Robust Video Based Iris Segmentation System in Less Constrained Environments

Nitin K. Mahadeo, Andrew P. Papliński, Sid Ray
Clayton School of Information Technology
Monash University
Email: {Nitin.Mahadeo, Andrew.Paplinski, Sid.Ray}@monash.edu

Abstract—One of the key challenges in traditional iris recognition systems is that they require substantial user cooperation. Several restrictions are imposed on positioning and motion of the subject during the image acquisition process so that an image of high quality can be captured. On the other hand, videos captured at a distance and on the move are less intrusive and more appealing to users. However, this extra convenience comes at a cost. Such videos suffer from significant degradation and are often of poor quality compared to images captured in controlled environments. In this work, we present a video based iris segmentation system for processing of images taken in less constrained environments. In the first part, frame alignment of face videos is performed for reliable and efficient extraction of the eye regions in Near Infrared (NIR) face videos. In the second section, we propose a new iris segmentation method aimed particularly at eye images captured in challenging environments. Reflections and out of frame iris regions are in-painted. A region based segmentation method is proposed for accurate eyelid detection in images with variable illumination and significant blur. Eyelashes are divided into two categories and eliminated. Experiments carried out on the Multiple Biometric Grand Challenge (MBGC) dataset demonstrate that the proposed system achieves higher accuracy than other recent state of the art video based iris segmentation techniques developed for less constrained environments.

I. INTRODUCTION

The system presented in this paper is built from our new components combined, for completeness, with the current state of the art techniques. In the text that follows, we clearly emphasize our novel contribution to this relatively complete iris segmentation system that works in less constrained environments.

As a result of the numerous advantages and the extra layer of security that biometric technology offers compared to traditional identity and verification systems, there is considerable interest from both governments and industries in this technology. The iris biometric is the most accurate biometric to date and has several advantages over its counterparts. The iris, due to its inherent properties is well positioned and is an attractive alternative compared to other biometrics [1]. In addition, the iris is not subjected to wear and tear and it is well protected. On the other hand, this also makes the task of capturing a high quality iris image even more challenging. Long-established iris recognition systems require close proximity of the subject to the camera. The image acquisition process is time consuming and often requires prepositioning by a trained staff. According to a U.K. Passport Service (UKPS) report [2], a relatively high failure rate was reported amongst users of traditional iris recognition systems. Some of the main reasons behind this poor performance compared to other biometrics include positioning problems and the time taken for enrolment.

A. Motivation

To overcome the limitations of conventional iris recognition systems, Matey *et al.* [3] proposed a new video based iris image acquisition system for less constrained environments. The Iris on the Move (IOM), also referred to as "the portal" in the literature, is capable of capturing iris images of moving subjects with significantly less user cooperation in less acquisition time together with more capture volume compared with conventional systems. Subjects are merely required to walk at normal speed ($<$1m/s) through the portal as the proposed system is capable of capturing iris images of moving subjects from a distance. Eyeglasses and contact lenses are allowed. The IOM system was not built to capture high quality images but instead the focus was more on capturing images in an unconstrained manner with a trade-off in iris image resolution. Images acquired using traditional iris recognition systems have 200 pixels or more across the iris and are considered to be of reasonable quality while IOM eye images have only 100 pixels along the iris region and are considered to be of marginal quality [3].

Previous studies evaluating the performance of iris recognition systems have been carried out on still images captured by conventional iris image acquisition systems. Video based iris recognition systems is a fairly new area and presents new challenges. The Multiple Biometric Grand Challenge (MBGC) dataset is organized by the National Institute of Standards and Technology (NIST) and University of Notre Dame (UND). It consists of video sequences of moving subjects captured by the portal. The segmentation stage is a particularly difficult one due to the presence of unexpected and inconsistent noise effects present in these video streams. In addition, poor recognition performance is reported in the literature in video to video iris image comparisons compared to still image comparisons [4]. To put things in perspective, three different still eye images of a subject captured in monitored environments are shown in the top row of Figure 1 and three eye regions extracted from the same IOM NIR face video stream are presented in the bottom row [5], [6]. The quality and the types of noise present in eye images captured at a distance and on the move differ considerably from each other
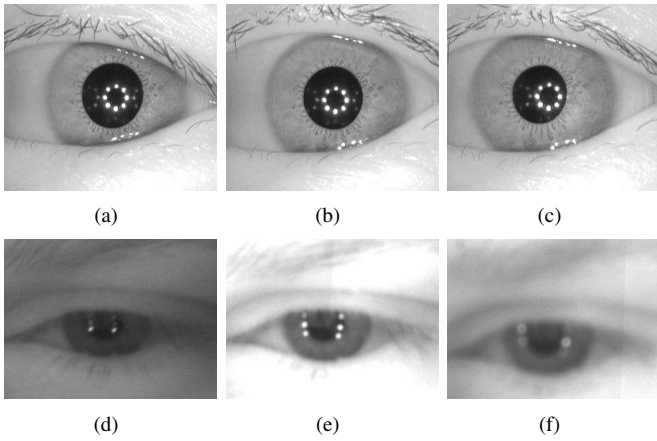
Fig. 1: Eye Image Quality : (a),(b),(c) are eye images captured in controlled environments-CASIA dataset [5]. (d),(e),(f) are eye images captured in less constrained environments-MBGC dataset [6].

and are significantly more challenging than iris still images captured using traditional iris recognition systems.

