Accurate Eye Localization in the Short Waved Infrared Spectrum through Summation Range Filters

Cameron Whitelam, Thirimachos Bourlai

*MILab, LCSEE, West Virginia University, 395 Evansdale Drive, Morgantown, WV 26506-6070, U.S.A.
(304)293-4326*

Abstract

The majority of facial recognition systems depend on the correct location of both the left and right eye centers in an effort to geometrically normalize face images. We propose a novel eye detection algorithm that efficiently locates the eye centers in five different bands of the SWIR spectrum, ranging from 1150nm up to 1550nm in increments of 100nm. Our eye detection methodology utilizes a combination of algorithmic steps, including 2D normalized correlation coefficients as well as summation range filters to effectively find the eyes in the aforementioned SWIR wavelengths. We validate our method by comparing our approach with currently available eye detection algorithms including a commercial face recognition software in which one of its capabilities is the extraction of the eye locations and a state of the art academic approach. Eye detection results as well as face recognition studies show that our proposed approach outperforms all other approaches, including the state of the art (originally designed to work in the visible band), when operating in the SWIR spectrum. We also show that our approach is robust to typical image degradation factors including spatial resolution changes, image compression, and image blurring. This is an important achievement that has also practical value for biometric operators. It is impractical to manually annotate thousands to millions of eye centers, therefore, a quick and robust method for automatically determining the eye center locations is needed.

Keywords:

Eye Detection, Face Detection, Short Waved Infrared Face Images, Face Recognition

1. Introduction

The past decade has seen significant progress in the field of automated face recognition as is borne out by results of the

2010 Multiple-Biometric Evaluation (MBE) organized by NIST [1]. For example, at a false match rate (FMR) of 0.001%, the false-nonmatch rate (FNMR) of the best performing face recognition (FR) system has decreased from 79% in 1993 to 0.003% in 2010. However, the problem of matching facial images remains to be a challenge when it is required to deal with diverse scenar-

Email addresses: cwhitela@gmail.com
(Cameron Whitelam),
Thirimachos.Bourlai@mail.wvu.edu
(Thirimachos Bourlai)

Preprint submitted to Elsevier

April 30, 2015

Figure 1: The electromagnetic spectrum shows the research gap in SWIR and sample images from a) 1150nm b) 1250nm c) 1350nm d) 1450nm and e) 1550nm. Notice how eye centers from both academic (green and blue) and commercial (red) algorithms are inaccurate and produce normalized face images with both rotational and scale problems (bottom row).

ios. Typical cases include the usage of different sensors to capture face images (e.g. 2D, 3D, visible, IR), large datasets, as well as variations in pose, illumination and expression. However, there are inherent problems that come along with face recognition in the visible band. The varying levels and nature of illumination is among visible FR’s most insidious problems. Recent FR trends are pointing to interest in higher spectra, i.e. Short Waved Infrared (SWIR) [2], Mid-Waved Infrared (MWIR) [4][3] and Long-Waved Infrared (LWIR) [6]. With the prices of these camera sensors decreasing rapidly, the practical benefit of using different spec-

tra for automatic FR is increasing. The potential to create more challenging (larger scale, different sensors) databases is becoming more of a reality for testing and developing face recognition algorithms with hopes of accurately matching visible gallery images to SWIR, MWIR, or LWIR probe images (namely the heterogeneous problem in FR, also known as cross-spectral matching).

In this paper, the focus of our work is in various wavelengths of the SWIR band. The SWIR spectrum has several advantages over other bands. It allows for covert capture in nighttime environments due to the fact that the illumination source cannot be seen (con-

sidering that its wavelengths are well above the visible spectrum). Also, certain pairs of sunglasses do not reflect SWIR wavelengths as they do visible light, allowing for both eyes to be seen unobtrusively. Because of the higher wavelength in the SWIR band, it becomes more tolerant to low level obscuring. This allows for imaging of scenes or faces through certain environmental conditions such as fog, smoke, etc. Finally, because facial features that are not observed in the visible spectra can be observable in the SWIR band, these two bands can be combined to generate a more complete image of the human face. This can allow for more facial features to be revealed and exploited and, as a result, FR performance can improve.

1.1. Motivation

In a typical FR system, one of the main challenges that must be overcome in order to achieve high FR identification rates is to successfully determine the face and eye locations of all face images in a database that are used for matching. Eye detection, in particular, is considered to be one of the fundamental steps in the majority of FR algorithms. With each passing year, the facial databases are enlarging to a point where millions of individuals need to be enrolled for research purposes but also to test the scalability of existing FR matching algorithms in practical scenarios. In general, commercial and academic facial recognition algorithms require that the face images of each individual be standardized (in terms of orientation, inter-ocular distance, masking etc.). Typically, feature points of sub-facial regions, more specifically the locations of human eye centers, are used to rotate and translate a face to a standard representation. While this operation can be man-

ually performed by an operator on a limited size dataset, when having to deal with larger databases (from thousands to millions of subjects), manually obtaining the eye locations of each subject and for each sample per subject available is not practical. Therefore, an accurate and robust eye detection method needs to be employed.

