Statistical Models of Appearance for Eye Tracking and Eye-Blink Detection and Measurement

Ioana Bacivarov, Student *Member,* IEEE, Mircea Ionita, Student *Member,* IEEE
and Peter Corcoran, Senior *Member,* IEEE

Abstract — *A statistical Active Appearance Model (AAM) is developed to track and detect eye blinking. The model has been designed to be robust to variations of head pose or gaze. In particular we analyze and determine the model parameters which encode the variations caused by blinking. This global model is further extended using a series of sub-models to enable independent modeling and tracking of the two eye regions. Several methods to enable measurement and detection of eye-blink are proposed and evaluated. The results of various tests on different image databases are presented to validate each model*[1].

Index Terms — eye blinking, eye tracking, active appearance model, digital imaging

I. INTRODUCTION

Recently digital cameras have begun to incorporate real-time face tracking technology, see for example [30], [31]. This new technology can be used to improve subject focus, image exposure and potentially image composition by providing user feedback. As the capabilities of digital imaging devices continue to improve it is clear that even more detailed analysis can be applied to the tracked faces within a digital image. In particular the occurrence of undesirable eye-blinking or partially closed eyes is the cause of many spoiled photographs.

Active Appearance Models [9] offer a form of 2D affine face model which can quickly match the texture and shape of a detected face region. Similar statistical models can be developed for sub-regions of a detected face such as the eyes [10]. Being one of the most expressive features of the human face, information about the eyes can play an important role in consumer applications. Examples include the analysis of facial expressions [1], computer animation [2], [3], driver awareness systems [4], [8], film and advertising industry [7] or assisting people with disability through an eye-based communication interface [5], [6]. For this reason, eye-related applications have received a good deal of attention in the literature.

Now current eye models function well for open eyes, but they are not robust against physiological changes in eye appearance, such as blinking or gazing. A determination of eye state, i.e. open or closed, is more complex than simply determining the eye location. This is due in part to the small size of the eyes relative to the face region and the weak contrast between the eye region and the surrounding skin.

In this article, a proof-of-concept model for the eye region is constructed, using Active Appearance Modeling (AAM) techniques [9] with the main goal to determine model parameters which can measure a degree of eye-blink. An initial model employing two dependent eye-regions is then expanded to provide a more self-consistent model by using the component-based technique presented in [12], [13]. These are adapted to our global model. After developing a reliable eye-model, a blink detector is proposed. Although AAM techniques were used by Hansen et al. [10] to model the eye region for gaze tracking, these techniques have not yet been applied to model eye blinking.

This paper is organized as follows. Section II presents a short survey of the existing methods in eye tracking and blink detection. The developed model for the eye-region is detailed in Section III, together with a brief theoretical introduction to statistical models of appearance and the extension of the model to the component-based model adapted for the eyes. Section IV gives an overview of the proposed system, including the blink detection method and a discussion of the requirements for an in-camera implementation. Section V presents some experimental results, while in section VI we draw some conclusions and discuss some possible future work.

II. RELATED WORK

Eye tracking techniques can be classified in two categories: intrusive methods which require direct contact with the eyes and non-intrusive methods which avoid any physical contact with the user. The latter is of a more interest for consumer applications, the advantage being the fact that the user is not constrained by uncomfortable equipment. A good survey of non-intrusive methods is offered in [16], that broadly classifies them into three categories: model-based [5], [6], [10], [21-24], appearance-based [16-20] and feature-based methods [24], [26], [27].

In model-based methods, firstly a generic eye model is designed and then a template matching is used to search the eye images. Deformable templates methods, as in [10], [22-24], are commonly used in this case.

[1] This work is supported under the Innovation Partnership Program (IP/2006/0361) of Enterprise Ireland under the National Development Plan. It is co-funded by FotoNation (Ireland) Ltd.

Peter Corcoran is with the Dept. Electronic Engineering, National University of Ireland, Galway (e-mail: peter.corcoran@nuigalway.ie).

Ioana Bacivarov and Mircea Ionita are with the Dept. Electronic Engineering, Natuonal University of Ireland, Galway (e-mail: ioana@wuzwuz.nuigalway.ie and mircea@wuzwuz.nuigalway.ie).

Manuscript received June 16, 2008

In appearance-based methods, the model is learned from a large set of original images, so it does not need to build further models for objects and no additional features need to be extracted. Infra-red (IR)-based methods can be applied successfully in the appearance-based algorithms [16], even solving traditional problems like the presence of glasses and day light. Methods based on active light are often used for such eye-tracking methods, as in [20] where the motion information between two consecutive frames is considered.

Feature-based methods are based on tracking individual features of the eyes, by exploring their characteristics such as edge and intensity of iris or their color and eye corners. A popular method is the extraction and tracking of the eyelids [26], although this becomes unreliable when the head is moving. In [27] a different feature is used; instead of detecting the eyes, the authors propose the detection of the mid-point between the two eyes, which is believed to be more stable and easier to spot. Eyes are subsequently detected as two dark parts, symmetrically located on each side of this mid-point.

Eye tracking should be robust to blinking or gazing, otherwise it could lose track due to involuntary actions of the subject. The eye blink reflex is one of the fastest and most consistent of human reflexes, and therefore a problematical topic. Consequently, numerous methods to detect blinking have been proposed. We enumerate only some recent work, from a considerable body of literature.

Using images from a high-resolution video-camera, Moryiama et al. [17] divides the eye-region in two: upper and lower half. By measuring the illumination intensity of each half over time, they determine the blinking when the two curves cross.