The focus of this work is therefore to develop a robust iris segmentation system capable of extracting reliable iris information from video sequences captured in unconstrained environments. The proposed system should be able to perform accurately on images with intermittent distortions and significant noise, which is vital to enhancing the performance of video based iris recognition systems. This paper is organized as follows. In Section II, we present an efficient and accurate way of extracting the iris region from face videos using alignment. In Section III, we propose a robust approach for eyelid segmentation in degraded images. Section IV includes our discussion and finally in Section V and Section VI, our results and conclusions are presented.

II. FRAME ALIGNMENT AND EYE IMAGE EXTRACTION

Extraction of the iris region in NIR face videos can be a challenging task due to low resolution and poor quality of eye images. A new model for eye image extraction in NIR face videos is proposed. Section II-A to Section II-D below describe each of the stages of the proposed approach. This is illustrated in the flowchart in Figure 2.

A. The Reference Frame

In [7], the eye regions in each frame are detected using thresholding. However, such an exhaustive frame by frame search for the iris region is clearly inefficient. Thresholding can also lead to poor detection due to the presence of artifacts such as jewellery and other reflective surfaces present in the image. In [8] and [9], the Viola Jones object detector which is based on the Adaboost framework is initially used for eye pair detection in the NIR face images [10]. In this work, the eye region is also detected in a similar manner. The following parameters are used in our experiments. The minimum size of the cascade object detector is set to 150×100 and the
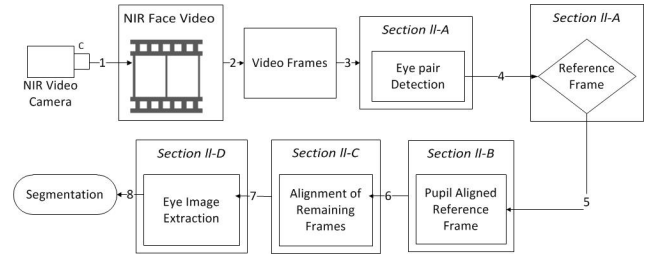

Fig. 2: Stages of eye image extraction in NIR face videos

dimension of the classifier is 45×11. In contrast with the work mentioned above where an eye pair is searched using the Viola Jones detector in each frame of the NIR Face Video, here the following undertaking is implemented to minimize false detection and improve efficiency. Our initial work on a subset of the MBGC dataset indicates that the radius of the left and the right pupil share an almost linear correspondence. Figure 3 is a plot of the radius of the right pupil versus that of the left pupil of hundred correctly detected eye pairs of hundred different subjects. A linear relationship can be observed between the radius of the right pupil and that of the left pupil. The slope of the linear fit is found to be 0.98 indicating an almost one to one correspondence between the two variables. Therefore, based on the above heuristic, an eye pair detected by Viola Jones object detector is considered to be genuine only if $|r_p - r_l| < 5$ pixels and $r_p, r_l > 10$ pixels where r_p and r_l are the radii of the right and left pupils. Taking into consideration that both eye pair detection and pupil localization on each frame can be computationally slow, detection is only performed until a genuine eye pair is identified. The frame belonging to the genuine eye pair is then selected as the reference frame, f_r and its pupil parameters are recorded to be used in the subsequent stages.

B. Pupil Alignment in the Reference Frame

Videos streams captured by the IOM image acquisition system contain significant head movement, partially occluded and out of frame faces. We first proceed by aligning both the right and the left eye in frame, f_r with the horizontal. This is
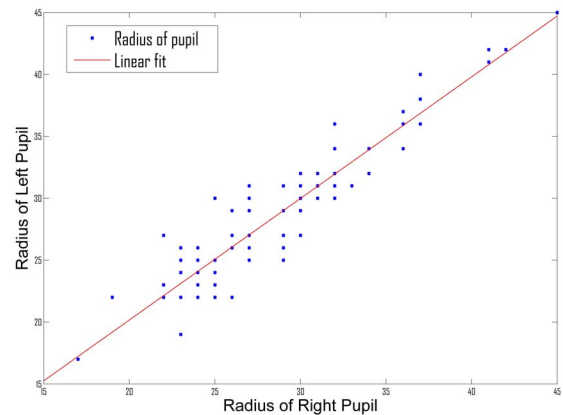

Fig. 3: Plot of the radii of right vs. left pupil in eye pairs

(a)   (b)

Fig. 4: Pupil Alignment: (a) is the reference frame, f_r (b) is the pupil aligned reference frame, f_a.

a vital step towards facilitating accurate iris region extraction in the remaining frames in addition to making eyelid detection easier. Successful detection of a genuine eye pair in a frame of a given video stream using the proposed method in Section II-A also entails computing the pupil parameters. The angle between the centre of the right and the left pupils can therefore be computed. The reference frame is subsequently rotated until both pupils are aligned with the horizontal axis. This is shown in Figure 4. The angle between the dotted red line and the horizontal axis in the reference frame is determined using the spatial centre coordinates of the pupils. The reference frame is then rotated about the centre point between the two pupils by the same angle. This output image is shown in the Figure 4(b). The red dotted line now overlaps the horizontal axis and both pupils are now aligned with the horizontal. This rotated reference frame, f_r is referred to as the aligned pupil reference frame, f_a in a given video. The technique used for pupil localization is discussed in Section III.

