Deformable Models for Eye Tracking

Martin Vester-Christensen, Denis Leimberg, Bjarne Kjær Ersbøll and Lars Kai Hansen
Department of Informatics and Mathematical Modelling
Technical University of Denmark
mvc@imm.dtu.dk, denis@kultvizion.dk, be@imm.dtu.dk, lkh@imm.dtu.dk

Abstract

A deformable template method for eye tracking on full face images is presented. The strengths of the method are that it is fast and retains accuracy independently of the resolution. We compare the method with a state of the art active contour approach, showing that the heuristic method is more accurate.

1 Introduction

Eye Tracking is the process of finding and tracking the eye of a human in a sequence of images. Specifically finding and tracking the iris or pupil can be used to infer the direction of interest of the human subject, this is denoted *gaze*.

Gaze is very important for human communication and also plays an increasing role for human computer interaction. Gaze can play a role, e.g., in understanding the emotional state for humans [1, 2], synthesizing emotions [5], and for estimation of attentional state [16]. Specific applications include devices for the disabled, e.g., using gaze as a replacement for a computer mouse and driver awareness monitoring to improve traffic safety [8].

It has been noted that the high cost of good gaze detection devices is a major road block for broader application of gaze technology, hence, there is a strong motivation for creating systems that are simple, inexpensive, and robust [7].

Eye tracking is an active area of research. COGAIN is a network of excellence on Communication by Gaze Interaction, supported by the European Commission's IST 6th framework program. COGAIN integrates cutting-edge expertise on interface technologies for the benefit of users with disabilities. The network aims to gather Europe's leading expertise in eye tracking integration with computers in a research project on assistive technologies for citizens with motor impairments[3]. The authors of this paper are members of this network, and it summarizes research presented in[11].

Figure 1: Examples of the dataset. The region surrounding the eyes can be found in various ways. We use a head tracking algorithm[8] based on Active Appearance Models. A subimage is extracted and subsequently processed by the eye tracking algorithms.

The paper is organized as follows. First a brief review of some of the methods used for eye tracking is given in section 2. Section 3 describes the proposed deformable template method. Section 4 describes the EM-contour method from [7] with additional constraints on the model. The two models are compared in section 5. Finally some concluding remarks are drawn in section 6.

2 Recent Work

Detection of the human eye is a difficult task due to a weak contrast between the eye and the surrounding skin. As a consequence, many existing approaches uses close-up cameras to obtain high-resolution images[7][19]. However, this imposes restrictions on head movements. The problem can be overcome by use of a two camera

setup[18][20]. One camera covering the head and controlling a second camera, which focuses on one eye of the person. Matsumoto and Zelinsky[12] utilizes template and stereo matching.

In many existing approaches the shape of iris is modeled as a circle [9][10][12][20]. Since the shape and texture of the object is known, a template model can be used with advantage[8][15]. J. Gracht et al.[17] utilizes an iris template generated by a series of wavelet filtering.

Wang et al.[18] detects the iris using thresholding, morphology and vertical edge operators. An ellipse is fitted to the resulting binary image.

A probabilistic formulation of eye trackers has the attraction that uncertainty is handled in a systematic fashion. Xie et al.[20] utilizes a Kalman filter with purpose to track the eyes. The eye region is detected by thresholding and the center of an eye is used for motion compensation. The center of this iris is chosen as tracking parameter, while the gray level of the circle modeled eye is chosen as measurement[21]. Hansen and Pece propose an active contour model combining local edges along the contour of the iris[7]. The contour model is utilized by a particle filter.

A generative model explaining the variance of the appearance of the eye is developed by Moriyama et al.[13]. The system defines the structures and motions of the eye. The structure represents information regarding size and color of iris, width and boldness of eyelid etc. The motion is represented by the position of upper and lower eyelids and 2D position of the iris. Witzner et al. utilizes an Active Appearance Model[6].