For video sequences, blinking can be detected from frame differencing by detecting the motion between two consecutive frames [19]. Head movements have to be taken into consideration as well. In [18], it is suggested to discriminate the second order change, such as the eye blink, from a first order change, such as the head movement observed in an image. After the pixels corresponding to the head motion have been filtered out by the second order change detection procedure, which uses three consecutives frames, the remaining motion pixels are used to decide whether a blink occurred or not.

Other blink detectors based on iris tracking [20] have been developed, in which the disappearance of the iris implies blinking.

Template matching techniques have also been used [6], [21]. Grauman et al. [21] calculate correlation scores between the actual eye detected by "blink-like" motion and a closed eye template. In [6], the creation of the open eye template is created on-line. The detection of blinking and the analysis of blink duration are based solely on observing the correlation scores generated by the tracking at the previous step using the on-line template of the user's eye.

In summary, traditional image-based eye tracking approaches track the eyes and detect eventual blinking by exploiting the differences between frames and making use of special characteristics of the eye such as dark pupil, white sclera, circular iris, eye corners, eye shape, etc.

The current work belongs to the model-based category, but ideas from the appearance-based methods are also incorporated. The main advantage of employing a statistical model of appearance is that the shape and texture of the eye are exploited directly as they appear in the original image. The method is also invariant to rotation, scale or pose and it is capable of successfully synthesizing eyes in unseen pictures.

III. Model Description

The appearance of the eye region is represented by a statistical model trained using a set of annotated image examples. The training set contains subjects with open and closed eyes, the model being able to synthesize all the states in between. A new image can be interpreted by finding the best match of the model to the image data. In this section a statistical model of appearance for the eyes is developed and fine-tuned by a component-based method. The manner in which the model is constructed will permit us to extract eye blinking parameters and consequently to develop, further on, a blink detector (section IV).

A. Statistical Models of Appearance

Statistical Models of Appearance [9] represent both shape and texture variations and correlations between them. The desired shape to be modeled is annotated by a number of landmark points. The shape is defined by the number of landmarks chosen to best depict the contour of the object of interest, in our case the eye region. A shape vector is given by the concatenated coordinates of all landmark points and may be formally written as, $s=(x1,x2,\ldots,xL,y1,y2,\ldots,yL)^T$, where L is the number of landmark points.

The shape model is obtained by applying Principal Component Analysis (PCA) on the set of aligned shapes:

$$s = \bar{s} + \varphi_s b_s \,, \tag{1}$$

where $\bar{s} = \frac{1}{N_s}\sum_{i=1}^{N_s} s_i$ is the mean shape vector, and N_S is the number of shape observations; φ_s is the matrix having the eigenvectors as its columns; b_s defines the set of parameters of the shape model.

The texture, defined as the pixel values across the object of interest, is also statistically modeled. Face patches are first warped into the mean shape based on a triangulation algorithm.

Then a texture vector $t=(t1,t2,\ldots,tp)^T$ is built for each training image by sampling the values across the warped, i.e. shape normalized patches, with p being the number of texture samples.

The texture model is also derived by means of PCA on the texture vectors:

$$t = \bar{t} + \varphi_t b_t , \qquad (2)$$

where $\bar{t} = \frac{1}{N}\sum_{i=1}^{N_t} t_i$ is the mean texture vector, with N_t as the number of texture observations; φ_t is the matrix of eigenvectors, and b_t the texture parameters.

The sets of shape and texture parameters $c = \begin{pmatrix} W_s b_s \\ b_t \end{pmatrix}$ are used to describe the overall appearance variability of the modeled object, where W_s is a vector of weights used to compensate the differences in units between shape and texture parameters.

After a statistical model of appearance is created, an Active Appearance Model (AAM) algorithm can be employed to fit the model to a new image. AAM is a fast technique used to interpret unseen images using the appearance model, by finding the best match of the model to the image. Therefore, the algorithm allows us to find the parameters of the model which generate a synthetic image as close as possible to the target image. The whole problem is treated as an optimization problem in which we want to minimize the difference between the new image and the image synthesized by the appearance model.

B. Eye Model Design - Extraction of Blinking Parameters

The model we have developed offers a detailed analysis of the eye region in terms of degree of eyelid opening, the position of the iris and the shape and texture of the eyes which are related to the identity of the subject.

A statistical model for the eye region was built using AAM techniques, following the procedure given in section III.A. The model is composed of the following elements: eyes, eyebrows and the iris. The latter allows the model to simulate eye-gaze in addition to eye-blink. In order to describe the desired shape, the model was annotated as shown in Fig. 1.

It is known that left and right eyes will typically blink in a closely synchronized fashion. Thus it would be sufficient to capture the trace for one eye in order to measure eye-blink for a single subject. However with a detailed two-eye model, we can determine a lot of additional information and distinguish other eye states such as winking. The large number of points, i.e. annotated landmarks, and the details of the eyebrows and eyelids considered in this model description are motivated by the fact that any extra information is useful in analyzing the state of the eye-regions.

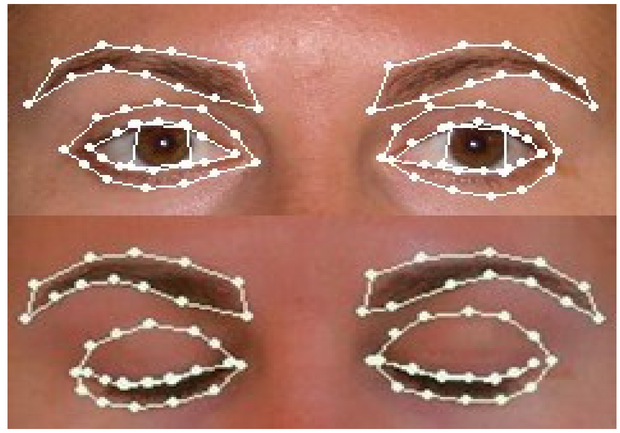

Fig. 1. Image annotations for open and closed eyes.

