Automatic Dry Eye Detection

Tamir Yedidya[1], Richard Hartley[1], Jean-Pierre Guillon[2], and Yogesan
Kanagasingam[2]

[1] The Australian National University, and National ICT Australia *
[2] Faculty of Medicine and Health Sciences, Lions Eye Institute, Australia

Abstract. Dry Eye Syndrome is a common disease in the western world,
with effects from uncomfortable itchiness to permanent damage to the
ocular surface. Nevertheless, there is still no objective test that provides
reliable results. We have developed a new method for the automated
detection of dry areas in videos taken after instilling fluorescein in the
tear film. The method consists of a multi-step algorithm to first locate the
iris in each image, then align the images and finally analyze the aligned
sequence in order to find the regions of interest. Since the fluorescein
spreads on the ocular surface of the eye the edges of the iris are fuzzy
making the detection of the iris challenging. We use RANSAC to first
detect the upper and lower eyelids and then the iris. Then we align the
images by finding differences in intensities at different scales and using a
least squares optimization method (Levenberg-Marquardt), to overcome
the movement of the iris and the shaking of the camera. The method has
been tested on videos taken from different patients. It is demonstrated
to find the dry areas accurately and to provide a measure of the extent
of the disease.

1   Introduction

One of the roles of the tear film is the maintenance of corneal epithelial integrity
and transparency which is achieved by keeping the ocular surface continuously
moist. The pre-ocular tear film in humans does not remain stable for long periods
of time [3]. When blinking is prevented, the tear film ruptures and dry spots
appear over the cornea [5].

The Fluorescein Break Up Time (FBUT) test was designed by Norn [8] and
called corneal wetting time. A moistened fluorescein strip is applied to the bulbar
conjunctiva and after two or three blinks to spread the fluorescein evenly, the tear
film is viewed with the help of a yellow filter in front of a slit-lamp biomicroscope.
When a dark area appears in the uniform colouration, it represents the rupture
of the tear film and the time elapsed since the last blink is recorded as break-up
time (BUT). The shorter the BUT, the more likely the diagnosis of dry eye [9].
The degree of blackness is related to the depth of the breakup. The deeper the
break, the greater the chances of ocular surface damage. If the eyes are kept
open, the area of the break will increase in size and breaks may appear in new
areas over the cornea.

Fig. 1. EyeScan portable imaging system for Fluorescein tear film recording

This is the test of tear film stability most commonly used by clinicians and a few improvements have been presented over the years [4]. A BUT inferior to 10 seconds usually reveals an anomaly of the pre-ocular tear film. If the film ruptures repeatedly in the same spot a superficial epithelial abnormality must be suspected. If the break up occurs over the center of the cornea, a decrease in visual acuity will be induced.

The location, shape, progression and depth of the original and successive breaks give clinical indications as to the cause of the break and is helpful to determine what treatment to choose. This is not possible with our current subjective observation and grading clinical routines. With the development of the EyeScan system, as seen in Fig. 1, an affordable and easy to operate system of video recording of anterior ocular structures is available.

To our knowledge, automatic methods towards finding the dry eye areas are sparse in the current literature. However, there are some existing approaches for locating the iris as part of another applications. Usually assuming its shape is a perfect circle, the methods mostly use circle fitting algorithms. Daugman [1] focused on iris recognition, but he first finds the pupil-iris and iris-sclera borders by searching over all circles in different radii for the ones that gives the maximum contour integral derivative. Ma and el [7] also locate the iris as part of their iris recognition procedure. They first locate the darker pupil and then the iris using Canny edge detector and Hough transform to estimate the iris location. However, in our videos the pupil is not visible at all as shown in Fig. 2, and the boundaries between the iris and the sclera are much fuzzier due to
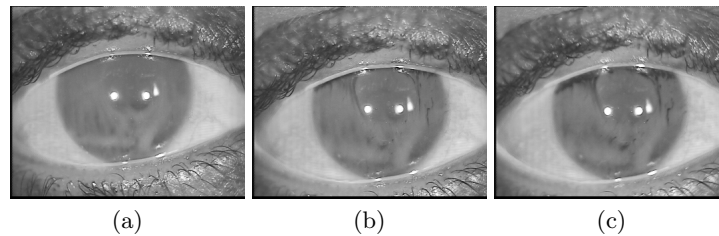