Based on the center of iris estimate, the gaze direction can be computed utilizing various methods. Stiefelhagen et al.[15] utilizes a neural network with the eye image as input. Witzner et al.[6] uses a Gaussian process interpolation method for inferring the mapping from image coordinates to screen coordinates. Ishikawa et al. [8] exploits a geometric head model, which translates from 2D image coordinates to a direction in space relative to the initial frame.

The present paper is inspired by the line of thinking mentioned above. We focus on some of the image processing issues. In particular we propose a robust algorithm for swift eye tracking in low-resolution video images. We compare this algorithm with a proven method[7] and relate the pixel-wise error to the precision of the gaze determination.

3 Deformable Template Matching

Modeling the iris as a circle is well-motivated when the camera pose coincides with the optical axis of the eye. When the gaze is off the optical axis, the circular iris is rotated in 3D space, and appears as an ellipse in the image plane. Thus, the shape of the contour changes as a function of the gaze direction and the camera pose. The objective is then to fit an ellipse to the pupil contour, which is characterized by a darker color compared to the iris. The ellipse is parameterized,

$$\mathbf{x} = (c_x, c_y, \lambda_1, \lambda_2, \theta)\,, \tag{1}$$

where (c_x, c_y) is the ellipse centroid, λ_1 and λ_2 are the lengths of the major and minor axis respectively. θ is the orientation of the ellipse.

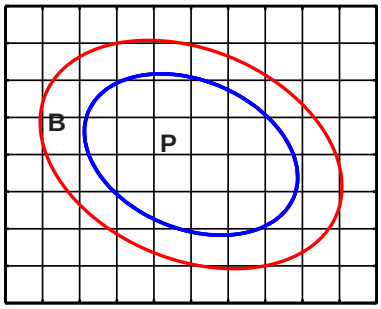

Figure 2: The deformable template model. Region P is the inner circle, and region B is the ring around it.

The model proposed here is based on the relationship between pixel values in two regions, see figure 2. The pupil region P is the part of the image I spanned by the ellipse parameterized by $\mathbf{x}$. The background region B is defined as the pixels inside an ellipse, surrounding but not included in P, as seen in figure 2. When region P contains the entire object, B must be outside the object, and thus the difference in average pixel intensity is maximal. To ensure equal weighting of the two regions, they have the same area. The area of the inner ellipse P is $A_P = \pi\lambda_1\lambda_2$. The shape parameters of B should satisfy the constraint on the area $A_{B/P} - A_P = A_P$. As a consequence, the param-

eters is defined as $\mathbf{x}_B = \left(c_x, c_y, \sqrt{2}\lambda_1, \sqrt{2}\lambda_2, \theta\right)$, while $\mathbf{x}_P$ is defined as (1).

The pupil contour can now be estimated by minimizing the cost function,

$$\mathcal{E} = \mathrm{Av}(P) - \mathrm{Av}(B), \tag{2}$$

where $\mathrm{Av}(B)$ and $\mathrm{Av}(P)$ are the average pixel intensities of the background - in this case the iris - and pupil region respectively.

The model is deformed by Newton optimization given an appropriate starting point. Due to rapid eye movements[14], the algorithm may break down if one uses the previous state as initial guess of the current state, since the starting point may be too far from the true state. As a consequence, we use a simple 'double threshold' estimate of the pupil region as starting point.

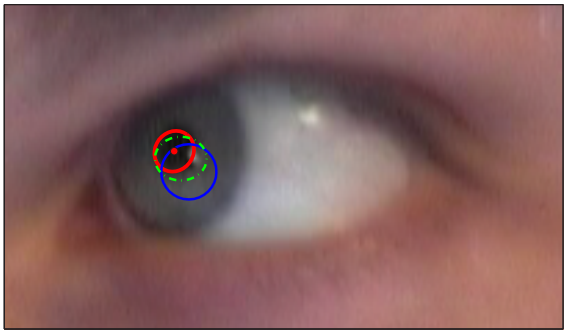

Figure 3: The blue ellipse indicates the starting point of the pupil contour. The template is iteratively deformed by an optimizer; one of the iterations is depicted in green. The red ellipse indicates the resulting estimate of the contour.

An example of the optimization of the deformable model is seen in figure 3.