C. Frame Alignment

In this section, we explain how the remaining frames in a given video are aligned with respect to frame f_a. Consider an
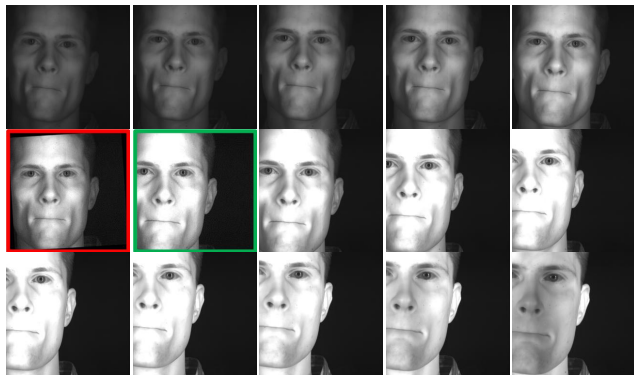

Fig. 5: Sequential Frames from a given IOM video stream where the red bordered frame is the aligned reference frame and the frame with the green border is next one to be aligned.

(a)   (b)

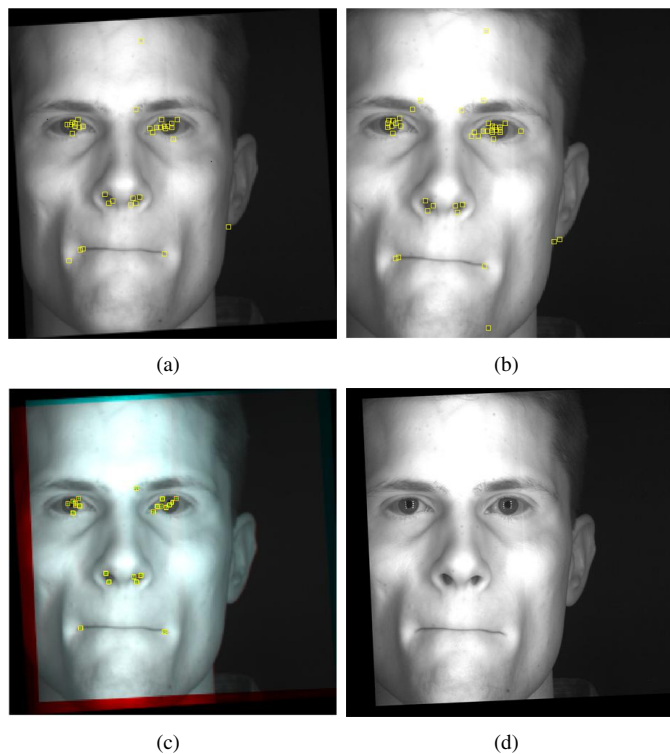(c)   (d)

Fig. 6: Estimate of distortion between 2 adjacent frames: (a) and (b) are the features detected in 2 adjacent frames by FAST (c) shows the matched features and (d) is the geometric transformation of frame (b).

NIR video of a given subject consisting of 26 frames. Only the first 15 frames, f_1 to f_{15} are shown in Figure 5. The remaining frames, f_{16} to f_{26} being too dark are unusable and are not shown here. Here, the aligned reference frame f_a for this particular video is frame f_6 (red box). For now, we will be examining only the alignment of frames to the right of frame f_6. Therefore, the next frame which needs to be aligned is frame f_7 (green box) and the amount of distortion between these two adjacent frames (red box and green box) needs to be estimated.

We proceed by finding the salient features present in both frames of relevance using the FAST (Features from Accelerated Segment Test) corner detector algorithm [11]. This rapid corner detector being able to produce stable points of interest makes it a suitable choice for frame alignment. For illustration purposes, we demonstrate a step by step approach of the distortion estimation between f_6/f_a and f_7 in Figure 6. The salient points detected in both the the aligned reference frame, f_a and its right adjacent frame f_7 have been identified. In the next step, correspondence between the feature points in frames f_a and f_7 is estimated. A 9×9 block centred around each point is extracted and the matching cost between points of the concerned region is computed by finding the sum of the squared differences. The affine transform between the two adjacent frames is then be estimated using RANSAC algorithm [12]. This is shown in Figure 6(c) where the
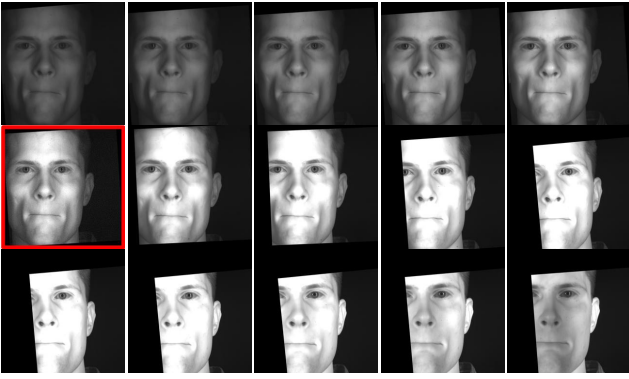
Fig. 7: Alignment of the remaining frames w.r.t the aligned pupil reference frame. The eye regions are properly aligned with each other thereby facilitating the eye region extraction.
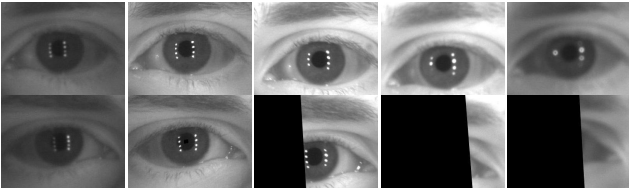

Fig. 8: Right (top row) and left (bottom row) eye regions from frames 3, 6, 9, 12 and 15 in Figure 7.

transform with the minimum cost is chosen and the affine transform is estimated. The above step is implemented using the *cvexEstStabilizationTform* function in Matlab [13]. Figure 6(d) shows the aligned frame f_7.