Fig. 2. *Samples of eye images after instilling fluorescein. (a) Immediately after a blink (b) half way between blinks (c) just before a blink. The darker areas are the dry eye areas which evolve through the sequence. In (c) the upper eyelid opened.*
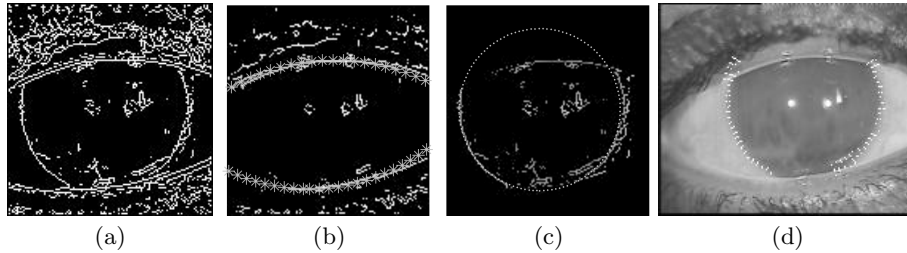
Fig. 3. *Steps for locating the iris (a) cropped edge map (b) cropped edge map for locating the eyelids and their detection (c) cropped image after removing outliers and the fitted iris (d) selected pixels used for LM on the original image*

the fluorescein spreading. The Hough transform is a commonly used method for circle detection, however it is very slow and needs a few iterations to find the correct radius. The Fast Radial Symmetry [6] finds strong gradients in opposite directions and picks the gradients according to a strictness parameter, but have difficulties coping with images where only parts of the circle are visible.

2 Algorithm

The dry areas always appear as darker areas in the fluorescent image, as seen in Fig. 2(c). Trying to apply a single threshold to the grey values does not work, since the spreading of the fluorescein and the lighting conditions change from patient to patient. Therefore, the intensity of the fluorescein cannot be predicted. Moreover, dry areas might become visible only at parts of the sequence, for example, as the eyelids open or close. Finally, the camera is hand held, therefore there is a considerable movement of the iris. To address these problems, we present our algorithm that is based on aligning the video.

2.1 Fitting a circular model

The images of interest in the video are those between blinks. To that end, we first find all the blinks and half-blinks and treat each sequence individually. Blinks are detected when consecutive images have big difference in intensity. An edge map of the image immediately after a blink is created using the sobel operator and non-maximal edges are suppressed. We use RANSAC [2] with three parameters, (x, y, r), on the edge map to locate the iris. At each iteration a set of three points is used for fitting a circle and the number of pixels passing through it is counted. In order for RANSAC to perform well the outliers need to occupy a small percentage of the total pixels as demonstrated in [2]. However, the edges of the iris are usually not strong and are hindered by the eyelashes and strong illumination changes around the iris's borders. In some cases, the ratio between the inliers (the iris) and the outliers is 1 to 10 as seen in Fig. 3(a).

Therefore, we initially segment the upper and lower eyelids and then remove all pixels above and below them respectively. The eyelids are detected by creating
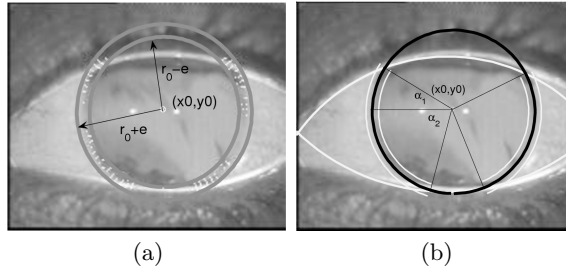
Fig. 4. *Explanation of the LM (a) the area between the two circles is used as the search area for the iris. The white pixels are those found by LM to best fit the iris (b) the detected eyelids and the iris are segmented, the white arcs are those corresponding to the circles from part (a). Only the area between the two arcs is used for searching*

another edge map with a higher threshold and then running RANSAC with three parameters and assuming the eyelids are circles with a very big radius, see Fig. 3(b). The method proves to be a reliable way to detect the eyelids. However, sometimes the lower eyelids might not be visible, so extra care is taken to ensure parts of the iris are not removed. Then it is possible for RANSAC to find the iris as seen in Fig. 3(c). At each iteration three points are randomly picked up and a few rules are imposed regarding the points distribution and limiting the search for the expected range of iris's radii existing in patients.