The choice of the region to model depends on the application, being a compromise between the speed of fitting the model, the number of model parameters and the amount of useful information obtained from the model. The same labeling was used to model both open and closed eyes, as shown in Fig. 1.

In order to model eye-blink, we have extended the conventional model to incorporate two different templates for open and closed eyes within the same model. This is achieved through using *the same number of points* for the two templates, mapped in the same order, where the upper and lower inner eyelids overlap. This overlapping is represented in Fig. 2 (upper left corner) by the red arrows: the first point is overlapped over the third one, point 12 over the point 4 and so on. The procedure emulates the physiological blinking action, when the upper eyelid closes over the lower eyelid.

One significant goal of this model is the extraction of the model parameters that encode the eye-blink information. Blinking uses mainly shape information, i.e. the eyelid shape is modified between the two states.

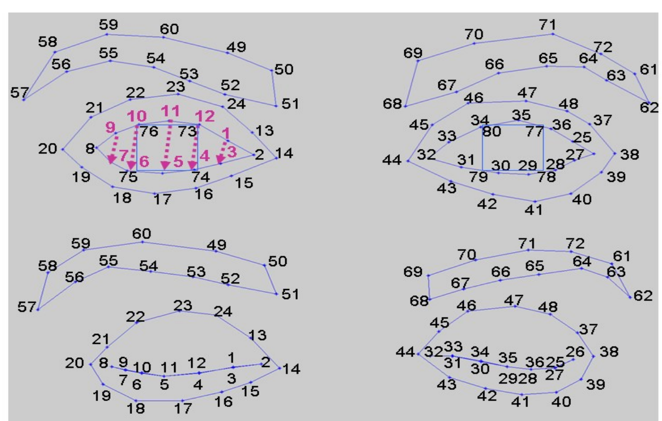

Fig. 2. The two annotations for the open/closed eyes (the latter obtained by overlapping the points from the upper inner eyelid).

Depending to a small extent on the complexity of the training set, shape parameters can be differentiated between closed and open eyes, gazing and pose, or shape-of-the-eye parameters.

Typically, only a small number of the shape parameters accounts for the eye-blink action, which corresponds to a large amount of the total variance of the shape parameters. The remaining model parameters encode other types of variability, such as pose or identity changes.

Fig. 3 displays an average over the texture (upper graph) and shape parameters (lower graph) of the eyes model for both images with open and closed eyes, respectively. There is a clear differentiation between the two curves for the first parameters and consequently between the states of the eyes. First texture parameter and the first two shape parameters were identified as depicting the blinking.

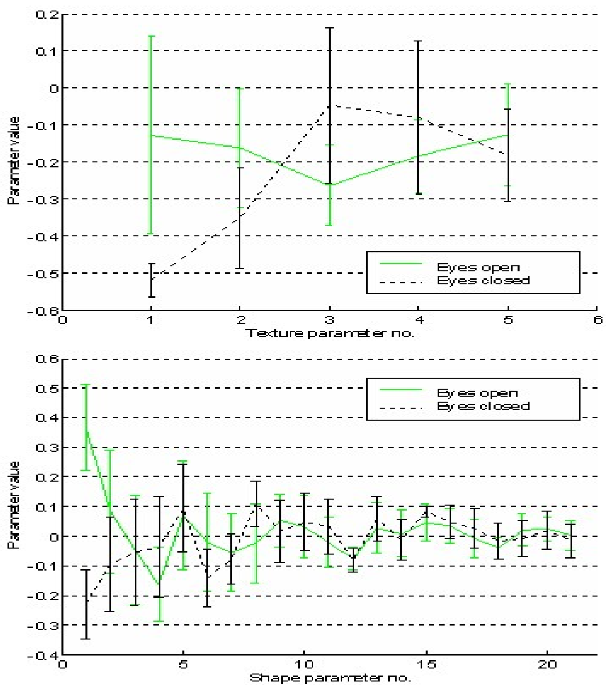

Fig. 3. Average values and the unit of standard deviation for the texture and shape parameters tested on images with eyes open (green curve) and closed (black curve), respectively.

A problem which generally appears in face modeling is that of robustness to different poses. Same as for full face modeling [15], the eyes model can be designed in order to include pose variability. To address this issue, we have included different poses of the eyes in the training set. Therefore, the model was enriched with a set of shape parameters, mainly accounting for pose variation. Thus the equation describing the shape model in (1) can be extended to the explicit form:

$$s = \bar{s} + \varphi_{blink} b_{blink} + \varphi_{pose} b_{pose} + \varphi_{eye\_shape} b_{eye\_shape} \quad (3)$$

Fig. 4 shows the effect of varying with +/- one unit of standard deviation for some representative parameters of the model.