3.1 Constraining the Deformation

Although a deformable template model is capable of catching changes in the pupil shape, there are also some major drawbacks. Corneal reflections, caused by illumination, may confuse the algorithm and cause it to deform unnaturally. In the worst case, the shape may grow or shrink until the algorithm collapses.

We propose to constrain the deformation of the model in the optimization step by adding a regularization term. Assume the parameters defining an ellipse is normally distributed with mean μ and covariance $\mathbf{\Sigma}$. The prior distribution of these parameters are then defined,

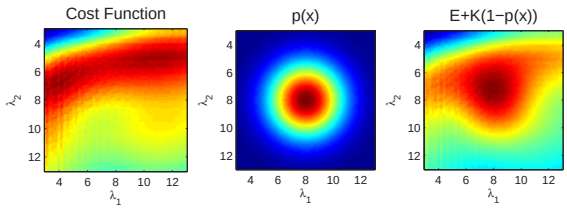

Figure 4: Given an appropriate starting point $\mathbf{x}$. The pose and orientation are kept fixed, while the shape parameters are varied. Note that the surface plots are not - as expected - smooth. This is due to rounding in the interpolation when evaluating the image evidence of the deformable template. *(Left)* The image confidence given the state - warmer colors means more likely. *(Middle)* The prior probability is a normal distribution with a given mean value μ and covariance Σ. *(Right)* Combining the image evidence and prior according to (4) yields the constrained estimate.

$$p(\mathbf{x}) = \mathcal{N}(\mu, \mathbf{\Sigma}) \propto \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right), \tag{3}$$

where the normalization factor has been omitted. The mean and covariance are estimated in a training sequence. At last the optimization of the deformable template matching method is constrained by adding a regularization term,

$$\mathcal{E} = \mathrm{Av}(P) - \mathrm{Av}(B) + \mathcal{K}\left(1 - p(\mathbf{x})\right), \tag{4}$$

where $\mathcal{K}$ is the gain of the regularization term.

The relevance of constraining the deformation is visualized in figure 4. A suitable starting point $\mathbf{x}$ is chosen. The pose and orientation are kept fixed, while the shape parameters are varied. In this case the true shape parameters λ_1 and λ_2 are approximately eight. The image confidence as a function of the shape parameters is depicted to the left, while the prior distribution is seen in the middle of figure 4. Combining the image confidence with a prior according to (4) yields the constrained estimate, which is depicted to the right in figure 4.

By use of the shape constraints, we incorporate prior knowledge to the solution. The robustness is increased considerably and the parameters are constrained to avoid the algorithm to break down due to infinite increase or decrease of parameters.

The deformable template matching method is seen applied with and without constraints in figure 5. The constrained estimate is seen to be less sensitive to noise due to reflections.

Figure 5: The deformable template matching method applied without constraints is seen in green, while the red ellipse depicts the constrained version . The constrained estimate is seen to be less sensitive to noise due to reflections.

4 EM Contour Tracking

The iris is circular and characterized by a large contrast to the sclera. Therefore, it seems obvious to use a contour based tracker. Witzner et al.[7] describe an algorithm for tracking using active contours and particle filtering. A generative model is formulated which combines a dynamic model of state propagation and an observation model relating the contours to the image data. The current state is then found recursively by taking the sample mean of the estimated posterior probability.

The proposed method in this paper is based on [7], but extended with constraints and robust statistics.

4.1 The Dynamic Model

The dynamic model describes how the iris moves from frame to frame. Again, the iris is modeled as an ellipse and the state vector $\mathbf{x}$ consist of the five parameters defining an ellipse as defined in equation 1.

