mirror is noted v. The normal vector to the surface at $m(x, y, f(x, y))$ is defined by the gradient of $f : W(m) = [-f_x, -f_y, 1](f_x^2 + f_y^2 + 1)^{-\frac{1}{2}}$ with $f_x = \frac{\partial f}{\partial x}$ and $f_y = \frac{\partial f}{\partial y}$.

For a specular mirror, the law of reflection implies that the vectors u, v and $W(m)$ are coplanar and that the angle of incidence must be equal to the angle of reflection. We can define the coordinates of v as Eq. 3.

$$v = u - 2\,(u \cdot W(m))W(m) = \begin{bmatrix} -2\,f_x(f_x^2 + f_y^2 + 1)^{-1} \\ -2\,f_y(f_x^2 + f_y^2 + 1)^{-1} \\ 2\,(f_x^2 + f_y^2 + 1)^{-1} - 1 \end{bmatrix} \tag{3}$$

The ray from the object plane striking the mirror can be expressed with the distance d between an object plane point and the mirror (Eq. 4).

$$\begin{bmatrix} x \\ y \\ f(x,y) \end{bmatrix} = \begin{bmatrix} \frac{\beta}{\alpha}x \\ -k\cos\theta - \frac{\beta}{\alpha}y\sin\theta \\ -k\sin\theta + \frac{\beta}{\alpha}y\cos\theta \end{bmatrix} + d\,v \tag{4}$$

A simplification of Eq. 4 is proposed in Eq. 5 by moving the object plane to infinity ($k \to \infty$ and $\beta = \lambda k$).

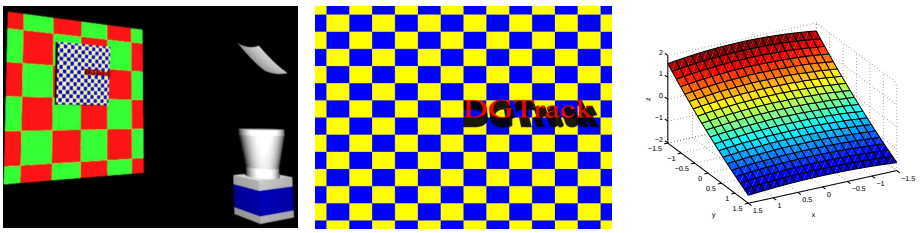
$$\begin{cases} f_y\frac{\lambda}{\alpha}x + (\cos\theta + \frac{\lambda}{\alpha}y\sin\theta)f_x = 0 \\ 2f_y(-\sin\theta + \frac{\lambda}{\alpha}y\cos\theta) - (\cos\theta + \frac{\lambda}{\alpha}y\sin\theta)(1 - f_x^2 - f_y^2) = 0 \end{cases} \tag{5}$$

There is no exact solution to this system. Hicks [13] proposed an approximated solution f considering a polynomial form $f(x, y) = \sum_{i=1}^{n} c_i(-\frac{\lambda}{2}x^2 + \alpha y\cos\theta + \frac{\lambda}{2}y^2\sin\theta)^i$ that satisfies the first equation of the system (Eq. 5). The coefficients c_i are found by minimizing the following functional (Eq. 6).

$$\mathcal{G}(f) = \int_{-\alpha}^{\alpha}\int_{-\alpha}^{\alpha}\left(2f_y(-\sin\theta + \frac{\lambda}{\alpha}y\cos\theta) - (\cos\theta + \frac{\lambda}{\alpha}y\sin\theta)(1 - f_x^2 - f_y^2)\right)^2 dxdy \tag{6}$$

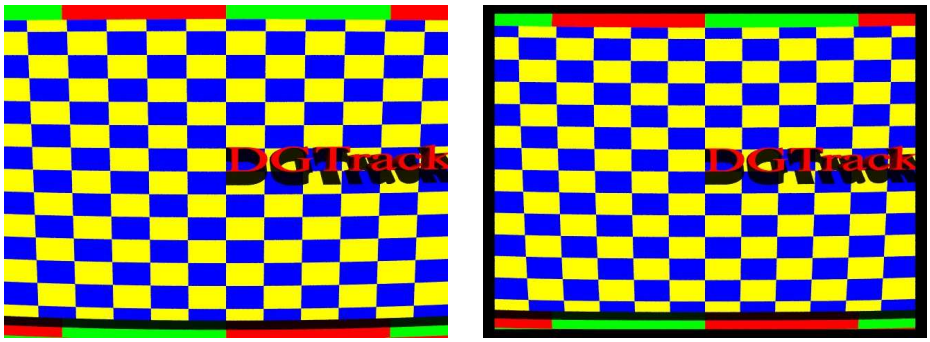
For the fixed positions of the catadioptric and the driver, we can define the values α, β, θ and k. We use a cubic polynomial form f_3 and apply the Nelder-Mead simplex method to minimize the previous functional (Eq. 6).