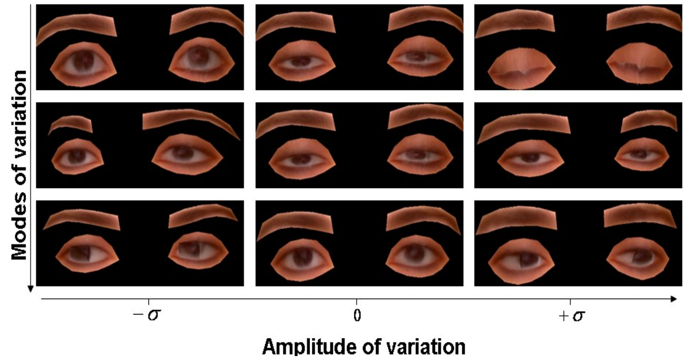

Fig. 4. Modes of variation for the eye model by +/- one standard deviation from the mean. The variation from the mean of the blink parameters is represented in the first row, while the variation of pose and gazing are represented in the last rows.

C. Component-Based AAM for the Eye Model

Although AAM models are a powerful tool for image interpretation, they present several drawbacks when used as a global appearance model. In particular there can be problems with under-trained models, as they can be too closely constrained to the model variations learned during the training phase.

Component-based AAM, as described in [12], [13] comes as a solution to this particular drawback of AAM. It combines a global model with a series of sub-models. These sub-models are typically component parts of the object to be modeled. This approach benefits from both the generality of a global AAM model and the local optimizations provided by its sub-models.

In the literature, component models are applied to face objects [12], [13]. We now extend this approach to our eye model and apply it as a strong initialization algorithm for robust eye tracking (the implementation details are given in section IV).

The component-based AAM, as customized for the eye-region, employs two distinct eye models which find their optimal positions and shape independently of each other. There are situations, especially when dealing with large pose variations, plane rotation, occlusions or strong facial expressions when the 2D projection of the eyes loses the property of global symmetry.

A global AAM model was trained using information extracted from the two eyes in the two states, i.e. open and closed, as described in III.B. Then a separate sub-model is created describing a single eye in all its states: i.e. open, closed or different variations in between. The sub-model is designed following the same principles as the global model. The two models, i.e. global model and sub-model, are trained and independently generated using the same AAM training procedure.

One valuable aspect of the eyes model is the symmetry between the two eyes. This characteristic permits, starting from one eye model, e.g. left eye, to auto-generate an equivalent model for the other eye (right eye, respectively), simply by mirroring the data. The advantage of having less memory space requirements is clear, as only one set of eye-model characteristics needs to be stored.

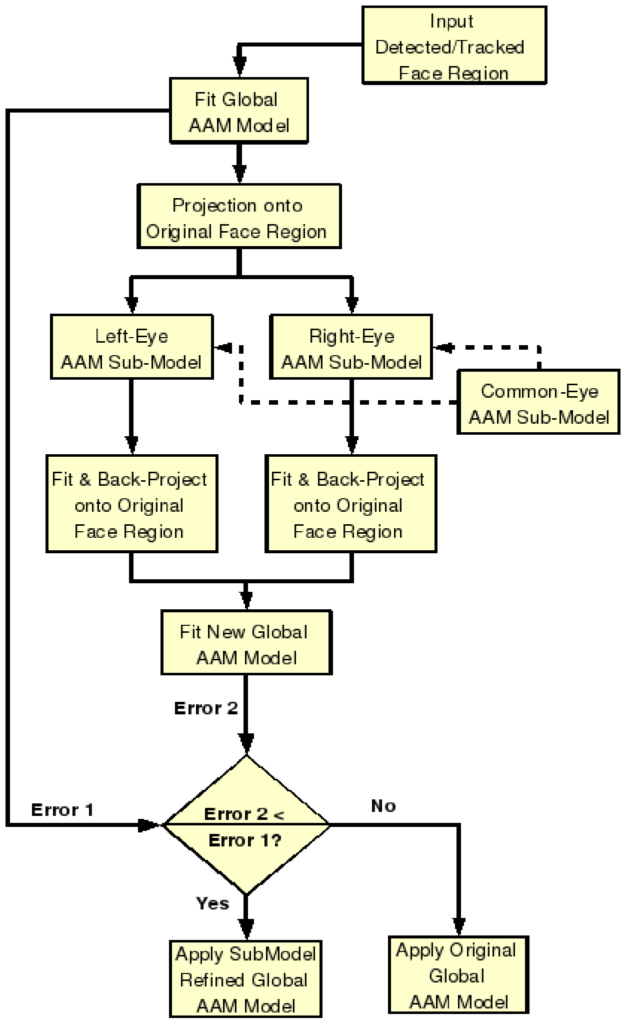

Fig. 5. Fitting algorithm for component based AAM eye model.

Fig. 5 describes the fitting algorithm adapted for the component-based procedure. At each iteration, the sub-model is inferred from the global model. Its optimum is detected by an AAM fitting procedure. Then the fitted sub-model is projected back into the global model. In a first step, the global AAM is fitted, roughly locating the two eyes. The points of the global model which correspond to the sub-models are first extracted to form the two shape vectors, providing a good initialization for each sub-model which is next fitted using the AAM method. As a second step we reintegrate the fitted points of the two independent sub-models back into the global model. Another projection of the global shape vector onto its principal components space is required. This step is necessary in order to constrain the two independently obtained eyes such that they remain within limits specified by the global model. The fitting error for this refined global model is compared with the original global model fitting error and a decision is taken to use the global model with the least error.

IV. System Overview and In-Camera Implementation

A. Blink Detection

Blinking is defined as the rapid closing and opening of the eyelid [15]. The average duration of an eye-blink is 0.3 to 0.4 seconds, with a frequency varying from once every two seconds up to several tens of seconds. The blinking rate can also be affected by external stimulus such as fatigue, eye injury, medication or disease.