Estimation of the amount of distortion, D in the remaining frames to the right of frame f_a is computed the following way. As shown in Figure 6, the distortion between two consecutive frames f_i and f_{i+1} can be estimated as an affine transform. Therefore, the collective distortion of a frame is equivalent to the product of the previous inter-frame transforms. This is shown in equation (1) where, $D_{c,i}$ is the collective distortion from frame f_a up to frame f_i where i is the frame index and $i > a$.

$$D_{c,i} = \prod_{j=a}^{i-1} D_i \tag{1}$$

For more information on the implementation of the above, we refer the reader to [13]–[15]. Note that the same procedure applies for frames on the left of frame, f_a. The only change is the frame number, i, which is computed from right to left in this case. The final output of the frame alignment is shown in Figure 7.

D. Extraction of the eye region

The result of the proposed approach using alignment is that the eye regions are now all roughly horizontally positioned at the same location in the video frames. In addition, video sequences with poor visibility have been eliminated. Therefore the iris region can be extracted more reliably and efficiently, even in cases where only one eye is present as opposed to the

use of only an eye pair detector for image extraction in [4], [8]. Figure 8 shows the extracted left and right eye regions from frames, $f_{3,6,..15}$ in Figure 7. All eye regions have the same dimensions and are extracted with respect to the pupil-centre coordinates in frame, f_a.

III. EYELID SEGMENTATION

This work builds upon the model based technique, for pupil and iris localization in severely degraded images, proposed by the authors in [16]. The spatial coordinates of the pupil and iris are obtained by maximizing the convolution of the eye image with a toroidal 2-D filtering shape derived from the 1-D Petrou-Kittler [17]. Figure 10(b) is an example where the pupil and the iris have been successfully detected.

In this remaining section, our main focus is on developing a robust and fast eyelid localization technique for eye images captured in unconstrained environments. As mentioned in [8], [18], [19], eyelid localization is particularly challenging. Iris images captured in such challenging environments suffer from significant degradation making transition between the iris and eyelid region less distinct. The eyelid region is also often covered by eyelashes and the shape of the eyelids varies from one individual to another. This makes the task of localizing the eyelids even more difficult.

A. Preprocessing

Due to poor image quality, image enhancement is performed prior to eyelid detection. The following preprocessing techniques were implemented in this study.

The first image enhancement step proposed is the inpainting of the iris region in images where insufficient information is available. This is shown in Figure 9. Transition from the out of frame eye region to that of the valid eye section is a sharp one and may lead to incorrect detection. In [7], out of range region is filled with pixels belonging to the last column of the boundary. This can however lead the edges being extended to other parts of the image as shown in Figure 9(b). Here, a smoother passage is established between the two regions by propagating pixels at the inner boundary of the valid iris region into the out of frame region using a sparse based PDE inpainting algorithm [20]. As shown in Figure 9(c), the change from the iris region to that of the out of frame region is much smoother.

The aim of the second image quality enhancement proposed in this work is to increase the contrast between the iris region
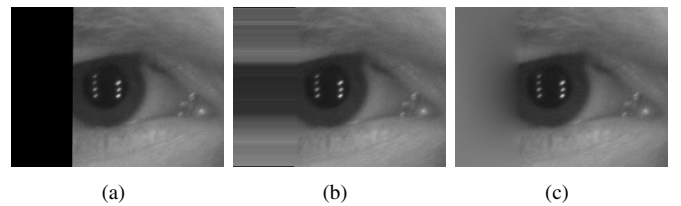

Fig. 9: Out of frame eye region:(a) Original Eye Region (b) Pixels in the last column extended (c) Inpainting

and that of the eyelid. Poor lighting and variable illumination, distinction between the eyelid region and that of the iris is often ambiguous. An example of such an eye image is shown in Figure 10(a). To address this problem, we first proceed by computing mean intensity in regions T, B, R, L in Figure 10(b). Here, T and B are 10×10 blocks at the topmost and bottommost of the detected iris circle while R and L are 10×10 patches at are the rightmost and leftmost of the pupil boundary. Both the top and the bottom regions are considered to compensate for illumination variation which may be present in the eye image. Similarly, the left and the right regions are considered in the iris as these patches are usually not occluded. Therefore, the mean intensity outside the iris region, μ_o is $(\mu_T + \mu_B)/2$ and the mean intensity inside the iris region, μ_i is $(\mu_R + \mu_L)/2$. In Figure 10(b), the input intensity values of μ_i is found to be 0.4 and that of μ_o is 0.6. The input intensity values are linearly mapped to output intensity values 0 and 1. Values lower μ_i are saturated to 0 while values higher than μ_o are saturated to 1. The output image is shown in Figure 10(c). This approach makes the distinction between the iris border and that of the eyelid much clearer thereby facilitating the detection process. Moreover, this process is an adaptive one i.e., the intensity remapping is done based on the intensity values detected in the four blocks.