To define the problem of tracking, consider the evolution of the state sequence

$$\mathbf{x}_{t+1} = \mathbf{f}_{t+1}\{\mathbf{x}_t, t \in \mathbb{N}\}, \tag{5}$$

of a target, given by

$$\mathbf{x}_{t+1} = \mathbf{f}_{t+1}(\mathbf{x}_t, \mathbf{v}_t), \tag{6}$$

where $\mathbf{f}_{t+1}$ is a possibly non-linear function of the state $\mathbf{x}_t$ and $\{\mathbf{v}_t, t \in \mathbb{N}\}$ is an independent identically distributed process noise sequence. The objective of tracking is to recursively estimate $\mathbf{x}_{t+1}$ from the measurements,

$$\mathcal{M}_{t+1} = \mathbf{h}_{t+1}(\mathbf{x}_{t+1}, \mathbf{n}_{t+1}), \tag{7}$$

where $\mathbf{h}_{t+1}$ is a possibly non-linear function and $\{\mathbf{n}_{t+1}, t \in \mathbb{N}\}$ is an i.i.d measurement noise sequence.

The pupil movements can be very rapid and is therefore modeled as Brownian motions(AR(1)). Thus the evolution of the state sequence (6) is modeled,

$$\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{v}_t, \quad \mathbf{v}_t \sim \mathcal{N}(0, \mathbf{\Sigma}_t), \tag{8}$$

where $\mathbf{\Sigma}_t$ is the time dependent covariance matrix of the noise. The time dependency compensates for scale changes, which affects the amount of movement. Larger movements is expected when the ellipse appears large, since the position of the eye is nearer to the camera. Contrary, when the eye is farther from the camera, smaller movements are expected. Hence, the first two diagonal elements of $\mathbf{\Sigma}_t$ corresponding to c_x and c_y are assumed to be linear dependent on previous sample mean.

4.2 The Observation Model

The observation model consists of two parts; a geometric component defining a probability density function over image locations of contours and a texture component defining a pdf over pixel gray level differences given a contour location. The geometric component models the deformations of the iris by assuming Gaussian distribution of all sample points along the contour. The gray level information is gathered by sampling a discrete set of points along the normals of all contour sampling points. Both components are joined and marginalized to produce a test of the hypothesis that there is a true contour present. The contour maximizing the combined hypotheses is chosen, see [7] for details.

4.3 Active Contour Tracking

The probabilistic formulation has the attraction that uncertainty is handled in a systematic fashion - Increased uncertainty results the particles to be drawn from a wider distribution, while increased confidence results the particles to be drawn from a narrower distribution.

The prediction stage involves using the system model (6) to obtain the prior pdf of the state at time $t+1$,

$$p(\mathbf{x}_{t+1}|\mathcal{M}_t) = \int p(\mathbf{x}_{t+1}|\mathbf{x}_t)p(\mathbf{x}_t|\mathcal{M}_t)\mathrm{d}\mathbf{x}_t \quad (9)$$

The observation $\mathcal{M}_t$ is independent of the previous state $\mathbf{x}_{t-1}$ and previous observation $\mathcal{M}_{t-1}$ given the current state $\mathbf{x}_t$. At time step $t+1$ a measurement $\mathcal{M}_{t+1}$ becomes available. This is used to update the prior via Bayes' rule,

$$p(\mathbf{x}_{t+1}|\mathcal{M}_{t+1}) \propto p(\mathcal{M}_{t+1}|\mathbf{x}_t)p(\mathbf{x}_{t+1}|\mathcal{M}_t). \quad (10)$$

With this in mind, the tracking problem is stated as a Bayesian inference problem by use of (9) and (10).

Particle filtering is used with the purpose to estimate the filtering distribution $p(\mathbf{x}_t|\mathcal{M}_t)$ recursively. This is done through a random weighted sample set $\mathcal{S}_t^N = \{(\mathbf{x}_t^n, \pi_t^n)\}$, where n is the n^{th} sample of a state at time t weighted by π_t^n. The samples are drawn from the prediction prior distribution $p(\mathbf{x}_{t+1}|\mathcal{M}_t)$. The samples are weighted proportionally to the observation likelihood $p(\mathcal{M}_t|\mathbf{x}_t)$ given by the contour hypotheses. This sample set propagates into a new sample set $\mathcal{S}_{t+1}^N$, which represents the posterior probability distribution function $p(\mathbf{x}_{t+1}|\mathcal{M}_{t+1})$ at time $t+1$.