Results for the Driver's Face Capture The image formation is evaluate in Fig. 6a with a checker pattern for the object plane (Fig. 6b). The geometry of the mirror shape designed for $\alpha = 1.5$ cm, $\beta = 15$ cm, $\theta = 6°$ and $k = 55$ cm is illustrated on Fig. 6c. The image formed with this mirror (Fig 7a) has a constant resolution with an image projection error on the edges. A similar image is obtained with the REFLEX Toolbox (Fig. 7b).



(a) Environment to evaluate the mirror design. (b) Object plane defined by a checker pattern. (c) Geometry of the mirror shape designed.

Fig. 6. Evaluation of the mirror shape design.



(a) Fixed surface method. (b) REFLEX Toolbox.

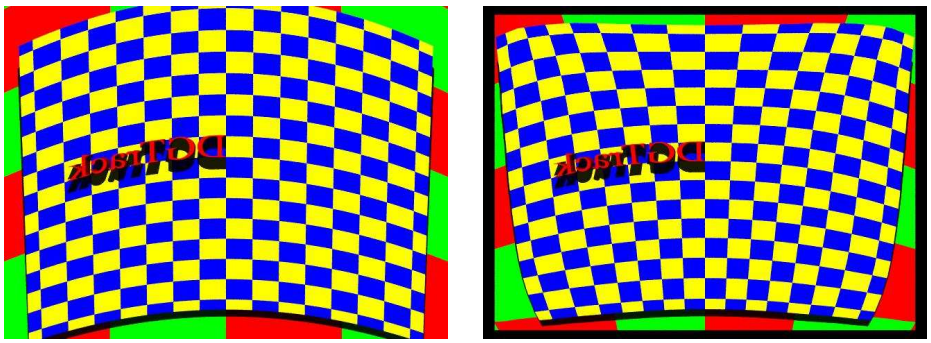
Fig. 7. Images formed with the two design methods for the prescribed projection.

It can be noted that the images obtained with the designed mirror are not reversed along the x-axis. We propose to modify the prescribed projection to obtain reverse images (the ray intersecting the image plane $(x, y, -c)$ starts from the object plane at a point $(-\frac{\beta}{\alpha}x, -k\cos\theta - \frac{\beta}{\alpha}y\sin\theta, -k\sin\theta + \frac{\beta}{\alpha}y\cos\theta)$) (Fig 8).

We observe more distortions for the images formed by the mirror designed with the two methods. Nevertheless, the images are formed as the ones provided by a traditional mirror.

2.4 Complete Mirror and Prototype

The position of our catadioptric device is not exactly in front of the driver. We need to rotate our mirror around z-axis (in our case, the rotation is about 22°). A simulation of the complete catadioptric sensor with reversed images of the driver face is proposed in Fig. 9.



(a) Fixed surface method. (b) REFLEX Toolbox.

Fig. 8. Reverse images like a traditional mirror.

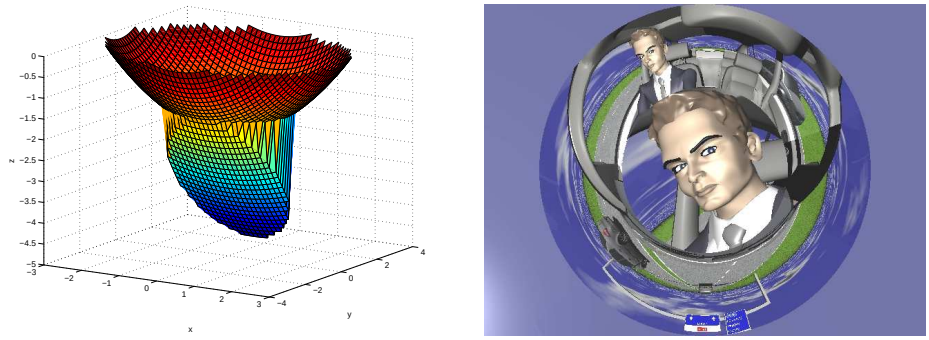


Fig. 9. Geometry and simulation of the complete catadioptric system in the vehicle.