B. Eyelid Detection using Region Growing

In this work, region growing was adopted for fast and accurate eyelid detection. This region based procedure works as follows. The region iteratively grows by comparing and appending unassigned neighbouring pixels to the initialized seed points based on a predefined criteria. The difference between the mean of the region and that of pixel intensity value is the criterion which determines if it is added to the growing region. The procedure stops when the difference between the pixel intensity value and the region's mean exceeds a given threshold [21]. This procedure is suitable for eyelid detection since a priori information is available. From the pupil and iris localization technique in [16], the centre coordinates and radius of the iris are known. Therefore, initialization of the seed points in the iris region is straightforward. In addition, the iris is also a connected region and the intensity value can be used as the stopping rule since the pixel intensity values do not vary significantly within the enhanced iris region. A threshold of 0.1 is used in our experiments.

To compensate for illumination variations which may be present on the image, two seed points are initialized close to the rightmost and leftmost part of the iris border as indicated by S_1 and S_2 in Figure 10(c). This region is usually unaffected by crypts and collarettes which are usually found close to the pupil boundary and may inhibit the region from growing. The circumference of the iris is set to pixel intensity values of 1. This acts as a control preventing the region from flowing into the scleric region. The final binary image is the combined sum of the left and right regions grown from each seed initialized. Morphological operations are used to close any holes which may be present in the final binary image. The output binary
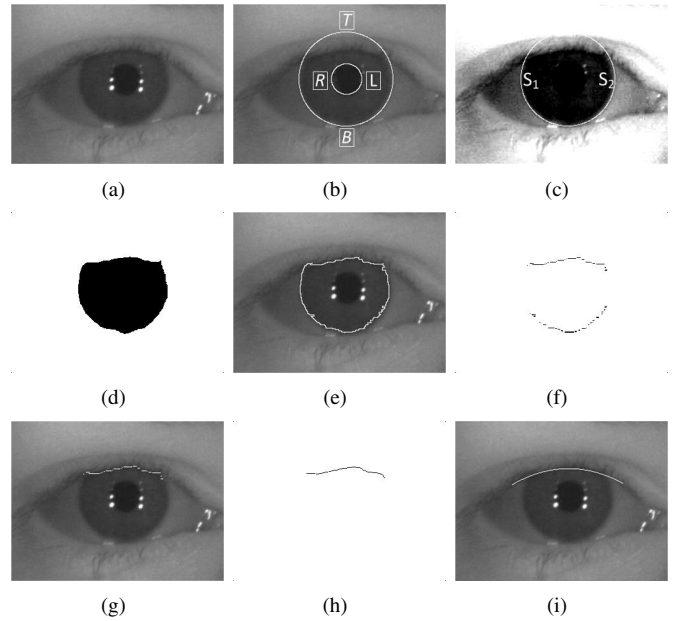

Fig. 10: Stages of Upper Eyelid Detection: (a) Original Eye Region (b) Pupil and iris localization (c) Enhanced eye image (d) Binary output image of region growing (e) Boundary of detected iris (f) Upper and lower eyelid detected (g) Upper eyelid genuine edges (h) Edge linking (i) Curve fitting

image of region growing is shown in Figure 10(d).

C. Eyelashes

Eyelashes usually occur predominantly on the upper eyelid region. They are usually vertically aligned and have lower intensity values than the iris. The proposed enhancements and the region based approach worked well on images without eyelashes. However, in images with prominent eyelashes, the image enhancement technique also caused the eyelash to be considered as part of the iris region since they have lower intensities. In order to deal with this issue, eyelashes on the eyelid region are divided into two categories, namely thin and thick eyelashes.

Thin Eyelashes

Thin eyelashes are typically vertically aligned and separated from each other. From the alignment process in Section II-B, the eyelid boundaries are horizontally aligned while the eyelashes are vertically oriented. This facilitates the eyelash detection process. In order to eliminate the influence of such eyelashes, vertical edges in the upper eyelid boundary detected by region growing are eliminated. As shown in Figure 10(e), the left and right sections of the eyelid boundary detected by region growing are influenced by the presence of eyelashes and is not very accurate. In order to reduce the effect of thin eyelashes present on the eyelid region, vertical edges present in the boundary of the region growing are eliminated by applying a morphological operation, namely erosion using a rectangular structuring element of size 1×3 on the iris

boundary detected in Figure 10(e). The output of the operation is shown in Figure 10(f). Note how the vertical edges caused by eyelashes are no longer present. In addition, elimination of the vertically aligned edges result in both the upper and lower boundaries to be clearly separated from each other. These edges are considered as genuine, i.e., they belong to the eyelid boundary. These detected upper eyelid edges are shown in Figure 10(g). In the next step, connectivity between the edges is established. The horizontal eyelid edges are linked to each other by dilation followed by a thinning operation. The edges are now connected horizontally along with the upper eyelid region as shown in Figure 10(h). Finally, the arc shape of the eyelid is estimated by fitting a second degree polynomial to the resulting connected edge as illustrated in Figure 10(i). It can be observed that the final result is very accurate although eyelashes and other types of noise are present in the image. It is also a more accurate fit than Figure 10(e) where only region growing is implemented. This approach is very fast and computationally inexpensive.