Reported work in the literature shows that automated and robust eye localization can have a positive impact on FR performance [7]. However, while available eye detection approaches perform well in the visible band, they perform poorly in the SWIR spectrum. This happens mainly due to the fact that these algorithms were not designed to work well in other bands than the visible. Also, the lack of publicly available SWIR face datasets is another challenge that needs to be overcome in the future. The eye locations found when using three well known eye detection methods, i.e. both commercial and academic (Viola and Jones and state of the art [8]) on our hyper-spectral SWIR face dataset (face images are captured in multiple wavelengths within the SWIR band) do not perform well. More importantly, when face images in a dataset are geometrically normalized based on the found eye locations, the resultant image cannot be used by face recognition algorithms due to the effect that rotations and scale have on face images. This problem is illustrated in Figure 1. Therefore, an efficient eye localization algorithm needs to be designed, developed, and applied on both the gallery (conditioning of enrolled individuals) and probe wavelength specific SWIR datasets.

1.2. Proposed Approach and Evaluation Process

In this paper, we propose a fast and efficient eye detection and localization algo-

algorithm that accurately determines the centers of the eyes across the SWIR band (i.e. 1150, 1250, 1350, 1450, 1550nm). The effect of the proposed method on FR performance is compared against normalized face images generated by manually annotated eye centers. Experimental results demonstrate that our methods are better than all tested algorithms across all datasets used. One point of note is that the authors recognize that in the SWIR datasets used, the number of samples and subjects is relatively small due to the fact that there are very few publicly available SWIR databases. It is also very difficult to perform a new SWIR data collection due to university data collection policies, cost, time, etc. Because of this limitation, we only managed to use the data that was collected in one of our recent projects.

The rest of this paper is organized as follows. Section 2 describes previous work done in the field of automatic eye detection while Section 3 describes the methodology of the proposed approach. Section 4 extensively validates our algorithm while drawn conclusions and future work is described in Section 5.

2. Background

In the field of eye detection, there are three main methodologies:

- *Feature Based methods* [8][14][19]: such methods use a variety of eye characteristics to detect the eyes. To identify distinctive features of the eyes, characteristics such as the color distribution of the sclera, the flesh of the eye, and the edge and intensity of the iris are used. Although these methods are typically efficient, they are not as accurate as other eye detection methods

(appearance and template based methods), especially for low contrast images.

- *Appearance Based methods* [7][25][27]: in such methods, eye detection is considered the problem of classifying each scanned subwindow as one of two classes (i.e., eye and non-eye). Appearance-based methods avoid difficulties in modeling 3D (depth) structure of the human eye by focusing on possible eye appearances under various conditions. In these methods, a classifier that determines an eye vs. a non-eye location may be learned from a training set composed of eye and non eye sample images. Different classification schemes, such as the AdaBoost algorithm, can be used to train the eye detector.
- *Template Based methods* [9][12][20][22]: such methods use a generic eye model, typically based on eye shapes, to determine the location of the eye. Then, the eye model is moved across the image of a detected face, computing a correlation score. The location with the highest score is then determined to be the location of an eye.

All of these methods can be used to accurately detect the locations of the eyes in a face region. In this paper, we follow a *template based matching* approach and show that our method achieves a higher performance rate than previous work.

Because there is little to no research done in the field of SWIR automatic eye detection, it is difficult to compare methodologies. However, there is a multitude of work done in automatic eye detection performed in the visible (350nm - 850nm) and Near Infrared (NIR - 850nm - 1050nm) spectra.

Table 1: Different methods for eye localization in the literature

Method	Pre-requirements	Uses Learning	Used Feature	Band
Valenti [8]	Face Region	X	Isophotes	Visible
Asteriadis [25]	Face Region	-	Edges	Visible
Wang [7]	Face Region	X	FDA Descriptor	Visible
Song [38]	None	-	Binary Edge Images	Visible
Shafi [14]	Face Region	-	Hybrid	Visible
Khairrosfaizal [10]	Face Region	-	Circular Hough Transform	Visible
Dowdall [19]	None	-	Integral Projections	NIR
Zhu [20]	None	X	Kalman Filtering	NIR
Previous Work [9]	Face Region	-	Normalized Correlation	SWIR
Proposed	None	-	Summation Range Filter	SWIR