4.4 Constraining the Hypotheses

Corneal reflections, caused by illumination, may confuse the algorithm to weigh some of the hypotheses unreasonably high compared to others. This issue is illustrated left in figure 6, where the relative normalized weighting is colored in a temperature scale - Blue indicates low, while red high scores. By using robust statistics, these hypotheses are treated as outliers and therefore rejected.

The contour algorithm may fit to the sclera rather than the iris. This is due to the general formulation of absolute gray level differences $\Delta\mathcal{M}$[4], which seeks to detect contours in a general sense. An example is depicted in figure 7, where the image evidence of the contour surrounding the sclera is greater than the one around the iris. It turns out that for a large number of particles, the maximum likelihood estimate prefers the contour around the white sclera when the gaze is turned towards the sides.

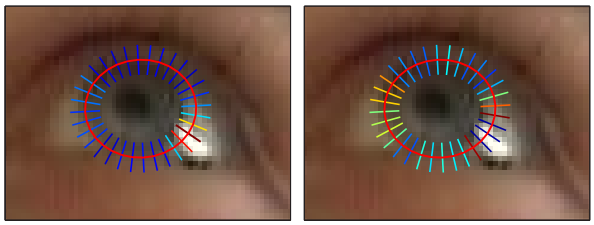

Figure 6: The relative normalized weighting of the hypotheses regarding one particle are colored in a temperature scale - Blue indicates low, while red high scores. (*Left*) Corneal reflections cause very distinct edges. Thus some hypotheses are weighted unreasonable high, which may confuse the algorithm. (*Right*) By use of robust statistics outliers are rejected. This results in a better and more robust estimate of the hypotheses regarding the contour.

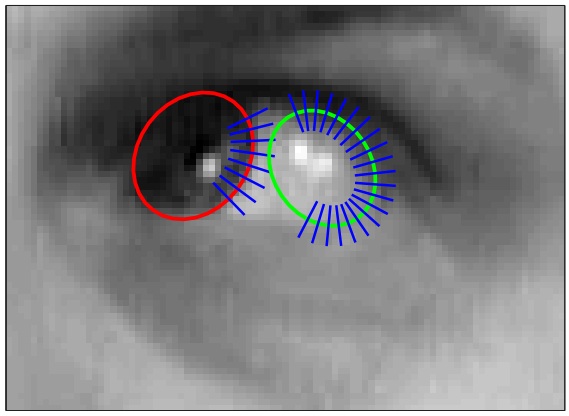

Figure 7: This figure illustrates the importance of the gray level constraint. Due to the general formulation of absolute gray level differences, the right contour has a greater likelihood, and the algorithm may thus fit to the sclera. Note the low contrast between iris and skin.

As a consequence, we propose to constrain the hypotheses. Intuitively, the average intensity value of the inner ellipse could be compared to some defined outer region as seen in expression (2). This is a poor constraint due to corneal reflection causing white blobs in the pupil area. The robustness of the active contour algorithm is increased by weighing the belief of hypotheses and utilizing robust statistics to reject outliers.

We propose to weigh the hypotheses through a sigmoid function, applied on the measurement line $\mathcal{M}$, defined as,

$$\mathcal{W} = \left(1 + \exp\left(\frac{\mu_i - \mu_o}{\sigma_w}\right)\right)^{-1} \quad (11)$$

where σ_w adjust the slope of weighting function, μ_i and μ_o are the mean values of the inner and outer sides of the contour respectively. The function is exemplified in figure 8. This has the

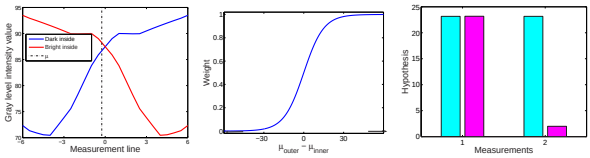

Figure 8: (*Left*) The two lines depicts the gray level intensity of two measurement lines - The blue one where the inner part of the ellipse is dark, and the red in the reverse case. (*Middle*) The shifted hyperbolic tangents is utilized as weighting function. Note, the limit values are in range $[-255; 255]$. (*Right*) The cyan bars indicates the hypothesis value before weighting, while the pink is after. Measurement *1* - The blue line - is nearly unchanged, while *2* - the red line - is suppressed.

effect of decreasing the evidence when the inner part of the ellipse is brighter than the surroundings. In addition, this relaxes the importance of the hypotheses along the contour around the eyelids, which improves the fit.