Thick Eyelashes

Contrary to thin eyelashes, thick eyelashes are often bundled together and longer. Occurrence of such eyelashes on the eyelid region cannot be eliminated in a similar manner to thin eyelashes as they can produce false edges which appear horizontal and mislead the proposed technique into believing that it is a genuine edge. The following approach is adopted to reduce the effect of thick eyelashes in eye images. As eyelashes are significantly darker than the eyelid region, this also suggests that the intensity difference between these two regions of interest would be significantly higher. Therefore, by applying a local range filter on the image, thick eyelashes present in the image can be detected. As its name suggests, a local range filter computes the range value in the neighbourhood of a given pixel in the eye image. A 7-by-3 neighbourhood was found to be optimal for the detection of thick eyelashes in eye images. Figure 11 illustrates the steps implemented for segmenting an eye image with thick eyelashes. Figure 11(b) shows the binary output image of the range filtering. As can be seen in Figure 11(b), most of the thick eyelashes have been detected. In the next stage, comparably to thin eyelash detection, the image is enhanced and region growing is applied on the iris region of the image. This enhanced image is shown in Figure 11(c). The binary output image of the region based method is shown in Figure 11(d). It can be observed that the region has grown into the eyelash region on the eyelid. The effect of eyelashes is eliminated using the following approach. A logical exclusive-OR operation is performed between Figure 11(b) (thick eyelashes) and Figure 11(d) (binary output image of region growing). The output of this logical operation is shown in Figure 11(e). The next steps are similar to the technique used for detection of thin eyelashes. The vertical edges are eliminated as shown in Figure 11(f) followed by horizontal edge linking and curve fitting.
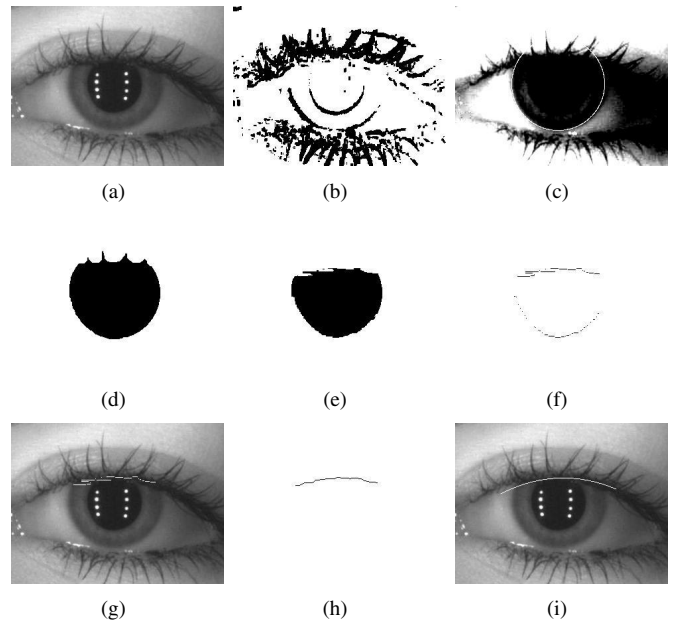

Fig. 11: Upper Eyelid Detection:(a) Original Eye Region (b) Thick eyelash detection (c) Image enhancement (d) Binary output image of region growing (e) Elimination of eyelashes (f) Erosion of vertical edges (g) Genuine upper eyelid edges (h) Edge linking (i) Curve fitting

IV. DISCUSSION ON THE PROPOSED SYSTEM

A. Alignment

In [7], eye regions in face videos are detected using thresholding. The distance between detected reflections in the binary image is then computed to determine if it belongs to the eye region. However, due to the presence of unexpected noise effects such as jewellery and other artifacts present in the frame, this is not always optimal. In [4], [8], [9], the occurrence of an eye pair in each frame is carried out using the Viola Jones object detector. This approach is liable to false detection due to poor quality of images and can be computationally inefficient. Also, this approach does not consider frames where only one eye is visible. The proposed alignment method has following advantages. Firstly, verification of the correspondence between the left and right pupil radii in the reference frame, f_r in Section II-A prior to extraction significantly reduces the possibility of false detection. Secondly, instead of frame by frame detection, in our experiments detection is performed only until a genuine eye pair is detected for each video stream. Since, the remaining eye regions are positioned in the same locations in the frames after alignment, they can be extracted more accurately and efficiently. Thirdly, as the subject moves towards the camera, more of the face region is exposed. Consequently, the size of the eye pair detected using the Viola Jones object detector vary significantly from one frame to another [8]. On the other hand, the affine transform being an s-r-t(scaling-rotation-translation) transformation ensures the extracted eye regions have the same dimensions in our experiments. Fourthly, due

to user movement, there are frames in which only one eye is completely visible but the quality is sufficient enough for iris recognition. This is shown in frame, f_{10} in Figure 5 where only one eye is present but the frame quality is sufficient. Alignment makes it possible for this frame to be extracted and used in our experiments while such frames would be discarded in [4], [8] where eye pair detection only is used. Finally, alignment also makes it possible to differentiate between the right eye and the left eye in images where only one eye is visible.