The advantage of modeling the eyes utilizing AAM techniques is that we obtain a detailed description of the eye and not just its rough location. Using an educated choice of geometrical distances of the synthesized shape, a blink criterion can be defined (as described in Fig. 6), by comparing the ratio between the width of the eye and its aperture with a predefined threshold. The chosen distances were between the eleventh and the fifth point on the vertical and the eight and the second point on the horizontal. The annotation of the points is shown in Fig. 3. The method is accurate as long as the active appearance fitting algorithm provides the correct localization of the eye.

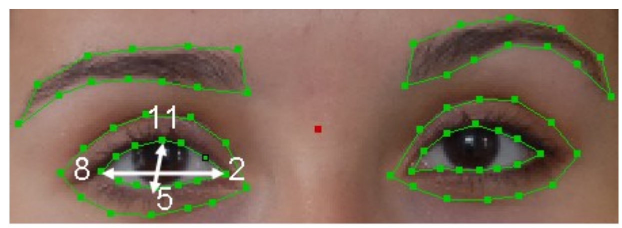

Fig. 6. The proposed method for blink detection.

We have developed alternatives to the proposed blink detector, as identifying if the blink parameters belong to the open-eyes curve (represented in green in Fig. 3) or to the closed-eyes one (black curve in Fig. 3) either for the first parameter of texture or for the first two parameters of shape. The drawback of this method is that it is designed to work for person-specific models dedicated to a single individual; also, the blink parameters have to be identified in advance.

B. Eye Tracking System Description

We propose a robust and accurate algorithm to track the eyes. The robustness refers to pose variations, eye actions, e.g. blinking, winking, gazing and occlusions.

The eye detection mechanism is inferred from a Viola-Jones face detection algorithm [11]. This involves using the statistical relation between the rectangle describing the surroundings of the face obtained after the detection step and the eye position in the rectangle learned from the training dataset. Thus, initial rough estimation of the eye position, 2D rotation angle and scale are deduced and are used as initial inputs for our eye-model.

As illustrated in Fig. 7 all the model parameters need to be optimized initially by the component-based AAM procedure described in section III.C. Those are the shape parameters including pose and blinking, texture parameters, as well as the parameters which characterize the modeled eye-patch inside the image frame. The purpose of this stage is to get an initial strong eye position from the first frame of the image sequence. This step is the most computationally expensive. Preferably it is performed on the first frame of a video sequence and is only periodically repeated. In our experiments this step is performed every 10 to 30 frames, but it depends on the application.

The main benefit of individual modeling of the eyes through the component-based AAM is that we can manipulate them independently. This is important in tracking applications where the subject is changing position from frame to frame. In some cases, one eye is partially or even completely occluded or the eyes lose their property of symmetry or co-linearity learned in the training stage from the global AAM model.

When one eye is occluded, the non-occluded eye, which can still be correctly fitted, provides sufficient information for the accurate positioning of both eyes using the global model. This confers robustness to the eye tracker.

Eye gazing must also be taken into consideration or the model will fail to differentiate between eyes with a wide gaze angle and closed eyes. We have found a simple solution, namely to introduce the iris into the model annotation, giving it the information on deformations which can result from a range of gaze angles during the training phase.

After parameters initialization, an update on each new image frame is made only on the model parameters that encode shape variations, position, 2D rotation and scaling. Since the computational requirements are substantially reduced, this operation is now performed on each separate frame.

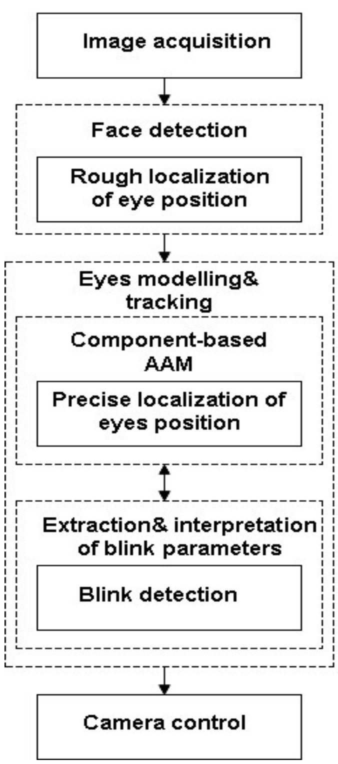

Fig. 7. System overview.

C. Eye Blink Detection for Consumer Imaging

As stated in the introduction, eye monitoring and analysis are key problems in a variety of consumer applications. One motivation for our current research is to investigate the feasibility of using AAM based eye tracking techniques in embedded imaging appliances. The reduced computational costs of the proposed eye tracking scheme, associated with the robustness of its output to blinking and gaze variations, could meet the requirements for real-time applications in such devices. As an example application we used our eye-model to analyze the aperture amplitude for the tracked eye. Within a digital camera this would determine if a subject had closed one of his/her eyes when a digital photo was captured.

Furthermore, our model would enable a replacement to be made of the subject's closed eye with an equivalent open eye. The degree of openness could be controlled using the model parameters and tuned to match that of the other tracked eye so that the substitution would be unnoticeable. Contrary, where both eyes are substantially closed, the degree of openness could be matched to the average degree of openness of the eyes over a sequence of video frames. A key point here is that once the local eye models are initialized most of the eye description will remain static and it is only necessary to adjust a handful of model parameters to adjust the openness of the eye or the direction of eye-gaze.