4.5 Maximum a Posteriori Formulation

The dynamic model may, in certain outlier cases, grow or shrink the contour to a degree, from where the algorithm gets lost. As a consequence, we propose to constrain on the shape of the ellipse in analogy to section 3.1. The parameters defining an ellipse is assumed normal distributed with mean μ and covariance $\mathbf{\Sigma}$. The prior distribution of these parameters are then defined,

$$p(\mathbf{x}) = \mathcal{N}(\mu, \mathbf{\Sigma}) \propto \exp\left(-\frac{1}{2}(\mathbf{x}-\mu)^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\mu)\right), \tag{12}$$

where the normalization factor has been omitted. The mean and covariance are estimated in a training sequence.

Combining the priors - presented in this section - with the likelihood, results in the *Maximum a Posteriori* formulation (MAP), where the goal is to maximize,

$$p(\mathbf{x}|\mathcal{M}) \propto p(\mathcal{M}|\mathbf{x})p(\mathbf{x}). \tag{13}$$

By incorporation of prior knowledge about the shape, with the prediction prior and observation likelihood (10), the robustness increases considerably and the parameters are constrained to avoid the algorithm to break down due to infinite increase or decrease of parameters.

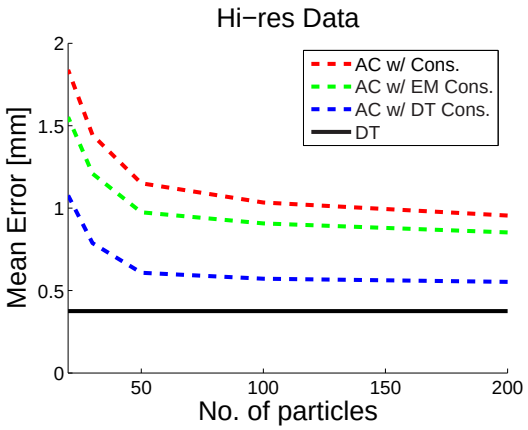

Figure 9: The error of the algorithms as a function of the number of particles for the high resolution data.

5 Results

A number of experiments have been performed with the proposed methods. We wish to investigate the importance of image resolution. Therefore the algorithms are evaluated on two datasets. One containing close up images, and one containing a down-sampled version hereof.

The algorithms estimate the center of the pupil. For each frame the error is recorded as the difference between a hand annotated ground truth and the output of the algorithms. This may lead to a biased result due to annotation error. However, this bias applies to all algorithms and a fair comparison can still be made.

Figure 9 and 10 depicts the error as a function of the number of particles used, for low resolution and high resolution images respectively. The errors for three different active contour (AC) algorithms are shows; basic, with EM refinement, with deformable template (DT) refinement. The error of the deformable template (DT) algorithm, initialized by double threshold, is inserted into the plot.

It can be seen that the proposed constraints on the active contour generally improves the accuracy of the fit. The refinement by the deformable template performs better than the EM method. The cost is an increased number of computations, which is resolution dependent. Nonetheless, the deformable template method, initialized by double thresholding, is seen to outperform all active contour algorithms.

The table in figure 5 lists the mean error in accuracy in centimeters and degrees. Also listed is the computation time in frames per section of

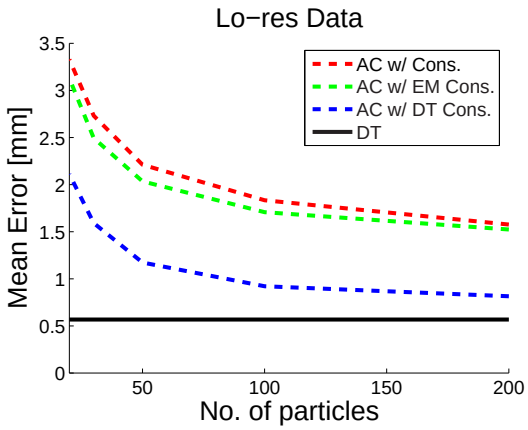

Figure 10: The error of the algorithms as a function of the number of particles for the low resolution data.