Our mirror prototype is a parabolic surface and the mirror designed to have reverse images of the driver. An illustration of the images captured in a real environment is presented in Fig. 10. We use the RemoteReality[6] OneShot optic and a CMOS camera with a resolution of 640×480 pixels. We obtain similar results in ray-tracing simulations and in real conditions for the driver's eye resolution (3×3 pixels in the panoramic view and 13×13 pixels in the central view of the driver (equivalent to an accuracy of 5.3°)).

[6] RemoteReality: http://www.remotereality.com/

(a) (b)

Fig. 10. On the left, we present our mirror prototype combining two different shapes and on the right an image acquired in the laboratory.

3 Evaluation on Real Images: Tracking the Driver's Face

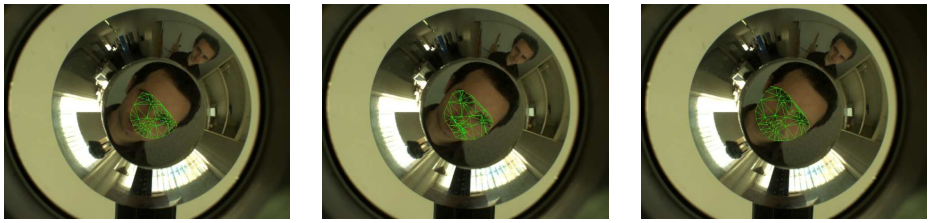
Face tracking is the first main part of a driver's gaze direction estimator for finding and extracting the region of eyes and providing the head pose. A good estimation of the 3D facial structure requires a good location of landmarks on the image. We propose to verify that a face tracking algorithm is applicable to the images obtained by our sensor (the central view of the driver's face has a resolution of 200×200 pixels).

In recent years, successful face tracking algorithms have been proposed using face models to track the face as a single object (feature based methods are very local and assume limited movements of the head). Active Appearance Models (AAM), proposed by [19], provide a statistical model of the faces by using a nonlinear parametric approach in terms of pixel intensity. A shape of an AAM is defined as a vector of landmarks forming a representative mesh of the object tracked. The shape can be expressed as a deformation of a mean shape by a linear combination of shape vectors. The appearance of an AAM is an image defined over the pixels of the mesh. The appearance can be expressed in the same way as the shape with a variation of a mean appearance by a linear combination of images. AAM requires two main steps: build the statistical models of shape and appearance and fit the deformable model to an image:

A Procrustes analysis and a Principal Component Analysis [20] are applied on training images with manual annotation of the shape, to compute the mean shape, appearance and their variability. The successfulness of the model fitting algorithm is dependent on the annotated training base.

Several works have been proposed to deform these models for face tracking. Matthews and Baker [21] have introduced a fast fitting algorithm based on the Lucas-Kanade image alignment algorithm on which we based our validation. It uses an inverse compositional approach to optimize the AAM alignment and deformation.

We use 50 images captured with our sensor to describe our statistical model. An initial model is created with the mean shape and the mean appearance (Fig. 11a). From this position, the AAM is iteratively deformed and aligned until the convergence to the face of the image is reached (Fig. 11b and 11c). Despite distortions introduced by the central mirror, face tracking is still possible and can be achieved using standard algorithms, originally designed for perspective images, without any modifications.



(a) Initial shape (b) Intermediate shape (c) Final shape

Fig. 11. Active Appearance Model Fitting on the driver face focus

4 Conclusion

In this paper, we presented a new compact catadioptric sensor combining two different mirror shapes. The sensor architecture offers new field of applications in the vehicle (hypo-vigilance, study of driver behavior and road infrastructure, biometric driver identification, video-surveillance).

The image reflected by the mirrors combine a large-field of view of the road scene as well as a high resolution image of the driver's face. This resolution is sufficient to apply standard face tracking algorithms without any modifications. Thus, our central mirror overcomes limitations of the omnidirectional sensors when gaze tracking is required.

