

# Automated Individualization of Deformable Eye Region Model and Its Application to Eye Motion Analysis

Tsuyoshi Moriyama  
Dept. of Media and Image Technology,  
Tokyo Polytechnic University  
1583 Iiyama, Atsugi,  
Kanagawa 243-0297, Japan  
moriyama@mega.t-kougei.ac.jp

Takeo Kanade  
Robotics Institute,  
Carnegie Mellon University  
5000 Forbes Avenue Pittsburgh,  
PA 15213-3891, USA  
tk@cs.cmu.edu

## Abstract

*This paper proposes a method of automated individualization of eye region model. The eye region model has been proposed in past research that parameterizes both the structure and the motion of the eye region. Without any prior knowledge, one can never determine a given appearance of eye region to be either neutral to any expression, i.e., the inherent structure of the eye region, or the result of motion by a facial expression. The past method manually individualized the model with respect to the structure parameters in the initial frame and tracks the motion parameters automatically across the rest of the image sequence, assuming the initial frame contains only neutral faces. Under the same assumption, we automatically determine the structure parameters for the given eye region image. We train Active Appearance Models (AAMs) for parameterizing the variance of individuality. The system projects a given eye region image onto the low dimensional subspace spanned by the AAM and retrieves the structure parameters of the nearest training sample and initializes the eye region model using them. The AAMs are trained in the subregions, i.e., the upper eyelid region, the palpebral fissure (the eye aperture) region, and the lower eyelid region, respectively. It enables each AAM to effectively represent fine structures. Experimental results show the proposed method gives as nice initialization as manual labor and allows comparative tracking results for a comprehensive set of eye motions.*

## 1. Introduction

Accurate analysis and anatomical interpretation of eye motions are useful for medical purposes, biometrics, and psychological studies. Image sensing techniques for them are inherently sensitive to eye's pose to the camera and illumination changes and facial expressions. Effects due to fa-

cial expressions are particularly difficult to remove because neutral face can look one with expression and vice versa. One needs either to assume a face given is neutral or to define "neutral" based on prior knowledge.

Moriyama *et al.* [7] proposed a meticulously detailed eye region model that parameterizes the individuality of the eye region by the structure parameters and the motion of the eye by the motion parameters. They assume the initial frame of the input image sequence contains "neutral" face and manually initialize the model with respect to the structure parameters in the frame and automatically estimates the motion parameters using a gradient descent algorithm for the rest of the image sequence.

Active Appearance Model (AAM) [2] is powerful in representing deformable image objects such as faces and organs in the low dimensional subspace and capable of constraining the appearance of a new sample generated by the AAM "within" those of the set of training data. This paper extends Moriyama's work using AAM so that the system can automate initialization of the eye region model by automatically determining the structure parameters in the initial frame. The system firstly represents the input eye region by AAM and secondly selects the nearest training sample to the input and thirdly initializes the structure parameters of the input with those of the selected training sample. Dividing eye region into subregions assures AAM to be able to represent detailed structure of the eye region that varies among individuals. The system finally tracks the motion of the eye such as blinking across the entire image sequence.

## 2. Eye Region Image and Its Individuality

An eye region divides into three subregions that include the upper eyelid subregion, the lower eyelid subregion, and the *palpebral fissure* subregion (the aperture between the eyelids).

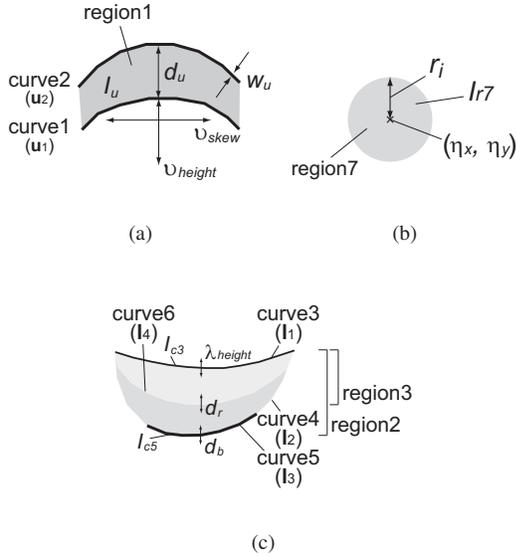


Figure 1. Model for each part in the eye region proposed by Moriyama *et al.* [7]. (a) Model for the upper eyelid. (b) Model for the iris. (c) Model for the lower eyelid.