Hi-res	$E(x,y)$[mm]	$E(\theta)$	[frame/s]
AC	0.9	4.1	0.54
AC w/EM	0.8	3.7	0.49
AC w/DT	0.5	2.3	0.25
DT	0.3	1.4	2.2
Lo-res	$E(x,y)$[mm]	$E(\theta)$	[frame/s]
AC	1.5	7.3	0.57
AC w/EM	1.5	6.9	0.55
AC w/DT	0.8	3.7	0.49
DT	0.5	2.3	8.4

Table 1: Speed and precision comparison of the algorithms. The active contour uses 200 particles.

a Matlab implementation run on a 2.4Ghz PC. In general, the accuracy improves with high resolution as seen in table 5. However, the methods utilizing deformable template matching are less sensitive. The computation time for the basic active contour and EM refinement methods are independent of resolution. A significant increase in speed is noticed for the deformable template methods.

6 Conclusion

In this paper we have presented heuristics for improvement of the active contour method proposed by [7]. We have shown increased performance by using the prior knowledge that the iris is darker than its surroundings. This prevents the algorithm from fitting to the sclera as seen in figure 7.

Also presented is a novel approach to eye tracking based on a deformable template initialized by a simple heuristic. This enables the algorithm to overcome rapid eye movements. The active contour method handles these by broadening the state distribution and thus recovering the fit in a few frames. Furthermore, the accuracy is increased by fitting to the pupil rather than iris. This is particularly the case when a part of the iris is occluded as seen in figure 11.

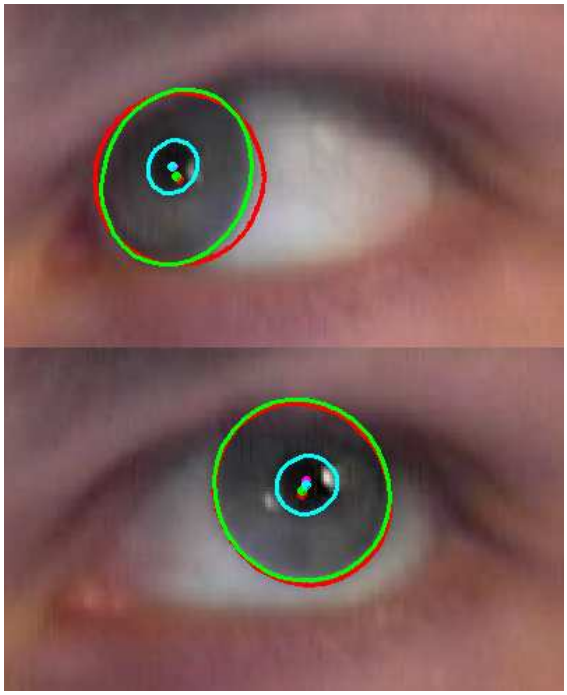

Figure 11: The resulting fit on two frames from a sequence - the red contour indicates the basic active contour, green indicates the EM refinement and the cyan indicates the deformable template initialized by the heuristic method. The top figure illustrates the benefit fitting to the pupil rather than the iris. Using robust statistic the influences from corneal reflections on the deformable template fit are ignored as depicted in the bottom image.

It is shown that the deformable template model is accurate independent of resolution and it is very fast for low resolution images. This makes it useful for head pose independent eye tracking.

Acknowledgements

We wish to thank Hans Bruun Nielsen, Department of Informatics and Mathematical Modelling - Technical University of Denmark, for providing the optimization implementation. Additionally, we wish to thank Dan Witzner Hansen, IT-University of Copenhagen, for inspiring and insightful discussions. This research is supported by the IST Network of Excellence - Communication by Gaze Interaction (COGAIN).