With regards to the state of the art, Valenti *et al.* [8] use isophote properties in the eye region to accurately detect the center eye location in visible images. Assuming the face region is known or obtained from a face detection algorithm, anthropometric measures are used to determine the eye region areas of interest. Then, this estimated eye region is used to test the isophote approach. The authors report a high accuracy rate at a low normalized error ($\sim 80\%$ accuracy across databases at a normalized error of 5%). Asteriadis *et al.* [25] assigns a vector to each pixel in the edge map of an eye region in a visible image pointing to the closest edge pixel. By using a training set created off line, the detection and localization of the eyes is done by matching the length and slope of the vectors with the training database. Wang *et al.* [7] statistically learned discriminative features (by using Fisher Discriminant Analysis) to characterize eye patterns. Probabilistic classifiers are then learned to separate eyes from non-eyes. Multiple classifiers are combined in AdaBoost to form an accurate and robust eye detector. Song *et al.* [38] used a robust Binary Edge Image (BEI) and light

spots to accurately determine the locations of the eyes. By using vertical projections, eye locations are determined and light spots are used for the final eye location. Shafi *et al.* [14] combine the work of [33],[34], and [37] to create a hybrid eye detection algorithm utilizing illumination, color, and edge-density information to determine the location of the eyes. Lu *et al.* [17] uses rectangular as well as pixel-pattern-based texture features while Khairrosfaizal *et al.* [10] uses circular Hough transforms.

In the infrared spectrum, Dowdall *et al.* [19] used the lower band of the IR illumination to find the eyes since it is darker than the rest of the face, and the upper IR band to find the eyebrows, considering that the eyebrows reflect light extremely well in this range. Zhu *et al.* [20] also operates in the IR spectrum, and uses Kalman filtering and the mean shift tracking algorithm to track and detect eyes in real time. Due to the appearance and disappearance of corneal reflections from the IR sensors, blob detection is used to narrow down the eye detection search space. Table 1 lists different techniques used to determine the location of the eyes. All aforementioned methods operated

Figure 2: Overview of the proposed fully automated pre-processing methodology. a) Query Image b) Photometric Normalization c) Sample Average Face and Eye Templates d) Face Detection e) Eye Detection f) Eye Center Localization using Summation Range Filtering g) Geometric Normalization

in either the visible or the IR spectrum, but were not designed to efficiently operate in the Short Waved Infrared Spectrum.

As indicated in the last row of Table 1, this paper will describe a template based matching algorithm that can empirically determine where the face and eye regions are within an image. It can quickly and robustly determine the eye center using Summation Range Filters in multiple bands of the SWIR spectrum. We study the accuracy and robustness of the proposed approach to scaling, image compression, and image blurring and compare these results with multiple available eye detection methods, including the state of the art. Furthermore, we test and show that our eye detec-

tion method yields superior face recognition rates than available methods.

3. Methodology

In this section, the methodology of performing automatic face/eye detection in multiple SWIR wavelengths (i.e. 1150, 1250, 1350, 1450, and 1550nm) will be described. The authors assumed the data used was full frontal images of an upper body area with ideal lighting conditions and did not consider more challenging cases, with head pose variation and eye rotations. An overview of the proposed approach can be seen in Figure 2. Given a query image (Figure 2(a)), the method-

ology consists of five main processes, i.e. preprocessing, automatic face detection, eye region localization, summation range filtering, and geometric normalization. This leads to an image that is suitable for a face recognition system. Sample images of subjects from each individual wavelength can be seen in Figure 1.

Preprocessing

SWIR images tend to have low contrast in the face region, especially in the upper wavelength bands (i.e. 1350, 1450, and 1550nm). Instead of the human skin reflecting those wavelengths back into the camera, the moisture in the skin tends to absorb higher SWIR wavelengths, causing the skin to appear very dark (even for very light skinned subjects). In order to compensate for this, photometric normalization techniques are used to increase the contrast. However, different photometric techniques bring out unique features that are beneficial for face/eye detection in different wavelengths. For each wavelength, a specific photometric normalization is used (Figure 2(b)). For a detailed description and justification on which photometric normalization is used on each individual wavelength, refer to Section 4.3.

Because the proposed technique is a *template based* method, average templates are needed. Therefore, for each wavelength, 7 subjects are randomly selected. The faces are geometrically normalized, cropped, and averaged together to create an *average face template* (Figure 2(c.1)). Then, the eye regions from this template are cropped and used as *average eye templates* (Figure 2(c.2)). These average templates are then saved and used on all images in the database.

Automatic Face Detection

Figure 3: Correlation map of query face image for automatic face detection. The highest peak of the correlation map corresponds to the location of the face.

