Reliable and Fast Eye Finding in Close-up Images

Theodore A. Camus
Sarnoff Corporation
Princeton, NJ 08543-3500
tcamus@sarnoff.com

Richard Wildes
Department of Computer Science
Centre for Vision Research
York University
Toronto, Ontario
wildes@cs.yorku.ca

Abstract

This paper describes a method for quickly and robustly localizing the iris and pupil boundaries of a human eye in close-up images. Such an algorithm can be critical for iris identification, or for applications that must determine the subject's gaze direction, e.g., human-computer interaction or driver attentiveness determination. A multi-resolution coarse-to-fine search approach is used, seeking to maximize gradient strengths and uniformities measured across rays radiating from a candidate iris or pupil's central point. An empirical evaluation of 670 eye images, both with and without glasses, resulted in a 98% localization accuracy. The algorithm has also shown robustness to weak illumination and most specular reflections (e.g., at eyewear and cornea), simplifying system component requirements. Rapid execution is achieved on a 750 MHz desktop processor. [1]

1. Introduction

The ability to automatically localize the eye of a human subject in a non- invasive fashion with accuracy, precision and speed has a number of significant real-world applications. Human-computer interaction can exploit such an ability to determine gaze direction [6, 7]. Driver alertness and attention can be monitored via assessment of eye behavior [1]. Iris recognition biometric-based human identification depends on localizing the iris of a subject in a larger field of view image [9]. In order to respond successfully to the demands of such applications a number of challenges must be surmounted. The exact shape of the eye varies with attitude of the subject relative to the sensor, eye-lid closure and individual biometric differences. The contours that define the shape (e.g., pupil and iris boundaries) can be imaged with low contrast. Specular reflections from eyewear can obscure large regions of interest. Since human eyes are often in motion, processing must be efficient if eye position is to be reported in a time relevant fashion.

In response to these motivations and challenges, this paper describes an algorithm for finding and localizing a subject's iris and pupil boundaries given a relatively close-up image of the eye, i.e., an eye within the range of approximately 130 to 160 pixels diameter in a 640x480 image. Thus, it assumes that a complete system has the means to localize the eye in a wider-angle field of view image, and subsequently capture a narrower field of view image containing the eye somewhere in the image. Several systems for real-time eyetracking at a macro scale have been presented in the literature. For example, [4] describes an iris identification system which finds the subject's eyes in 3-D using stereo and facial template finding algorithms with continuous operation at 2 Hz. Also, [5] describes a real-time eye tracker using the red-eye effect caused by retinal reflection of on-axis light. For present purposes, it will be assumed that the full iris is contained within the image.

A number of previous efforts have considered precision pupil/iris localization in close up images. An approach was documented based in a hierarchical neural network, which during its latter stages of processing made use of classifiers tuned for the features of present concern [10]. A deformable template was proposed that fit a parameterized geometric model to high contrast contours [12, 3]. An approach based on an analysis of the image gradient pattern corresponding to an eye, including motion information also has been presented [8]. The pupil/iris boundaries also have been detected and localized via edge detection following by a Hough transform defined to capture the contours of interest [11]. In the light of previous research, the outstanding features of the work presented in this paper are as follows. The developed approach is capable of accurate and precise localization, even in the presence of low contrast images, eccentric views and partial obscuration by specularities. Performance has been evaluated relative to a database of 670 iris images with 98% correct localization of the eye. The

implementation executes rapidly on a commercially available 750MHz desktop computer.

2. Technical approach

The input to the algorithm consists of a single greyscale image containing a single human eye. The output of the algorithm is the position of the detected pupil and iris boundaries. The iris boundary is modelled as a circle, described by its (x, y) center and radius r_i. The pupil boundary is modelled as an ellipse of the form

$$r'_p = r_p/(1 + \epsilon(1 + cos(2(\phi + \theta)))), \qquad (1)$$

where r_p is the half the length of the major axis, ϕ is the angle of the minor axis, r'_p is the length of the ray from the point on the ellipse at angle $\phi + \theta$ to its center, and ϵ is a measure of eccentricity. Thus, the algorithm output consists of the five values $(x, y, r_i, r_p, \phi, \epsilon)$, i.e. the shared (x, y) center of the iris and pupil, the iris radius r_i, the radius of the circle circumscribing the pupil r_p, the pupil orientation ϕ, and the eccentricity parameter ϵ. (See below for discussion of the circle/ellipse modeling of the iris/pupil.)