## 2.1. Upper Eyelid Subregion

The upper eyelid subregion is a skin region above the *palpebral fissure* and consists of the *orbicularis oculi* muscle covered by fat [4][6][1]. The *orbicularis oculi* muscle divides into three parts (from the top, the *orbital orbicularis*, the *preseptal orbicularis*, and the *pretarsal orbicularis*) and the *pretarsal orbicularis* folds at the boundary between the upper and the lower parts to make a crease (the *sulcus suprapalpebralis*) along the boundary when the eye opens.

When the position of the crease is low, the eye looks a single eyelid fold that is common in East Asians (or one that goes to double eyelid fold when it closes that 50% of East Asians have), when high, it looks a double eyelid fold. The depth of the crease is determined by both the strength of connection between the *superior tarsal* plate and the skin above it (the lower part of the upper eyelid) and the amount of the skin and the fat above the connection (the upper part of the upper eyelid). When the connection is strong and the amount of the skin and the fat is much, the skin of the lower part of the upper eyelid, the *orbicularis oculi* muscle, and the *superior tarsal* plate are all lifted higher as a whole and more of the skin above it sticks out to make a deeper crease. The individuality of the upper eyelid subregion thus depends on the structure of both the position and the depth of the crease.

The eye region model proposed by Moriyama *et al.* shown in Fig.1(a) has two polygonal curves for the lower end of the upper eyelid and parameterizes the individuality

by the distance between those curves ( $d_u$ ) and the boldness of the crease ( $f$ ; the thickness of the crease and the darkness of the region surrounded by those curves). They determined those parameters so as to make the best accuracy of tracking by the eye region model but they are also anatomically sufficient to represent the individual difference.

## 2.2. Lower Eyelid Subregion

The lower eyelid subregion has a hammock-like structure where the *inferior tarsal* plate located along the upper edge of the lower eyelid supported by both ends of the eye (the medial *canthal* tendon and the lateral *canthal* tendon) lifts the skin that covers the region below it. There occurs a bulge resulting from the shape of the covered portion of the eye, shortening of the inferior portion of the *orbicularis oculi* muscle on its length, and the effects of gravity and aging. The bulge creates the *infraorbital furrow* (a diagonal furrow above the cheek) below the skin region lifted by the hammock. The individuality of the lower eyelid subregion thus depends on both the width and the depth of the bulge.

The eye region model proposed by Moriyama *et al.* shown in Fig.1(c) has two polygonal curves for the upper end of the lower eyelid and parameterizes the individuality by the distance between those curves by  $d_b$ . They also considered the environmental illumination that may brighten the lower eyelid. Their model has a polygonal curve for modeling the lower edge of the illuminated region and parameterizes the distance to the upper edge of the lower eyelid by  $d_r$ .

## 2.3. Palpebral Fissure Subregion

The *palpebral fissure* subregion is the aperture surrounded by both the upper and the lower eyelid subregions and consists of the region of the iris and that of the *sclera*. The inner corner of the *sclera* has a reddish organ that provides tear. The size of the iris differs over individuals. The center of the iris has the pupil that is a black cavity that changes its size depending on the amount of light that reaches the *retina*. The color of the iris changes from blue to brown depending on the amount of pigment. When minimum (the iris to appear blue), the texture inside the iris is more visible and the iris reflects the reverse image of the environment on the *cornea* and the lens. The individuality of the *palpebral fissure* subregion thus depends on both the size of the iris and the texture inside of the iris.

The eye region model proposed by Moriyama *et al.* shown in Fig.1(b) has a disk with a single color for modeling the iris and parameterizes the individuality by the size ( $r_i$ ; the radius) and the brightness ( $I_{r7}$ ) of the disk. They did not consider modeling the texture inside the iris but simply represented it by a mean brightness within the region and achieved accurate tracking of the iris.