B. Segmentation

It should be noted that the image enhancement technique implemented significantly improved the performance of the eyelid detection by increasing the contrast between the iris and the eyelid regions. Thus, the region growing procedure was also able to correctly identify the eyelid border in a reliable and efficient manner. In addition, horizontal alignment of the pupils in the videos corrected images with tilt or head movement facilitating the distinction between the vertically aligned thin eyelashes and the horizontal eyelid boundaries. The same technique applies for lower eyelid detection. The proposed eyelid segmentation is also compared with two other popular iris segmentation techniques in the literature. In [1], Daugman extended the integro differential operator to an arc shape to search for the eyelid and in [22] Masek implemented an open source version Wilde's segmentation technique using Canny edge detection and Hough transform. As shown in Figure 12, the proposed detection technique is able to detect the eyelid boundary more accurately while valid iris information is lost in the case of Masek's method and the maximum of the convolution operation in Daugman's algorithm detects the region above the upper eyelid.

Some examples of the proposed iris segmentation technique are shown in Figure 13 in images with variable illumination, thick eyelashes, reflections and subjects wearing glasses. Images of failed iris segmentation are shown in Figure 14. These images contain significant amount of unexpected noise effects and blur.

V. Experimental results

In this study, the MBGC V1.0 dataset is used [6]. It consists of 149 NIR Face videos of 114 subjects captured at a distance and on the move by the IOM system [3]. The videos captured are of resolution 2048×2048 and the frame rate is 15 fps. The portal was designed with the aim of capturing at least
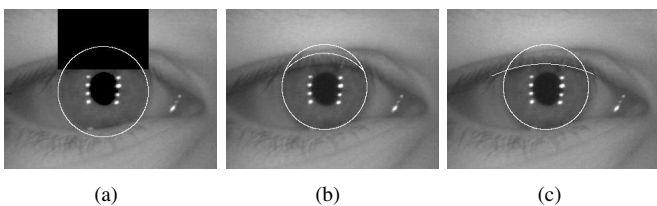

Fig. 12: Upper Eyelid Detection comparisons with (a) Masek's technique (b) Daugman's method (c) Proposed
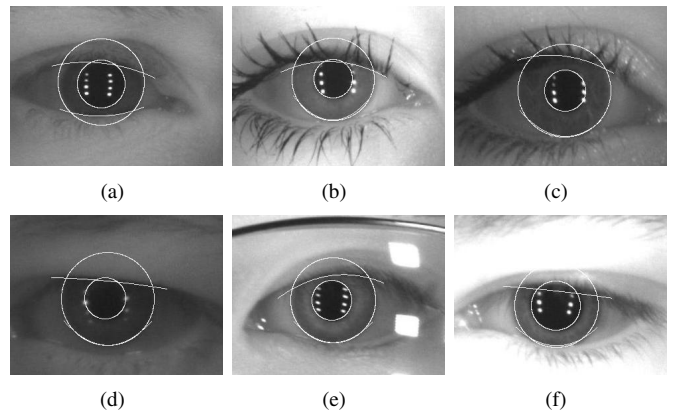

Fig. 13: Examples of successfully segmented iris regions in eye images captured in less constrained environments.
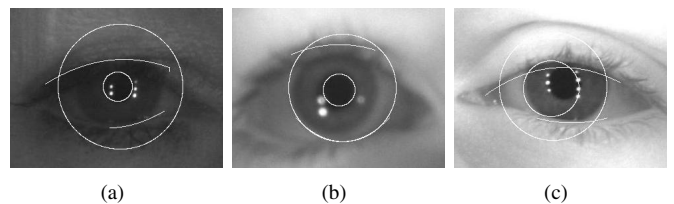

Fig. 14: Examples of failed iris segmentation in eye images with (a) poor illumination (b) significant blur (c) ellipse shaped pupil

one image in focus. On average, each video consists of about 20 frames and only 1-2 images in each video were found to be useful for iris recognition. The quality of the remaining eye images were found to be too dark, blurred, out of focus and do not contain sufficient iris information. The extracted eye region in our experiments is of size 240×320 as shown in Figure 8. The proposed system was implemented in Matlab [13]. The average alignment rate is 0.5 seconds per frame. The mean segmentation time is 3.7 seconds i.e., 2.6 seconds for pupil and iris localization and 1.1 seconds for the proposed eyelid segmentation technique. Our results are compared with other popular techniques developed specifically for iris segmentation of eye images captured in less constrained environments. The same dataset is used for comparison purposes. In [7], iris images are extracted by measuring the distance between reflections in the binary image obtained by thresholding. Using this approach, the number of images extracted will vary depending on how strict the settings are. Our results are shown in Table I. The proposed method in this work achieves

Method	Eye Images Extracted	Mean Eye Images per Subject	Accuracy (%)
Chen *et al.* [7]	657	4.4	94.2
Proposed	776	5.2	94.4

TABLE I: Segmetation accuracy w.r.t the number of eye images extracted from IOM video streams

Method	Eye Images Extracted	Accuracy(%)
VASIR(2009) [4]	254	65.9
VASIR(2013) [8]	252	81.7
Proposed	252	93.7

TABLE II: Segmentation accuracy using the HBIS approach

comparable performance with an even higher number of eye images.