After the iris has been detected we use Levenberg-Marquardt(LM) [2] to do the fine-tuning by iteratively minimizing a non-linear function for fitting a circle. The initial estimation (x_0, y_0, r_0) is the one found for the iris in the RANSAC step. We define the maximum error e to be a few pixels, usually around four, and then compute the gradient's magnitude for all pixels in the annulus formed by the circles of radii $r_0 - e$ and $r_0 + e$ and the center (x_0, y_0), as depicted in Fig. 4(a). However, since the iris is usually only partly visible because of the eyelids, we find the angles α_1 and α_2 between the estimated iris and the upper and lower eyelids respectively, see Fig. 4(b). Only the pixels that are on the arc between the eyelids are taken. If most of the iris is visible (as in Fig. 5(a)), all pixels are taken. These pixels are sorted according to their magnitude and the strongest 180 are considered for the LM minimization. The diameter of the iris in most of the population is between $9mm$ to $13mm$, of the population The parameter to the function to be minimized is the vector (x, y, r) and at each iteration the set of pixels found are used to compute the sum of squares of the Huber distance $D_h(x_k, y_k)$ of each pixel (x_k, y_k) to the estimated iris:

$$\text{define} \quad d = \sqrt{(x_0 - x_k)^2 + (y_0 - y_k)^2} - r$$

$$D_h(x_k, y_k) = \begin{cases} d & \text{when } |d| \leq \epsilon \\ \text{sign}(d) \sqrt{(2|d| - \epsilon)\epsilon} & \text{when } |d| > \epsilon \end{cases} \tag{1}$$

The use of the Huber distance makes the penalty linear above the threshold ϵ, avoiding quadratic penalties to faraway outliers. Results of the LM with the chosen pixels are depicted in Fig. 3(d). The detection of the iris in the rest of the sequence is done in a similar way, each time using the previous estimation as

the initial guess. Since, we do not expect drastic movements of the iris, the LM converges quickly. It is worth mentioning that the time-consuming RANSAC is computed only once, for the first image.

2.2 Computing Image Translation

After locating the iris in each of the images, it is possible to align the images over the iris area. The need for such alignment is demonstrated in Fig. 5. Since, the iris is represented as a circle, we first simply align by scaling and translating one circle to another circle. Then a second alignment is done by using the grey levels of the iris area to find the best homography between the images. Every few images are treated as a block. The first image in the block is aligned to an image in the previous block and the other images are aligned to the first image in the block to avoid accumulating error. The alignment procedure is as follows:

1. Initialize the homography matrix H_0 to be the identity matrix. Since, we are interested only in translation and scaling the matrix has three unknowns. Define the initial region of interest(ROI) as the area that includes the iris and the eyelids, since they differ in grey level from the iris and help the alignment process.
2. The homography H_0 is used as our initial guess. A pyramid of scaled images for both images is built. For each scale the best homography is found starting from the smallest scale. Going up the pyramid is done by rescaling the transformation and the ROI.
3. Compute the homography by minimizing:

$$C(H) = \sum_{x \in ROI} (I_x^{(j)} - I_{Hx}^{(k)})^2 \qquad (2)$$

   where Hx is the transformed pixel in I_k, and j and k are indices for the two images being aligned.
4. The process is repeated for every image in the sequence.

It is difficult to show the results of an alignment unless in a video. However, Fig. 5 shows the averaged images over a sequence. The non-aligned image is completely blurred and cannot provide information regarding the dry areas. It is expected since the camera is hand held and the patient's gaze can move.

2.3 Segmentation of the dry area

Based on the two previous steps, the task of segmenting correctly the dry areas becomes possible. In section 2.1 we found the circle equations of the iris and the upper and lower eyelids. Since, the dry areas appear only on the iris, the area to be searched for dryness can be extracted. The pixels of interest will be those that become darker during the sequence. However, the intensities of the dry areas can
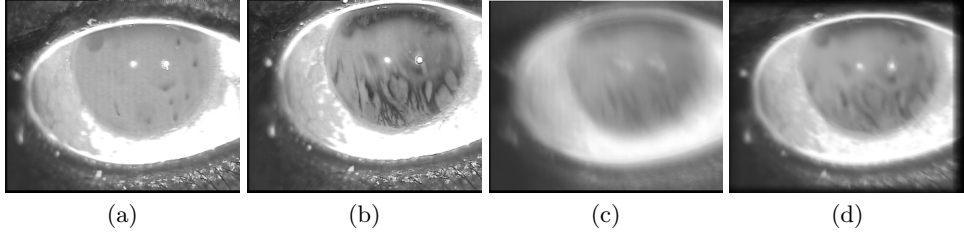
| (a) | (b) | (c) | (d) |