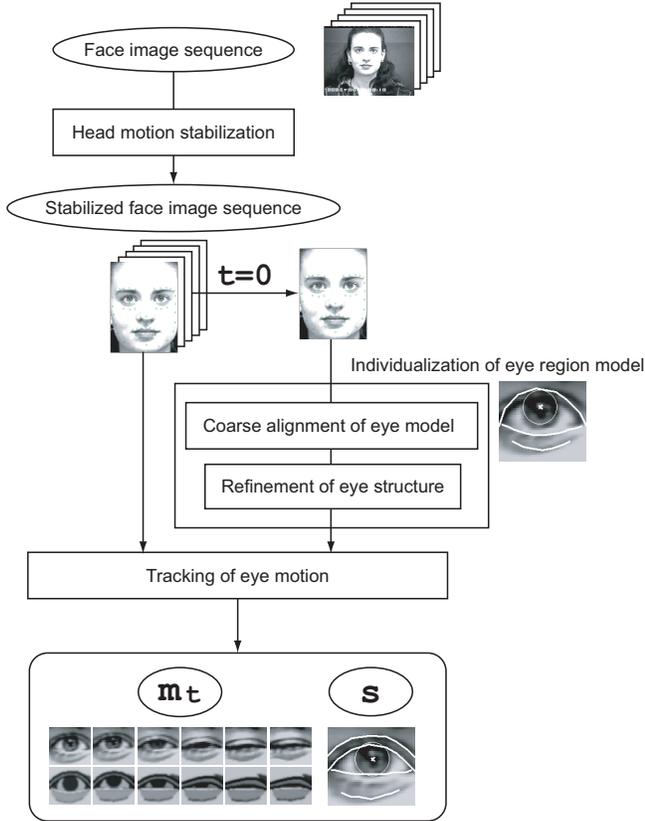


Figure 2. Model-based eye image analysis system.

### 3. Model-Based Eye Image Analysis

Parameters  $d_u$ ,  $f$ ,  $d_b$ ,  $d_r$ ,  $r_i$ , and  $I_{r7}$  are called the structure parameters (denoted by  $\mathbf{s}$ ) and parameterize the static and structural detail in the eye region model. The dynamic detail of the model is parameterized by the motion parameters (denoted by  $\mathbf{m}_t$ ,  $t$ : time) that include the motion of the upper eyelid ( $\nu_{height}$  and  $\nu_{skew}$ ), that of the lower eyelid ( $\lambda_{height}$ ), and that of the iris ( $\eta_x$  and  $\eta_y$ ).

Fig. 2 shows an overview of the model-based eye image analysis system that uses the eye region model [7]. The system tracks head motion across the input image sequence and warps the faces to a canonical head pose (upright and frontal), that they refer to as the stabilized images. In a particular frame of the stabilized image sequence, the eye region model is aligned to the input eye region and individualized by adjusting the structure parameters  $\mathbf{s}$ . The system automatically tracks the motion parameters  $\mathbf{m}_t$  frame by frame starting from the initial frame using a gradient descent algorithm. Individualization of the eye region model was manually done in their report.

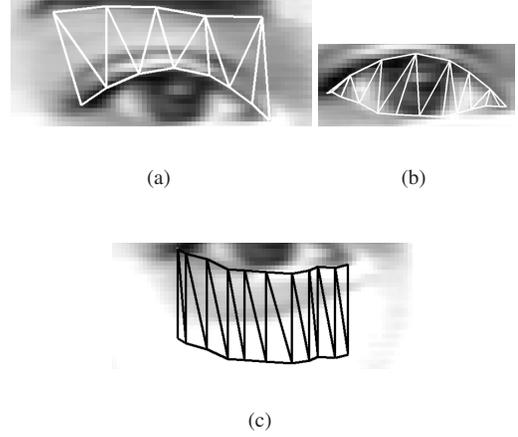


Figure 3. Subregions in an example eye region image and their triangulation. (a) The upper eyelid region has 5 points on the eyebrow and 8 on the upper eyelid. (b) The *palpebral fissure* region has 8 points on the upper eyelid and 11 on the lower eyelid. (c) The lower eyelid region has 11 points on the lower eyelid and the same number parallel to and at half the width of the eye region below it.