Our methodology was applied as follows, in order to replace the closed eyes: an AAM fitting was performed on both pictures, i.e. with open and closed eyes. This procedure enables segmenting the eyes in the two photos and retrieving their shape, pose and texture parameters. The original texture for the open eyes is transferred into the closed eyes photos, respecting the closed eyes pose parameters, i.e. translation, rotation, and scaling. The pictures including open eyes can be obtained from prior information in the camera, for instance by reviewing already existing photos.

Some initial experiments have been performed. The main requirement for this application is that both the shapes in the open-eyes picture and the closed-eyes picture are fitted with high accuracy. The results can be judged from Fig. 8.

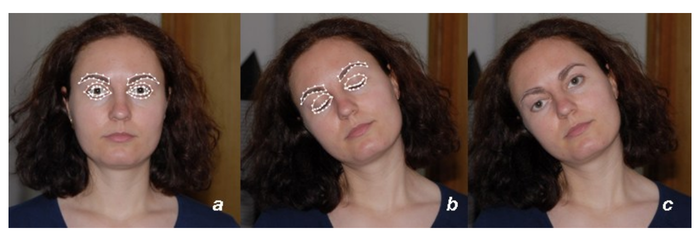

Fig. 8. The original pictures with a. open and b. closed eyes and c. the processed one, where closed eyes were replaced by open eyes.

V. Experimental Results and Challenges

Examples of the results obtained for different pictures and video sequences are presented in Figs. 9 and 11. For our initial training of the eye model only 44 annotated images have been used obtained from various subjects. From each person different states of the eyes were chosen, i.e. closed, open, half open, gazing and diverse head poses. The resulting model, constructed at a resolution of 30×10 pixels, presents 26 parameters: 21 shape-parameters, as 95% of the shape variance was allowed after the PCA, and 5 texture parameters, with 65% allowed texture variance. The choice of the resolution and the allowed variance is motivated by the minimal resolution for which the model stills performs successfully, from the point of view of correct blink detection and tracking accuracy, taking into account a minimal size for the model.

The model was evaluated using a set of 200 consumer pictures taken especially for this experiment where the subjects were asked to blink, gaze, rotate their head etc. In addition, two specialized databases, namely Georgia Tech Face Database [27] and VidTIMIT Video Dataset [28] were used to provide reference data sets. A number of 40 individuals have been tested from the Georgia Tech Face Database. The faces were captured at different dates, scales and orientations. The VidTIMIT Video Dataset is comprised of video recordings of different subjects reciting short sentences, but it is also useful for blinking or gazing detection. None of the pictures from the testing set have been used in the training set.

For our tests we have used an adaptation of the appearance modeling environment FAME [29], modifying and extending it to accommodate the techniques described in this paper.

A. Testing the Component-based AAM vs. the Classical AAM Algorithm for the Eye Model

The component-based AAM, an alternative to the classical AAM was adapted for the eye-region model and tested. Both the classical and the component-based AAM worked well for men, women or children, and for different skin and eye color. However, the latter model showed a better performance than the classical AAM in case of under-trained models or in more complex situations that deal with:

-*Pose:* e.g. for head rotations, the left eye will have a slightly different shape from the right one;

-*Occlusions:* Using the left/right eye sub-models, the eyes will fit independently and theoretically the non-occluded eye will present correct fitting, information that can be used further to fit the occluded eye.

-*Face expression:* One eye might change its shape independently of the other, e. g. winking.

-*Unseen details of the eye:* not seen in the training set, like unseen shapes or distances between the eyes, distances between the eyebrow and the eye.

-*Plan-rotation:* as represented in Fig. 9.

While the conventional AAM is based on a rough initialization inferred from the face detection step (Fig. 9b), the component-based uses the AAM fitting as starting point. Moreover, the fact that each eye is aligned independently (Fig. 9c) permits that each sub-model finds its optimum separately. The optimal result is obtained by constraining them within a two-eyes-together limit (Fig. 9d), which represents a new AAM alignment using as initialization the result from the independently fitting of the eyes.

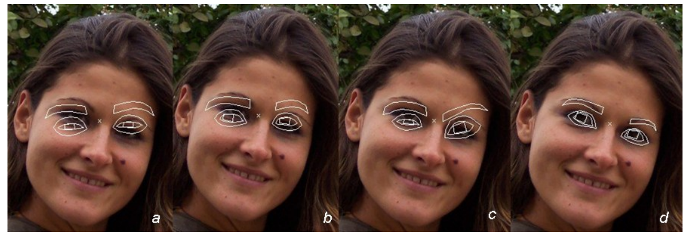

Fig. 9. From left to right: a. the initialization eye-shape inferred from face detection; b. the eye-shape from global AAM fitting; c. the eye shapes from AAM sub-model fitting; d. the fitted shapes of the component sub-models refitted by the global AAM – note how eye-symmetry has been restored in this case.

Tables I and II show the tested performance of the proposed system. Table I gives a summary of the performances obtained by our conventional statistical model of appearance in terms of eye tracking, gazing and blinking detection accuracy. Table II reports the improvements brought by the Component-based AAM for the enumerated actions. It is obvious that the component-based model improves overall accuracy for the system for all the databases used in the test set.