Another interesting point of our catadioptric device is the possibility to contemplate, on the same image, two different projections of the driver's head. Therefore, the design of the mirror allows a stereo-vision of the driver's face. Future works will be focused on calibration, gaze tracking and gaze projection in the road scene.

References

1. Trivedi, M.M., Gandhi, T., McCall, J.: Looking-in and looking-out of a vehicle : Computer-vision-based enhanced vehicle safety. IEEE Transactions on Intelligent Transportation Systems **8** (2007) 108–120

2. Sivak, M.: The information that drivers use: is it indeed 90% visual? Perception **25** (1996) 1081–1089
3. Ishikawa, T., Baker, S., Matthews, I., Kanade, T.: Passive driver gaze tracking with active appearance models. In: Proceedings of the 11th World Congress on Intelligent Transportation Systems. (2004)
4. Huang, K.S., Trivedi, M.M., Gandhi, T.: Driver's view and vehicle surround estimation using omnidirectional video stream. In: Proceedings of the IEEE Intelligent Vehicles Symposium (IV'03). (2003) 444–449
5. Baker, S., Nayar, S.K.: A theory of single-viewpoint catadioptric image formation. International Journal of Computer Vision **35** (1999) 175–196
6. Chahl, J.S., Srinivasan, M.V.: Reflective surfaces for panoramic imaging. Applied Optics **36** (1997) 8275–8285
7. Conroy, T.L., Moore, J.B.: Resolution invariant surfaces for panoramic vision systems. In: Proceedings of the 7th IEEE International Conference on Computer Vision (ICCV'99). (1999) 392–397
8. Hicks, R.A., Bajcsy, R.: Reflective surfaces as computational sensors. In: Proceedings of the 2nd Workshop on Perception for Mobile Agents. (1999) 82–86
9. Gaechter, S., Pajdla, T., Micuvik, B.: Mirror design for an omnidirectional camera with a space variant imager. In: Proceedings of the 2nd Workshop on Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing (OMNIVIS'01). (2001) 99–105
10. Gaspar, J., Decco, C., Okamoto, J.J., Santos-Victor, J.: Constant resolution omnidirectional cameras. In: Proceedings of the 3rd IEEE Workshop on Omnidirectional Vision (OMNIVIS'02). (2002) 27–34
11. Marchese, F.M., Sorrenti, D.G.: Mirror design of a prescribed accuracy omnidirectional vision system. In: Proceedings of the 3rd IEEE Workshop on Omnidirectional Vision (OMNIVIS'02). (2002) 136–142
12. Hicks, R.A., Perline, R.K.: Equiresolution catadioptric sensors. Applied Optics **44** (2005) 6108–6114
13. Hicks, R.A., Perline, R.K.: The fixed surface method for catadioptric sensor design. Technical report, Drexel University Department of Mathematics (2004)
14. Hicks, R.A.: Designing a mirror to realize a given projection. Journal of the Optical Society of America **22** (2005) 323–330
15. Swaminathan, R., Grossberg, M., Nayar, S.: Designing mirrors for catadioptric systems that minimize image errors. In: Proceedings of the 5th Workshop on Omnidirectional Vision, Camera Networks and Non-classical cameras (OMNIVIS'04). (2004)
16. Yagi, Y., Egami, K., Yachida, M.: Map generation for multiple image sensing sensor miss under unknownrobot egomotion. In: Proceedings of the IEEE Conference on Intelligent Robots and Systems (IROS'97). (1997) 1024–1029
17. Yin, W., Boult, T.E.: Physical panoramic pyramid and noise sensitivity in pyramids. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'00). (2000) 90–97
18. Hicks, R.A., Perline, R.K.: Geometric distributions for catadioptric sensor design. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'01). (2001) 584–589
19. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Proceedings of the 5th European Conference on Computer Vision (ECCV'98). (1998) 484–498

20. Cootes, T., Taylor, C.: Statistical models of appearance for computer vision. Technical report, Imaging Science and Biomedical Engineering, University of Manchester (2004)
21. Matthews, I., Baker, S.: Active appearance models revisited. International Journal of Computer Vision **60** (2004) 135–164