## 4. Automated Individualization of Eye Region Model

The proposed method automates individualization of the structure parameters  $\mathbf{s}$  by using AAM (Active Appearance Models) [3]. It defines three subregions in the eye region image, as shown in Fig.3. Manually labeled contour points of each subregion for all the training samples are used for generating AAM. AAM represents both the shape and the texture of each subregion of an eye region image in the low dimensional subspace by using only the significant principal components.

For each subregion of the input eye region image, the system first searches an optimal AAM that represents the appearance of the input in the subspace while leaving the positions of the corner points of both the eye and the eyebrow fixed (constrained) [3]. It then finds the nearest training sample to the input in the subspace and initializes the structure parameters of the input with those of the selected training sample.

### 4.1. Training AAM for Subregions

Denoting the coordinates of manually labeled contour points by  $\mathbf{x}$ , a new shape is represented by the linear combination of the mean shape  $\bar{\mathbf{x}}$  and the variance from it using principal component analysis on training samples as (1). All the textures warped to a canonical shape (*e.g.* mean shape)

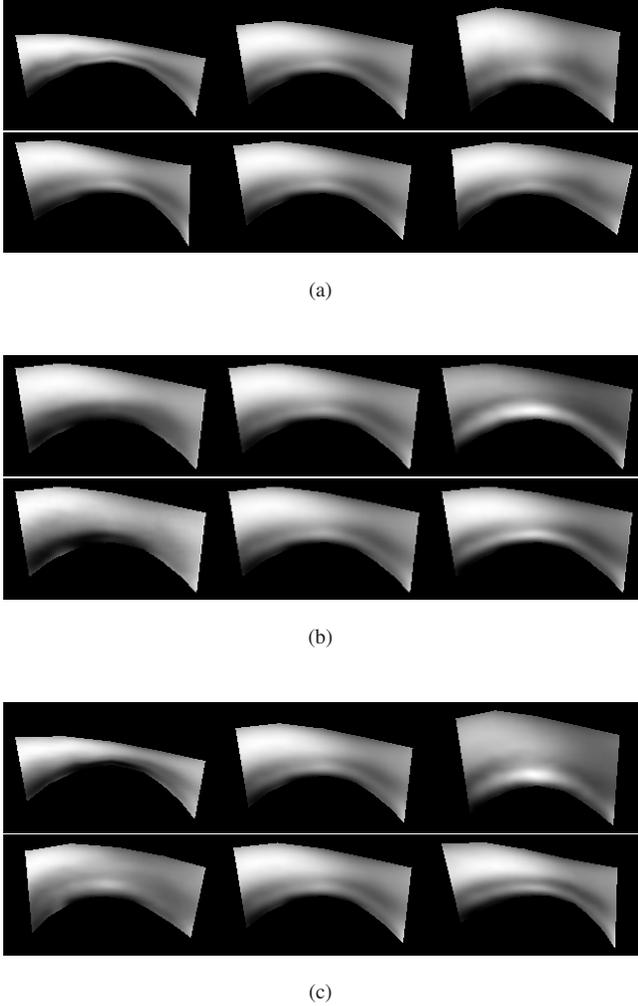


Figure 4. Modes for shape, texture and appearance for the upper eyelid subregion. AAM when systematically changing each mode from  $-3\sigma$  to  $+3\sigma$  ( $\sigma$ : standard deviation). (a) shape parameters. (b) texture parameters. (c) appearance parameters.

are similarly represented as (2):

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}_s \mathbf{b}_s \quad (1)$$

$$\mathbf{g} = \bar{\mathbf{g}} + \mathbf{P}_g \mathbf{b}_g, \quad (2)$$

where  $\mathbf{b}_s$  and  $\mathbf{b}_g$  are the shape parameters and the texture parameters that give the amount of deviation from the mean, respectively. After normalizing the dimensionality by  $\mathbf{W}_s$ , both shape and texture further reduce the dimensionality by

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_g \end{pmatrix} = \begin{pmatrix} \mathbf{W}_s \mathbf{P}_s^T (\mathbf{x} - \bar{\mathbf{x}}) \\ \mathbf{P}_g^T (\mathbf{g} - \bar{\mathbf{g}}) \end{pmatrix} = \mathbf{Q} \mathbf{c} \quad (3)$$

where  $\mathbf{Q}$  is the eigenvectors and  $\mathbf{c}$  controls both the shape and the texture of the eye region image simultaneously.