TABLE I
SUMMARY OF THE SYSTEM ACCURACY (%) FOR THE CONVENTIONAL MODEL OF APPEARANCE

Method	Georgia Tech Face Database	VidTIMIT Dataset	Our database
Eyes tracking	92	98	75
Eyes gazing	68.75	52.08	40.25
Blink detection	90.625	100	67.92

TABLE II
SUMMARY OF SYSTEM ACCURACY (%) FOR THE COMPONENT-BASED MODEL

Method	Georgia Tech Face Database	VidTIMIT Dataset	Our database
Eyes tracking	97.5	98	79.24
Eyes gazing	78.125	83.33	69.18
Blink detection	96.875	100	79.24

It is worth remarking that our database has been designed to test the model more rigorously as can be seen from a poorer level of performance. However these are initial results and we expect to further refine and optimize the model. Controlled conditions of illumination or pose, as usually found in databases alleviate the task of the algorithm, so an important aim is to make it robust to general conditions.

Fig. 10 presents the histogram of point-to-point (Pt-Pt) shape errors calculated with respect to the manual ground-truth annotations. The boundary errors for the tested fitting algorithms are concentrated within lower values as compared to the initial point generated from the detection algorithm, showing the methods improvement for eyes localization. Furthermore, it can be noticed that the shape boundary errors are concentrated within lowest values, indicating that the full component-based AAM performs the best in terms of fitting accuracy, thus resulting in a clear improvement over the initial position as well as over the other fitting algorithms.

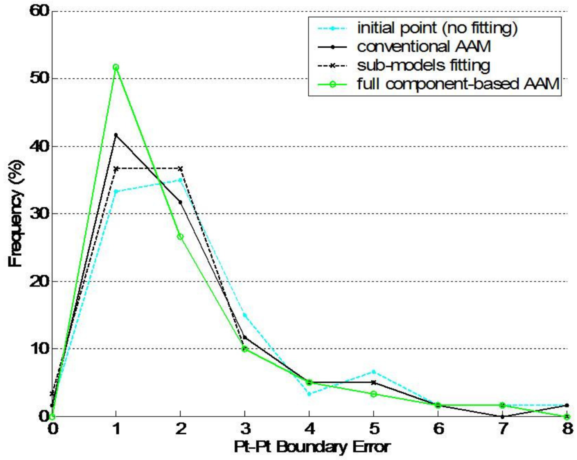

Fig. 10. The histogram of the boundary error for the three algorithms: conventional AAM, sub-model fitting and the component-based AAM.

B.Accuracy Testing of the Blink Detector and Eye Tracker

Various test video sequences, among which there are rotation of the head, blinking of the eyes and partial occlusions, were used to measure the performances during tracking. The tracking system was automatically initialized by the component-based AAM procedure and achieved an average accuracy of 91.5%.

The iris, indicating gazing movements, and eye contours, indicating the blinking or head movements, are tracked correctly by our method, depending on the complexity of the changes. The results proved to be encouraging (Fig. 11), namely the model was able to analyze correctly the eye, fact that infers an accurate tracking and a precise blink detection. A besetment of the model and its algorithm is the accuracy to gazing, where future improvements need to be done. Failures of the tracking algorithm are typically due to unfocused images, large face offsets between frames, partial occlusions and, in particular gazing issues.

VI. Conclusions & Future Work

In this article, we have described a system which uses statistical models of appearance to recover the eye blinking parameters of an eye model as well as tracking the location of the eyes in image sequences for various subjects with various expressions or poses.

Two eye states were considered in the same training set for the model, i.e. open and closed, instead of using two different models. The advantage is that the resulting model is able to synthesize all the states in between, facilitating the blink parameters extraction.

The results show that our method works well for the images taken under controlled environment. However, in more difficult situations like large variations of pose, plan-rotation, occlusions etc., the conventional statistical model of appearance performs poorly. The proposed solution was to extend the conventional statistical model of appearance to the component-based AAM by combining a global model (the two eyes together) with a series of sub-models (the two eyes separately). The method was applied successfully to robust eye tracking.

Fig. 11. Eye tracking in two example video sequences using our model, when blinking, face expression and small pose variations are involved.

References

[1] J. Orozco, J. Gonzàlez, P. Baiget, J. J. Villanueva, “Human Emotion Evaluation on Image Sequences ", *CogSys II Conference*, Radboud University Nijmegen, Netherlands, April 2006.

[2] S. P. Lee, J. B. Badler, N. I. Badler, “Eyes alive”, *ACM Trans. Graph*, vol. 21, no. 3, pp. 637-644, 2002.

[3] A.M. Tekalp, J. Ostermann, "Face and 2-D mesh animation in MPEG-4", *Signal Processing: Image Communication*, Vol. 15, pp. 387-421, 2000.

[4] T. Ishikawa, S. Baker, I. Matthews, T. Kanade, “Passive Driver Gaze Tracking with Active Appearance Models”, In Proc. World Congress on Intelligent Transportation Systems, February 2004.

[5] M. Betke, J. Gips, P. Fleming,” The Camera Mouse: Visual tracking of body features to provide computer access for people with severe disabilities”, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, March 2002.

[6] M. Chau, M. Betke, “Real Time Eye Tracking and Blink Detection with USB Cameras”, Boston University Computer Science, Technical Report No. 2005-12.

[7] Ssontech SynthEyes: http://www.ssontech.com/

[8] Q. Ji , X. Yang, “Real Time Visual Cues Extraction for Monitoring Driver Vigilance”, ICVS '01: *Proceedings of the Second International Workshop on Computer Vision Systems*, pp. 107-124, 2001.

[9] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active appearance models”, Lecture Notes in *Computer Science*, vol. 1407, pp. 484–, 1998

[10] D. W. Hansen, J. P. Hansen, M. Nielsen, A. S. Johansen, ” Eye Typing using Markov and Active Appearance Models”, WACV '02: *Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision*, pp. 132, 2002.