References

[1] Jr. Adams, R.B. and R.E. Kleck. Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, 2003.

[2] R.B. Jr. Adams, H.L. Gordon, A.A. Baird, N. Ambady, and R.E. Kleck. Effects of gaze on amygdala sensitivity to anger and fear faces. *Science*, 300:1536–1537, 2003.

[3] Communication by Gaze Interaction. www.cogain.org.

[4] Jens Michael Carstensen. *Image analysis, vision and computer graphics*. Technical University of Denmark, Kgs. Lyngby, 2 edition, 2002.

[5] Jonathan Gratch and Stacy Marsella. Tears and fears: Modeling emotions and emotional behaviors in synthetic agents. *Proceedings of the 5th International Conference on Autonomous Agents, Montreal, Canada*, June 2001.

[6] D. W. Hansen, J. P. Hansen, M. Nielsen, A. S. Johansen, and M. B. Stegmann. Eye typing using markov and active appearance models. In *IEEE Workshop on Applications of Computer Vision - WACV*, pages 132–136, dec 2002.

[7] Dan W. Hansen and Arthur E. C. Pece. Iris tracking with feature free contours. In *Proc. workshop on Analysis and Modelling of Faces and Gestures: AMFG 2003*, October 2003.

[8] Takahiro Ishikawa, Simon Baker, Iain Matthews, and Takeo Kanade. Passive driver gaze tracking with active appearance models. In *Proceedings of the 11th World Congress on Intelligent Transportation Systems*, October 2004.

[9] T. Kawaguchi, D. Hidaka, and M. Rizon. Detection of eyes from human faces by hough transform and separability filter. In *ICIP00*, pages Vol I: 49–52, 2000.

[10] Kyung-Nam Kim and R.S. Ramakrishna. Vision-based eye-gaze tracking for human computer interface, 1999.

[11] D. Leimberg and M. Vester-Christensen. Eye tracking. Master's thesis, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2005. Supervised by Prof. Lars Kai Hansen.

[12] Y. Matsumoto and A. Zelinsky. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *AFGR00*, pages 499–504, 2000.

[13] Tsuyoshi Moriyama, Jing Xiao, Jeffrey Cohn, and Takeo Kanade. Meticulously detailed eye model and its application to analysis of facial image. In *Proceedings of the IEEE Conference on Systems, Man, and Cybernetics*, pages 629 – 634, 2004.

[14] J. Pelz, R. Canosa, J. Babcock, D. Kucharczyk, A. Silver, and D. Konno. Portable eyetracking: A study of natural eye movements, 2000.

[15] Rainer Stiefelhagen, Jie Yang, and Alex Waibel. Tracking eyes and monitoring eye gaze. In *Workshop on Perceptual User Interfaces*, Banff, Canada, October 1997.

[16] Rainer Stiefelhagen, Jie Yang, and Alex Waibel. Estimating focus of attention based on gaze and sound. In *PUI '01: Proceedings of the 2001 workshop on Perceptive user interfaces*, pages 1–9. ACM Press, 2001.

[17] Joseph van der Gracht, V. P. Pauca, Harsha Setty, Ramkumar Narayanswamy, Robert Plemmons, Sudhakar Prasad, and Todd Torgersen. Iris recognition with enhanced depth-of-field image acquistion. volume 5438, pages 120–129. SPIE, 2004.

[18] J.G. Wang and E. Sung. Study on eye gaze estimation. *IEEE Transactions on Systems, Man and Cybernetics - Part B: Cybernectics*, 32(3):332–350, June 2002.

[19] Jian-Gang Wang, Eric Sung, and Ronda Venkateswarlu. Eye gaze estimation from a single image of one eye. In *ICCV*, pages 136–143, 2003.

[20] X. Xie, R. Sudhakar, and H. Zhuang. A cascaded scheme for eye tracking and head movement compensation. *T-SMC*, A28:487–490, 1998.

[21] X.D. Xie, R. Sudhakar, and H.Q. Zhuang. Real-time eye feature tracking from a video image sequence using kalman filter. *SMC*, 25(12):1568–1577, December 1995.