The algorithm operates in the following fashion. First, bright specularities in the image are filled in to ameliorate their effect on subsequent processing. Second, candidate seed locations are generated to provide initial conditions for the pupil/iris boundary fitting operation. Third, for each candidate seed, (x, y), pupil and iris boundary parameters are recovered. The parameters associated with the best fit across all seeds are returrned. The remainder of this section documents each of these steps in more detail.

Specularity filling The first step of the developed approach is to attenuate the influence of specularities in the image, which can disrupt a pupil-iris or iris-sclera boundary. These very bright areas are formed from reflections off of the cornea of the eye, or off of the lenses of glasses or contacts if present. The detection of regions for filling in is accomplished via a simple threshold on the local intensity value; for all cases discussed in this paper, it was set to 250 out of a range of [0,255]. Given a detected region of specularity, filling in is accomplished via bilinear interpolation from the surrounding image intensities.

Seed point selection Seed points for the boundary fits are selected as points that correspond to local minima of image intensity. For current purposes, a local minimum is defined as an image intensity value below a certain global threshold and also one that is the smallest value within a 5x5 pixel region. For cases discussed in this paper, the global threshold was set to 128 from a range of [0,255].

Boundary representation The anatomy of the human eye dictates that the pupil and iris boundaries are approximately circular and concentric. Owing to the imaging process, these contours will be imaged as ellipses, with the pupil contour constrained to lie within the iris contour.
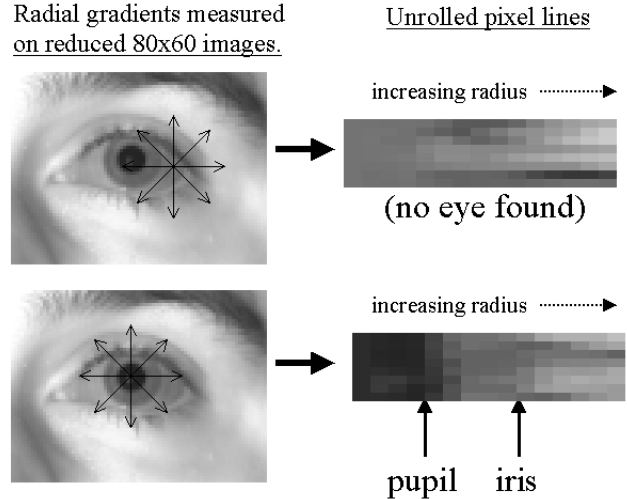

Figure 1: Mapping of radial rays into a polar representation; pupil and iris boundaries become vertical edges.

For the present purposes, only the pupil's boundary will be modelled as a non-circular structure, due to generally greater variability in its naturally occurring shape, as well as deformability due to naturally occurring constrictions and dilations (future work will include a more flexible representation of the iris's contour as well). These constraints are exploited in the boundary fitting algorithm in terms of both the geometry of the proposed representation as well as the photometric pattern that is taken as indicative of the contours of interest.

Owing to the geometry of the pupil and iris boundaries, it is advantageous to make use of a polar representation of the eye region, e.g., as centered about a candidate center seed returned by seed point selection. In particular, the input Cartesian image representation is remapped into (r, θ)-space, with r the distance from the center and θ the polar angle, see Figure 1. Notice that the pupil and iris boundary contours are mapped to (near) vertical edges in this representation, so that the detection and localization process is now simplified to the extraction of (near) vertically contiguous loci of significant contrast in image intensity, with nearness to strict linearity/verticality depending on the eccentricity of the imaged contours.

Within the polar representation, the pupil and iris boundaries are characterized in similar fashions. In both cases, three observations are exploited. (i) Since the reflective properties of the eye change as one crosses from pupil to iris and from iris to sclera, there should be a local maximum in the radial derivative of the corresponding image intensities. (ii) Since the reflective properties are relatively homogeneous on each side of these boundaries, the intensity derivatives should be relatively uniform along the polar direction. (iii) The inner region of each boundary should, typically, be darker than the outer region.

For both the pupil and iris boundaries, the noted observations can be formalized via a component-goodness-of-fit metric for a candidate boundary parameters being considered with respect to a given center for the polar coordinate system. The component-goodness-of-fit is defined as

$$\sum_{\theta=1}^{n} \left((n-1)\|g_{\theta,r})\| - \left(\sum_{\phi=\theta+1}^{n} \|g_{\theta,r} - g_{\phi,r}\| \right) - g_{\theta,r}/8 \right)$$
(2)

with n standing for the number of discrete values of the polar variable θ that are considered (in current implementation, $n = 8$) and $g_{\theta,r}$ standing for the directional derivative of image intensity in the radial direction. In the goodness-of-fit (2), the first term captures the weighted summed strength of the gradients across the boundary, the second term captures the uniformity of the gradients along the boundary, and the last term captures a slight preference for darker regions on the boundary interior.