Fig. 5. *Since the camera is hand held and the iris can move, it is necessary to align the video: (a) image after a blink (b) last image before the next blink. The averaged image of all the images in the (c) non-aligned video (d) aligned video. The non-aligned result is blurred showing that the dry areas cannot be detected without aligning the images or by just taking the difference between images (a) and (b)*

change a lot from patient to patient as the fluorescein spreading varies and light conditions are not the same. To that end, we calculate a difference image:

$$D(x,y) = \sum_{k=i_0}^{i_0+\Delta} I_k(x,y) - \sum_{k=n-\Delta}^{n-1} I_k(x,y) < T_1 \tag{3}$$

assuming the number of images in the sequence is n, Δ is a small number of images and i_0 is chosen to be close to zero.

Ideally each pixel will have a monotonically decreasing intensity function through the sequence. However, it is not always the case. It is often that the dark areas are revealed only in parts of the sequence, for example, if the eyelids open or if a few pixels are aligned incorrectly especially near the borders of the iris with the eyelids. To overcome this, for each frame k we compute an average value $V_k(x,y)$ for the pixels in the frame averaged over a small neighborhood about a pixel (x,y). A smoothing is done by convolving with a gaussian to minimize the effects of erratic movement. Then we compute the following downs-ups ratio:

$$R(x,y) = \frac{\#downs}{\#ups} = \frac{\sum_{k=0}^{n-1}[V_{k+1}(x,y) < V_k(x,y) - d]}{\sum_{k=0}^{n-1}[V_{k+1}(x,y) > V_k(x,y) + d]} \tag{4}$$

where d is a small number around 0.5. The ratio $R(x,y)$ provides information how the pixel changes through the sequence and is dependent on the sequence length. Unlike Eq. 3, $R(x,y)$ detects pixels that only mildly become dry and also dry pixels near the iris's borders that their intensity fluctuates through the sequence. These pixels usually have low or unexpected $D(x,y)$ value. A weighted average of equations (3) and (4) and the intensity in the last frame serve as the criterion for a dry pixel. Finally, isolated pixels are deleted, since they have a high chance of being erroneous. The result is a confidence map of the dry areas, computed as follows:

$$\tilde{I}(x,y) = \lambda_1 D(x,y) + \lambda_2 R(x,y) + \lambda_3 I_{n-1}(x,y) \tag{5}$$

where the parameters are at present chosen according to our empirical observations. In ongoing work, we are determining ways to choose them through learning
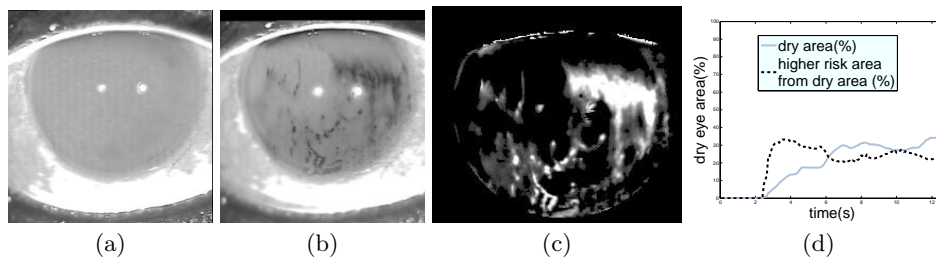
Fig. 6. *Dry eye area segmentation results (a) image after a blink (b) last image before the next blink (c) Confidence map of the dry area. Brighter shades represent drier areas (d) graph of the evolution of the dry area in the sequence. About 35 percent of the eye became dry, but from that only 22 percent is severely dry*

techniques. However, λ_3 should have a small value, since low intensity at the final frame does not always mean that the pixel became dry. The brighter the pixel is in $\tilde{I}(x, y)$, the drier it is likely to be. Darker pixels still exhibit some dryness, but might not be so substantial. The area of the dryness is calculated and we separate between higher risk and lower risk areas as seen in Fig. 6.