[11] P. Viola, M. J. Jones, “Robust Real-Time Face Detection”, *Springer Netherlands for Computer Science*, Volume 57, Number 2, May, 2004.

[12] C. Zhang and F. S. Cohen, “Component-based Active Appearance Models for face Modelling”, *Proceedings of the International Conference of Advances in Biometrics*, , ICB 2006, Hong Kong, China, January 5-7, 2006.

[13] M.G. Roberts, T.F. Cootes, J.E. Adams, “Linking Sequences of Active Appearance Sub-Models via Constraints: an Application in Automated Vertebral Morphometry”, *Proc.BMVC*, vol. 1, 2003, pp. 349--358.

[14] P. Corcoran, M. Ionita, I. Bacivarov, “Next Generation Face Tracking Technology Using AAM Techniques”, *Signals, Circuits and Systems, ISSCS* 2007, International Symposium, Iasi,13-14 July 2007, Romania, vol.1, pages 1-4, 2007.

[15] Robert A.Moses., “Adler's Physiology of the eye clinical application”, Ed. (Mosby, 1981), chap. 1, pp. 1-15.

[16] Z. Zhu, Q. Ji “Robust Real-Time Eye Detection and Tracking Under Variable Lighting Conditions and Various Face Orientations”, *Computer Visage and Image Understanding,* vol. 98, pp. 124-154, 2005.

[17] T. Moriyama, T. Kanade, T. Cohn, J.F. Jing Xiao; Z. Ambadar, Jiang Gao, H Imamura, “Automatic recognition of eye blinking in spontaneously occurring behaviour”, *Pattern Recognition*, 2002, Proceedings. 16th International Conference, Vol. 4, pp 78 – 81, 2002.

[18] D. O. Gorodnichy, “Second Order Change Detection, and its Application to Blink-Controlled Perceptual Interfaces”, *Proceedings*

of the International Association of Science and Technology for Development (IASTED) Benalmadena, Spain. September 8-10, 2003, pp. 140-145.

[19] S. Kawato and N. Tetsutani, “Detection and tracking of eyes for gaze-camera control”, *Proc. Intern. Conf. on Vision Interface* (VI’2002), Calgary, May 2002, pp. 348–355.

[20] K. Grauman, M. Betke, J. Gips, G. R. Bradski, “Communication via Eye Blinks- Detection and Duration Analysis in Real Time”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Lihue, HI, December 2001.

[21] D. W. Hansen, A. E.C. Pece, “Eye tracking in the wild”, *Computer Vision and Image Understanding* 98 (2005) pp. 155–181.

[22] K. Grauman, M. Betke, J. Gips, G. R. Bradski, “Communication via Eye Blinks- Detection and Duration Analysis in Real Time”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Lihue, HI, December 2001.

[23] A. Yuille, P. Hallinan, D. Cohen, “Feature extraction from faces using deformable templates”, *International Journal of Computer Vision* 8 (2) (1992) pp. 99-111.

[24] H. Tan; Yu-Jin Zhang; R. Li, “Robust eye extraction using deformable template and feature tracking ability”, Proceedings of the 2003 Joint Conference of the Fourth International Conference on Information, Communications and Signal Processing, Volume 3, Issue , 15-18 Dec. 2003.

[25] M. Pardas, “Extraction and tracking of the eyelids”, CASSP '00: *Proceedings of the Acoustics, Speech, and Signal Processing*, 2000, pp. 2357-2360.

[26] S. Kawato, J. Ohya, “Real-time detection of nodding and head-shaking by directly detecting and tracking the “between-eyes””, Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000 Volume , Issue , pp. 40 – 45.

[27] Georgia Tech Face Database, ftp://ftp.ee.gatech. edu/pub/users/hayes/facedb/.

[28] C. Sanderson and K.K. Paliwal, “Identity Verification Using Speech and Face Information”, *Digital Signal Processing* 14(5):449-480, 2004.

[29] M. B. Stegmann, B. K. Ersbøll, and R. Larsen, “FAME - a flexibleappearance modelling environment,” IEEE Transactions on Medical Imaging, vol. 22, no. 10, pp. 1319–1331, 2003. [Online]. Available: http://www2.imm.dtu.dk/pubdb/p.php?1918

[30] http://rfdesign.com/news/fotonation-facetracker-nokia-camera-0920/

[31] http://www.fotonation.com/index.php?module=product&item=23

Ioana Bacivarov received the M. Eng. Degree in Signal Processing from the National Polytechnic Institute of Grenoble, France and the University "Politehnica" of Bucharest, Romania in a double diploma agreement in 2006, 2007 respectively. She is currently pursuing a Ph.D. degree in Image Processing & Computer Vision at NUI, Galway. Her research interests include signal processing and pattern recognition with applications in face recognition.

Mircea C. Ionita received the B.Sc. degree in Computer Science & Electrical Engineering from University "Politehnica" of Bucharest, Romania, in 2003. He received the M.Sc. degree in Images, Patterns & Artificial Intelligence from the same university in 2004. He is currently pursuing a Ph.D. degree in Image Processing & Computer Vision at National University of Galway, Ireland. His research interests include color image processing, statistical face modeling, and face recognition and tracking.

Peter Corcoran received the BAI (Electronic Engineering) and BA (Math’s) degrees from Trinity College Dublin in 1984. He continued his studies at TCD and was awarded a Ph.D. for research work in the theory of Dielectric Liquids. In 1986 he was appointed to a lectureship in Electronic Engineering at NUI, Galway. His research interests include embedded systems, home networking, digital imaging and wireless networking technologies.