Given a goodness-of-fit measurement for the pupil and a goodness-of fit measurement for the iris, each calculated according to (2), the overall-goodness-of-fit is a sum of the two component measures, modulated by a pair of constraints on expected relationships between the pupil and iris. The first constraint helps restrict the iris-pupil solution to lie within expected radii ratios for normal human eyes. In particular, if the iris-pupil radii ratio is greater than 4.0 or less than 1.75, then the quality measure is multiplied by a ratio of $(4.0/ratio)^3$ or $(ratio/1.75)^3$, respectively. The second constraint reflects the observation that the average iris intensity, $iris$, is generally greater than the average pupil intensity, $pupil$. In cases where this is not true, the quality measure is multiplied by $(iris - pupil)^2$; this prevents the pupil boundary finder from latching onto a very bright region, e.g., a specularity.

Parameter estimation The recovered parameters of the pupil and iris boundaries are estimated with respect to the defined overall-goodness-of-fit metric. This result is achieved in two stages. (i) For each point returned by seed point selection, boundary parameters are recovered so that the overall-goodness-of-fit is maximized locally. (ii) The global maximum is taken across all initial seeds.

The local estimation operates as follow. (i) For a given center seed, the polar representation is constructed. (ii) An initial pupil radius is calculated by searching over radius values to maximize the component-goodness- of-fit metric (2). For this preliminary fit, the pupil boundary is approximated as a circle. (iii) An initial iris radius is recovered by searching over radius values to maximize the same component goodness metric, but with the search for radius values constrained to lie outside the calculated pupil radius and polar samples, θ, restricted to lateral directions to avoid eyelid intrusion. (iv) An initial overall-goodness-of-fit is calculated as the weighted sum of the initial component mea-

sures, according to the weighting described earlier. (v) The initial estimate is improved via a gradient search by considering four additional center locations to the north, south, west and east of the current center location. For each additional location, the component and overall goodness-of-fit measures are calculated and the results that yield the best fit are taken as the new estimate. This procedure iterates until a local maximum is located. (vi) The ellipse parameters for the pupil are recovered via brute force search for the values of ϕ and ϵ in parameterization (1) that maximize the goodness of fit (2). The resulting boundary parameters are then returned. The entire boundary fitting procedure is embedded in a coarse-to-fine refinement scheme using Gaussian image pyramids [2], whereby the result recovered at coarser spatial resolution can be used to initialize processing at finer resolution for improved precision with modest expense.

Relationship to other approaches The developed method most closely resembles deformable template techniques [12], in that it iteratively adjusts the parameters of a geometric contour model to best fit the captured pattern of image intensity. In this regard, significant differences are in terms of the use of a polar image representation to simplify boundary representation, details of the photometric pattern structure that is considered representative of the boundaries of interest and a particularly efficient implementation. Combined, these attributes allows for accurate and precise detection and localization with rapid execution, even in the presence of challenging acquisition scenarios, as documented in the next section. Also, of interest is the fact that the present approach makes no assumptions regarding the presence of strong retroreflections from the interior eye [5]; this allows the approach to have greater generality. Similarly, the approach used notably does not require the use of the presence and location of any reflections or glints on the eye to assist in the localization of the pupil or iris; thus, this algorithm is also more flexible than those requiring glints to find the eye.

3. Empirical results

To evaluate the described algorithm, the eyes of 11 human subjects were imaged, capturing a variety of natural eye shapes. Subject's placed their head in a chin rest for stable, in focus imaging. The eyes were imaged with iris diameter in the range [130,160] pixels in an overall 640×480 image. Variable obliquity of view was achieved by having subjects systematically adjust their gaze direction from extreme right to extreme left while their head pointed toward the camera. Near infra-red illumination (≈ 880nm) was employed for noninvasive acquisition; a matched filter was placed in front of the imager. Across several occasions, 640 images without glasses were acquired; an additional 30 images were acquired by having 6 of the subjects don their glasses.