Table 1. The number of dimensions that cover over 90% of the total variance for each mode.

| subregions               | shape | texture | appearance |
|--------------------------|-------|---------|------------|
| upper eyelid             | 5     | 14      | 19         |
| <i>palpebral fissure</i> | 3     | 20      | 23         |
| lower eyelid             | 7     | 16      | 23         |

Fig.4 shows how AAM changes its appearance when systematically manipulating the shape, the texture, and the appearance parameters, respectively. AAMs for the lower eyelid subregion and the *palpebral fissure* subregion are also similarly computed.

## 5. Experiments

The proposed method is tested on Cohn-Kanade AU-Coded Facial Expression Database [5] that contains 490 image sequences of facial behaviors from 102 subjects, all but one of which were from the publicly released subset of the database. The subjects are adults that range from 18 to 50 years old with both genders (66 females and 35 males) and a variety of ethnicities (86 Caucasians, 12 African Americans, 1 East Asian, and two from other groups).

Here used leave-one-out cross validation where all the subjects but the input's were used for training AAM. Automatically selected training sample that is nearest in the AAM subspace gives the structure parameters of the input. Evaluation of the proposed method is done by comparing the tracking results between manual and automated initialization of the eye region.

### 5.1. Trained AAMs

The same image frames that were used for manual initialization were collected (one for each subject) as training samples. The number of dimensions that cover over 90% of the total variance for each mode is shown in Table 1.

### 5.2. Results of Automated Initialization

Example results of automated initialization of the eye region model are shown in Table 2. The third column of the table shows the training samples that were found nearest in the subspace of AAM (from the top, the upper eyelid region, the *palpebral fissure* region, and the lower eyelid region). Those that have similar appearances were automatically selected for the subregions in each example except Table 2(d) where the training sample with double eyelid fold was selected for the input with single eyelid fold. This occurred because there was no other example of single eyelid fold in the training samples than the input (The database used contains only 1 East Asian with single eyelid fold.). Including more subjects with single eyelid fold would solve this problem.

Table 2. Example Results of Automated Structure Initialization. (a) Revealing double-fold, dark iris, and no bulge below the eye. (b) Double-fold, dark iris, and no bulge below the eye. (c) Double-fold, bright iris, and bulge below the eye. (d) Single fold, dark iris, bulge and reflection on it below the eye.

|     | input eye region images | selected training samples | individualized structure parameters   |  |
|-----|-------------------------|---------------------------|---|--|
| (a) |                         |                           | $d_u = 0.3$<br>$f = 1.0$<br>$I_{r7} = 86$<br>$r_i = 10$<br>$d_b = 0.0$<br>$d_r = 0.0$   |  |
| (b) |                         |                           | $d_u = 0.5$<br>$f = 1.0$<br>$I_{r7} = 50$<br>$r_i = 10$<br>$d_b = 0.68$<br>$d_r = 0.0$  |  |
| (c) |                         |                           | $d_u = 0.3$<br>$f = 0.7$<br>$I_{r7} = 50$<br>$r_i = 9$<br>$d_b = 0.05$<br>$d_r = 0.48$  |  |
| (d) |                         |                           | $d_u = 0.6$<br>$f = 0.75$<br>$I_{r7} = 50$<br>$r_i = 9$<br>$d_b = 0.56$<br>$d_r = 0.15$ |  |

### 5.3. Tracking Results by Automated Initialization

Tracking results for these examples are shown in Table 3. In 3(a) and (b) performance for automated and manual initialization were comparable. Tracking in 3(c) was unstable because the upper eyelid became brighter in the last frame and illumination reflection on the upper eyelid was not parameterized in the model. In 3(d) tracking error occurred in the last frame because the model for the double eyelid fold matched with the single eyelid fold in the input. Other examples of tracking with automated initialization for a comprehensive set of eye motions are shown in Table 4. Automated initialization again led to good tracking. This means that neutral eyes with similar appearances similarly change the appearance in their motions, which suggests that

Table 3. Example Comparison of Tracking Results between Manual and Automated Individualization

|     | first frame | last frame | manual | automated |
|-----|-------------|------------|--------|-----------|
| (a) |             |            |        |           |
| (b) |             |            |        |           |
| (c) |             |            |        |           |
| (d) |             |            |        |           |

the appearance changes caused from motion can be predicted from the structure in the neutral expression.