3 Results

We used 8 videos of different patients to evaluate our results. The videos are of size 355 x 282 pixels in RGB format. Each video usually had 3 to 5 good sequences that could be checked for the dry areas. The patients varied from having a very dry eye to an eye with no visible dryness. The videos were taken at different times with different illumination conditions.

Evaluating the robustness of our detection algorithm is a difficult task, since the area of dryness is subjective and cannot be exactly defined. This is the advantage of building the confidence map. In order to measure our performance, an optometrist was given the videos and hand segmented the dry areas over the last image (for reasons of convenience) in 12 sequences taken from 3 different patients with different degree of dryness (see Fig. 7(a) and (c)). We chose a fixed threshold for our confidence maps and used the following figure of merit:

$$S(\tilde{I}_x^t, I_x^h) = \frac{\sum_{x \in ROI} \tilde{I}_x^t == I_x^h}{|ROI|} \qquad (6)$$

where $\tilde{I}_x^t$ results from thresholding the confidence map, I_x^h is the hand segmented image and $|ROI|$ is the number of pixels in the iris. On average our method had 91% of accuracy, ranging from 84% to 96% and a standard deviation of 4%. Our algorithm found 61 out of the total 68 areas that were hand segmented for dryness, and in two images segmented an extra area. The main differences between the segmentations were in the exact boundaries of the dry areas and

the difficulty of our method to find thin areas next to the eyelids that were revealed only partly through the sequence. By choosing different thresholds for the confidence maps and changing the parameters in Eq. 5 the results can be adjusted to match better to specific hand segmentations.
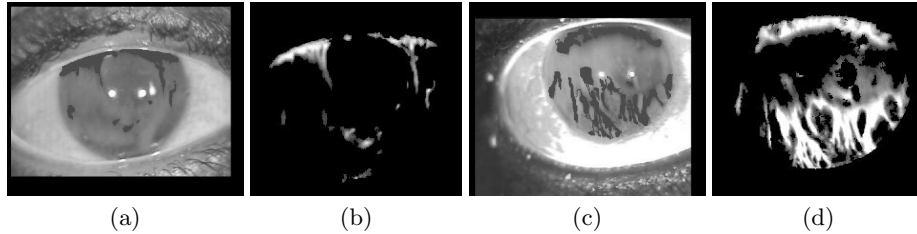

<div align="center">(a)   (b)   (c)   (d)</div>

Fig. 7. *Comparison of the dry eye area segmentation results. (a) & (c) Hand segmented areas for Fig. 2 and Fig. 5. (b) & (d) Our confidence maps for Fig. 2 and Fig. 5*

4   Conclusion and Future Research

In this paper, a new method for automatic detection of the tear film breakup area was developed. We demonstrated that the method can find the relevant area and size and build a confidence map.

This method will be tested on larger data sets and provide a comprehensive evaluation including the BUT value and confidence level for that value and a new standard to distinguish between the thinning of the tear film and the break based on the progression through the sequence. This should prove this method clinically useful by the optometrists. Our long term goal is to implement a non-invasive method to automatically detect the stability of the tear film.

References

1. John Daugman. The importance of being random: statistical principles of iris recognition. *Pattern Recognition*, 36(2):279–291, 2003.
2. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision.* Cambridge University Press, ISBN: 0521540518, second edition, 2004.
3. F. J. Holly. Formation and rupture of the tear film. *Exp Eye Res.*, 15:515–525, 1973.
4. D. Korb, J. Greiner, and J. Herman. Comparison of fluorescein break-up time measurement reproducibility using standard fluorescein strips versus the dry eye test (det) method. *Cornea*, 20(8):811–815, 2001.
5. M.A. Lemp and J.R. Hamill. Factors affecting tear film break up in normal eyes. *Arch Ophthalmol*, 89(2):103–105, 1973.
6. Gareth Loy and Alexander Zelinsky. A fast radial symmetry transform for detecting points of interest. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I*, pages 358–368, London, UK, 2002. Springer-Verlag.

7. Li Ma, Tieniu Tan, Yunhong Wang, and Dexin Zhang. Efficient iris recognition by characterizing key local variations. *IEEE Transactions on Image Processing*, 13(6):739–750, 2004.

8. M.S. Norn. Desiccation of the pre corneal film. *Acta Ophthalmol*, 47(4):865–880, 1969.

9. K. Tsubota. Tear dynamics and dry eye. *Progress in Retinal and Eye Research*, 17:556–596, 1998.