Illustrative successes are shown in Figure 2. The upper left panel shows an good localization in the presence of extreme pupil constriction, with an iris-pupil diameter ratio of 3.5. The upper right panel shows good localization in the presence of very low contrast contours: 8 and 26 grey-level differences across the iris and pupil boundaries, respectively. The middle left panel shows good localization in the presence of high contrast, in particular the case where the iris-pupil contrast is lower than the iris-sclera contrast. The middle right panel shows good localization in the presence of specular reflections that obscure major portions of the boundaries of interest. The lower left and lower right panels shows good localization in the presence of obliquely imaged eyes.

Illustrative failures are shown in Figure 3. The upper left panel shows a failure due to the red-eye effect, caused by direct on-axis active illumination. The upper right shows a boundary localization failure due to extreme off-axis gaze direction. The lower left panel shows an example of failure due to extreme obscuration of the eye by specular reflections off of glasses. Finally, the lower right panel shows a correct eye localization, but slightly imperfect elliptical pupil boundary localization.

A total of 670 eye images were evaluated, comprising of 640 images from 11 subjects without eyewear, and a further 30 images from a subset of 6 of those subjects wearing glasses. Localization performance of the eye itself resulted in 99.5% accuracy for naked-eye cases and 66.6% accuracy for those subjects wearing glasses, for an overall average accuracy of 98% (based on visual inspection). 94% of these hits were then judged to be either good or very good elliptical fits of the pupil boundary, as shown in Figure 2, with the remaining 6% considered bad fits, as shown in Figure 3. All errors for the eyewear condition were due to severe specularity occlusion, as also shown in Figure 3. The depicted examples of specularities and off-axis gaze direction, respectively, in Figure 2 and Figure 3 give an indication of the breaking points of the algorithm for those two conditions. Ultimately, the measure of the precision and accuracy of the localization must be a quantitative measure of the intended application, such as iris recognition error rates [3, 11].

The current implementation of the described algorithm uses almost entirely fixed-point arithmetic. This, in combination with the multi-resolution approach, results in rapid execution for the circular-model fitting stage: Execution speed is 55 Hz on a 750 MHz UltraSPARC-III processor, including the specularity filling and software pyramid-processing. This is about 3.5 times as fast as the combined specularity filling, iris boundary finding, and pupil boundary finding components reported in the work of [3] (after scaling for clock speed). The ellipse-fitting modelling of the pupil currently runs much slower than the circular-only model fitting algorithm, at 3 Hz on the same platform, due to the search over the ϕ and ϵ parameter space.

4. Summary

This paper has presented a robust, real-time algorithm for localizing the iris and pupil boundaries of an eye in a close-up image. It uses a multi-resolution approach to detect the boundary contours of interest quickly and reliably, even in cases of very low contrast, specular reflections and oblique views. Empirical results have demonstrated a 98% accuracy rate with rapid execution on a 750 MHz desktop computer. It is suggested that this algorithm can serve as an essential component for applications such as iris recognition and gaze direction determination.

References

[1] M. Betke, W. Mullally, "Preliminary Investigation of Real-Time Monitoring of a Driver in City Traffic", *Proc. IEEE Intel. Vehicles Symp.*, 563-568, 2000.

[2] P. Burt, E. Adelson, "The Laplacian Pyramid as a Compact Image Code", *IEEE TC*, 31(4):532-540, 1983.

[3] J. Daugman, "Statistical Richness of Visual Phase Information: Update on Recognizing Persons by Iris Patterns", *International Journal of Computer Vision,* 45(1):25-38, 2001.

[4] M. DellaVecchia, T. Chmielewski, T. Camus, M. Salganicoff, M. Negin, "Methodology and Apparatus for Using the Human Iris as a Robust Biometric", *SPIE Oph. Tech.,* 1998.

[5] A. Haro, M. Flickner, I. Essa, "Detecting and Tracking Eyes By Using Their Physiological Properties, Dynamics, and Appearance", *Proc. IEEE CVPR,* 163-168, 2000.

[6] T. Hutchinson, K.P. White, W.N. Martin, K.C. Reichert, and L.A. Frey, "Human-Computer Interaction Using Eye-Gaze Input", *IEEE TSMC,* 19(6):1527-1534, 1989.

[7] F. Jacob, "The use of eye movements in human-computer interaction techniques", *ACM TIS,* 9(93):152-169, 1991.

[8] R. Kothari, J. Mitchell, "Detection of eye locations in unconstrained visual images", *Proc. IEEE ICIP*, 519-522, 1996.

[9] M. Negin, T. Chmielewski, M. Salganicoff, T. Camus, U. von Seelen, P. Venetianer, G. Zhang, "An Iris Biometric System for Public and Personal Use", IEEE Computer, 70-75, 2000.