## 6. Conclusion

This paper extended the model-based tracking of eye motions that has originally been proposed by Moriyama *et al.* [7] so that the structural individuality of the input eye region is automatically parameterized prior to tracking while the original implementation has manually done it. Dividing the eye region into three subregions allowed AAM to effectively reduce the dimensionality in recovering the shape and the texture of the eye region image retaining the individual differences. The proposed method also constrained initialization of the structure parameters of the eye region model within the range possible to neutral eyes. The experimental results showed automated individualization achieved as accurate tracking of eye motions comparably with the manual setting both for diverse set of individual differences and for comprehensive set of motions spontaneously occurring in the eye region.

## Acknowledgement

This research was supported by grants R01 MH51435 from the National Institute of Mental Health, U.S. and the Special Coordination Funds of the Ministry of Education, Culture, Sports, Science and Technology, Japan. Special thanks to Jeffrey F. Cohn and Jing Xiao for their help and support.

## References

- [1] S. R. Carter. Eyelid disorders: Diagnosis and management. <http://www.aafp.org/>

Table 4. Example Results of Tracking Eye Motions with Automated Structure Individualization

|   | initial frame | selected training samples | individualized structure parameters   | tracking results |           |
|---|---------------|---------------------------|---|------------------|-----------|
|   |               |                           |   | manual           | automated |
| (a) Upper eyelid raising                |               |                           | $d_u = 0.5$<br>$f = 0.7$<br>$I_{r7} = 50$<br>$r_i = 10$<br>$d_b = 0.0$<br>$d_r = 0.0$   |                  |           |
| (b) Gaze change and cheek raising       |               |                           | $d_u = 0.6$<br>$f = 0.7$<br>$I_{r7} = 95$<br>$r_i = 10$<br>$d_b = 0.54$<br>$d_r = 0.0$  |                  |           |
| (c) Cheek raising                       |               |                           | $d_u = 0.4$<br>$f = 0.7$<br>$I_{r7} = 50$<br>$r_i = 10$<br>$d_b = 0.1$<br>$d_r = 0.66$  |                  |           |
| (d) Blinking and eyelid tightening      |               |                           | $d_u = 0.7$<br>$f = 0.8$<br>$I_{r7} = 77$<br>$r_i = 10$<br>$d_b = 0.05$<br>$d_r = 0.09$ |                  |           |
| (e) Cheek raising and eyelid tightening |               |                           | $d_u = 0.3$<br>$f = 1.0$<br>$I_{r7} = 50$<br>$r_i = 10$<br>$d_b = 0.05$<br>$d_r = 0.48$ |                  |           |

[afp/980600ap/carter.html](http://afp/980600ap/carter.html).

- [2] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In H. . B. Neumann, editor, *Proc. of the 5th European Conference on Computer Vision*, volume 2, pages 484–498. Springer, 1998.
- [3] T. F. Cootes and C. J. Taylor. Constrained active appearance models. In *International Conference on Computer Vision 2001*, volume I, pages 748–754, Vancouver, Canada, July 2001.
- [4] EyePlastics.com. Blepharoplasty - anatomy. [http://www.eyeplastics.com/topics/blepharoplasty/blepharoplasty\\_anatomy.htm](http://www.eyeplastics.com/topics/blepharoplasty/blepharoplasty_anatomy.htm).
- [5] T. Kanade, J. F. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. In *Proc. IEEE Face and Gesture '00*, pages 46–53, 2000.
- [6] F. Meronk. Chapter 4: Eyelid anatomy for those who hated high school biology. <http://www.drmeronk.com/insidersguide/04.html>.
- [7] T. Moriyama, T. Kanade, J. Xiao, and J. F. Cohn. Meticulously detailed eye region model and its application to analysis of facial images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5):738–752, 2006.