We also compare our work with the segmentation techniques proposed in [4], [8]. Here, the best eye image from the left and right eye in each video stream is selected by HBIS (Human Best Image Selection) i.e., the best image quality in this case is selected by human visual inspection. Eye images are selected in a similar fashion in our experiments. The results are shown in Table II and a considerable improvement can be observed using the proposed iris segmentation system. It should be highlighted that this image selection process is challenging as even the best quality eye image in some videos is extremely poor. IrisBEE whose underlying algorithm is based on Masek's only achieves only 48.6% accuracy on the MBGC dataset and is not included in the Table II [8], [23].

The above results demonstrate that performance was enhanced on two fronts. The results in Table I where eye images are solely detected using reflections show that our segmentation algorithm is robust. Table II indicates that both a robust segmentation algorithm and an efficient eye image extraction is required to enhance performance. This is due to the fact that the number of eye images available as a result of the alignment process is significantly higher using the proposed approach.

VI. CONCLUSION

In this paper, a new system for the extraction of eye images in NIR face videos is presented. We have shown that eye images can be extracted in a more efficient and accurate manner by exploiting spatial correspondence between the left and the right pupils. The amount of distortion in the remaining frames is then estimated by matching features in adjacent frames. A novel eyelid detection technique is also proposed. Image enhancement and a region growing are implemented for accurate detection of the eyelid borders. Both thin and thick eyelashes are correctly detected and eliminated. As shown by our results in Table I and Table II, the proposed iris segmentation system performs significantly better than previous systems developed for NIR face videos captured in less constrained environments.

ACKNOWLEDGEMENT

REFERENCES

[1] J. Daugman, "How iris recognition works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, pp. 21–30, 2002.

[2] "U.K. Passport Service Biometric Enrolment Trial," *Biometric Technology Today*, vol. 13, no. 7, pp. 6 – 7, 2005. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0969476505703684

[3] J. Matey, O. Naroditsky, K. Hanna, R. Kolczynski, D. LoIacono, S. Mangru, M. Tinker, T. Zappia, and W.-Y. Zhao, "Iris on the move: Acquisition of images for iris recognition in less constrained environments," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1936–1947, 2006.

[4] Y. Lee, R. Micheals, and P. Phillips, "Improvements in video-based automated system for iris recognition (vasir)," in *Motion and Video Computing, 2009. WMVC '09. Workshop on*, 2009, pp. 1–8.

[5] "CASIA Iris Database V4," http://www.cbsr.ia.ac.cn/china/Iris

[6] "NIST: Multiple Biometric Grand Challenge (MBGC) V1.0," http://face.nist.gov/mbgc, Jul. 2009.

[7] Y. Chen, J. Wang, C. Han, L. Wang, and M. Adjouadi, "A robust segmentation approach to iris recognition based on video," in *Applied Imagery Pattern Recognition Workshop, 2008. AIPR '08. 37th IEEE*, 2008, pp. 1–8.

[8] Y. Lee, R. Micheals, J. Filliben, and P. Phillips, "Vasir: An open-source research platform for advanced iris recognition technologies," *Journal of Research of the National Institute of Standards and Technology*, vol. 118, 2013.

[9] K. Nguyen, C. Fookes, S. Sridharan, and S. Denman, "Quality-driven super-resolution for less constrained iris recognition at a distance and on the move," *Information Forensics and Security, IEEE Transactions on*, vol. 6, no. 4, pp. 1248–1258, 2011.

[10] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 1, 2001, pp. I–511–I–518 vol.1.

[11] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2, 2005, pp. 1508–1515 Vol. 2.

[12] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed. New York, NY, USA: Cambridge University Press, 2003.

[13] "Matlab," Natick, Massachusetts, US.

[14] B. Tordoff and D. W. Murray, "Guided sampling and consensus for motion estimation," in *Proceedings of the 7th European Conference on Computer Vision-Part I*, ser. ECCV '02. London, UK, UK: Springer-Verlag, 2002, pp. 82–98. [Online]. Available: http://dl.acm.org/citation.cfm?id=645315.649190

[15] K.-Y. Lee, Y.-Y. Chuang, B.-Y. Chen, and M. Ouhyoung, "Video stabilization using robust feature trajectories," in *Computer Vision, 2009 IEEE 12th International Conference on*, 2009, pp. 1397–1404.

[16] N. Mahadeo, A. Paplinski, and S. Ray, "Model-based pupil and iris localization," in *Neural Networks (IJCNN), The 2012 International Joint Conference on*, June 2012, pp. 1 –7.

[17] A. Paplinski, "Directional filtering in edge detection," *Image Processing, IEEE Transactions on*, vol. 7, no. 4, pp. 611–615, April 1998.

[18] Z. He, T. Tan, Z. Sun, and X. Qiu, "Robust eyelid, eyelash and shadow localization for iris recognition," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, oct. 2008, pp. 265 –268.

[19] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Comput. Vis. Image Underst.*, vol. 110, no. 2, pp. 281–307, May 2008. [Online]. Available: http://dx.doi.org/10.1016/j.cviu.2007.08.005

[20] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, ser. SIGGRAPH '00. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 2000, pp. 417–424.

[21] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2003.

[22] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, University of Western Australia, 2003.

[23] P. Phillips, K. Bowyer, P. Flynn, X. Liu, and W. Scruggs, "The iris challenge evaluation 2005," in *Biometrics: Theory, Applications and Systems, 2008. BTAS 2008. 2nd IEEE International Conference on*, 2008, pp. 1–8.