[10] J. Vincent, J. Waite, D. Myers, "Precise location of facial features by a hierarchical assembly of neural networks", *Proc. IEEE Conf. on ANN*, 69-73, 1991.

[11] R. Wildes, "Iris Recognition: An Emerging Biometric Technology", *Proc. IEEE*, 85(9):1348-1363, September 1997.

[12] A. Yuille, P. Hallinan, D. Cohen, "Feature extraction from faces using deformable templates", *IJCV* 8(2): 99-111, 1992.
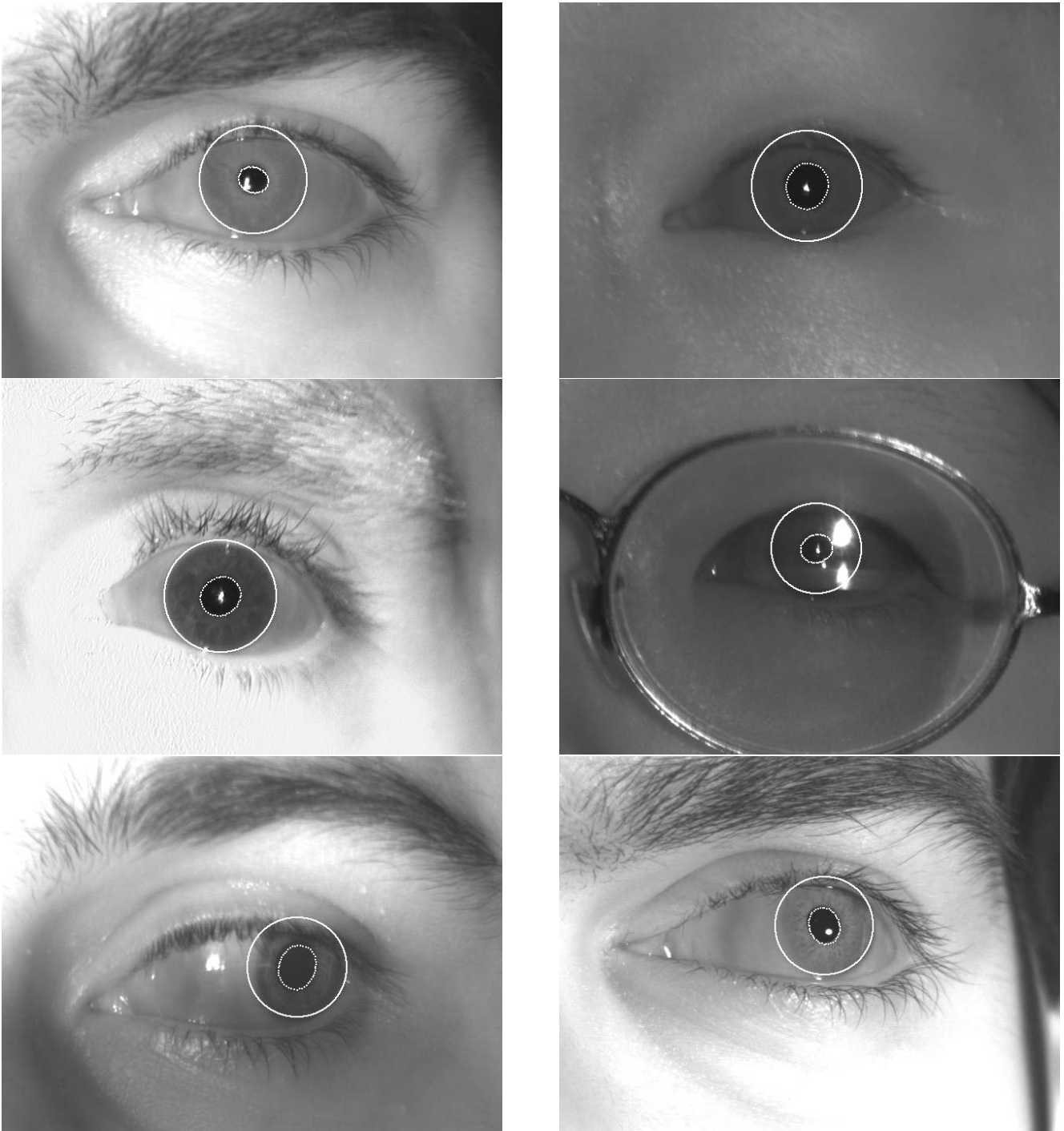
Figure 2: Localization Results. Correct localization for a highly constricted eye (upper left). Correct localization in the presence of low contrast (upper right). Correct localization in the presence of high contrast (middle left). Correct localization in the presence of specularity-based occlusion (middle right). Correct localization in the presence of oblique view (lower left). Correct localization in the presence of another oblique view (lower right). Localization indicated via white contour overlays.
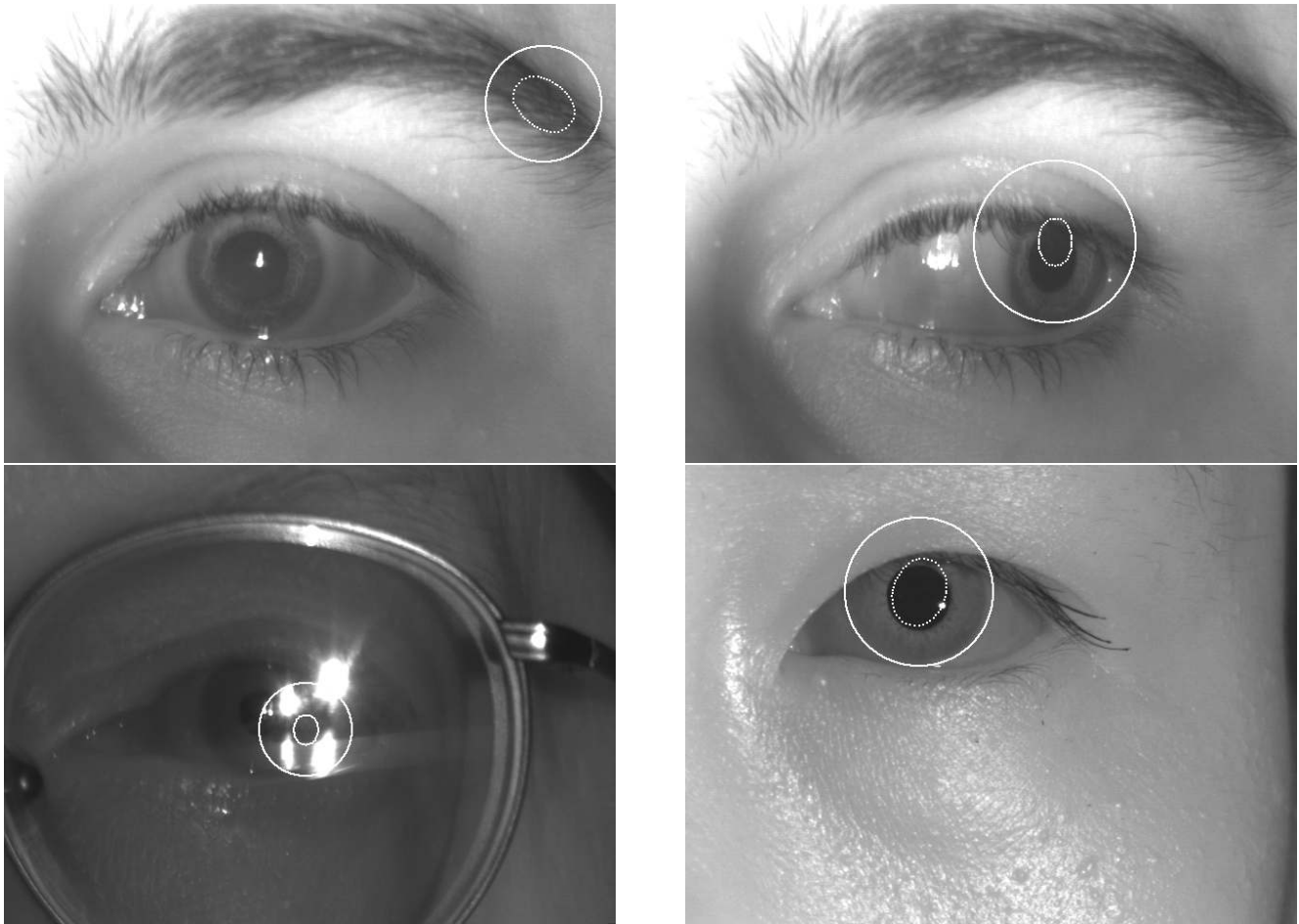
Figure 3: Localization Results. Failure in the presence of red-eye retinal reflection of active illumination (upper left). Failure due to extreme oblique view (upper right). Failure in the presence of specularity-based occlusion (lower left). Imperfect elliptical fit of pupil (lower right). Localization indicated via white contour overlays.