Because of the unique qualities that SWIR images have, typical face detection algorithms perform poorly and could not be used. Therefore, a template based face detection algorithm was developed to spatially locate the face. For each pixel in the query image, the 2D normalized cross correlation is computed between the region of that pixel and the average face template. Mathematically, the 2D normalized cross correlation can be described as:

$$\delta(u, v) = \frac{\sum_{x,y} [f(x,y) - \bar{f}_{u,v}][t(x-u, y-v) - \bar{t}]}{\{\sum_{x,y} [f(x,y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x-u, y-v) - \bar{t}]^2\}^{0.5}} \quad (1)$$

where f is the image, $\bar{t}$ is the mean of the template, and $\bar{f}_{u,v}$ is the mean of $f(x,y)$ in the region under the template. The convolution of the image and average template then yields a correlation map. Figure 3 shows the correlation map in correspondence with the query image. The highest location within the correlation map (i.e. the peak) is the location of the face. However, different average templates yield different results. Because of this, multiple average templates (in our case 5) are created and used to increase the chance of finding the

correct location. Mathematically, the final location of the face can be described as:

$$\hat{\delta}(u, v) = \underset{x}{\operatorname{argmax}}(\delta_x(u, v)) \quad (2)$$

where $\delta_x(u, v)$ is the location of the highest correlation coefficient obtained from average template x (in our case $x = 1, \dots, 5$). $\hat{\delta}(u, v)$ then corresponds to the upper left hand point of the face region and the face can be cropped to the size of the average templates (Figure 2(d)). This approach helps narrow down the search space when locating the eye regions.

Eye Region Localization

Since the location of the face is now known, the location of the eye regions can be easily determined. In order to further reduce the search space, the face is split into four equal regions (i.e. top left, top right, bottom left and bottom right). Assuming that the face region is found correctly through the method described above, the right and left eye should be located in the top right and top left regions respectively. Therefore, to obtain the left and right eye regions, the average eye templates are convolved with their respective quadrants using Equation 1. As stated above, different average eye templates yield different results. Therefore, the process is repeated multiple times (in our case 5) using unique templates to increase the chance of obtaining the correct region (Figure 2(e)). Justification for this practice is described in Section 4.4. Then, Equation 2 can be used to find the final location of the eye regions. Note that no assumptions as to where the eye regions are is made and if the face detection fails, the eyes will not be in the proper region and the effect is shown in the overall results

(see Section 4.1).

Summation Range Filter

Although the region of the eye can be easily found, the center of the eye cannot always be said to be the center of the found region. Therefore, an accurate way of determining the correct center of the eye must be employed. During the data collection process, subjects were asked to face and look directly at the camera as the samples were taken. Because an illumination source was used for an ideal capturing scenario, corneal reflections were assumed to be present and located in the center of the subjects eyes. Taking advantage of this, summation range filters can be used to more accurately determine the center of the eye by determining the location of those corneal reflections. The summation range map ($S(x, y)$) can mathematically be described as:

$$S(x, y) = \sum_{x=-1}^1 \sum_{y=-1}^1 R(x, y) \quad (3)$$

where

$$R(x, y) = \operatorname{argmax}(I(x-1 : x+1, y-1 : y+1)) - \operatorname{argmin}(I(x-1 : x+1, y-1 : y+1)) \quad (4)$$

where $I(x, y)$ is the original cropped eye region. Then, the eye final eye center is said to be:

$$P(x, y) = \operatorname{argmax}(S(x, y)) \quad (5)$$

This process is done for both the right and left eye regions to determine the final locations for the right and left eye respectively (Figure 2(f)). Figure 4 shows a sample found right and left eye region and its corresponding range filter while Figure 5

Figure 5: Sample summation range maps and their corresponding detected eye regions. Note that the peaks are not in the center of the found regions.

Figure 4: Sample detected left and right eye regions and their corresponding summation range images.

illustrates $S(x, y)$ in both the $1150nm$ and $1550nm$. Notice that the peaks in Figure 5 are not directly located in the middle of the found regions.

Geometric Normalization

In order to assist facial recognition systems, the left and right eye locations are used to geometrically normalize the image. By setting a standard interocular distance, the eye locations can be centered and aligned onto a single horizontal plane and resized to fit said distance. This en-

sures that if the eyes are found correctly, the left and right eyes are guaranteed to be in the same position every time, an assumption that is crucial to facial recognition algorithms. Therefore, all face images were geometrically normalized based on the found locations to have an interocular distance of 60px with a resolution of 130x130px (Figure 2(g)). The geometrically normalized images can then be used in a facial recognition system (Figure 2(h)).

4. Evaluation

In this section, our method is tested on multiple bands of the SWIR spectrum. Additionally, our method is tested against well-known eye detection algorithms as well as for robustness to change in scale, compression, and blur. Finally, face recognition tests based on the found eye locations are performed. Because there are no publicly available datasets collected in the SWIR

Figure 6: Results of running our proposed algorithm with (red) and without (blue) the summation range filter. When the summation range filter is not used (blue), the center of the found eye region is used. At $e = 0.05$, the summation range filter increases the accuracy of our original approach by 25.3%, 28.5% and 16.4% for the 1150, 1350, and 1550nm bands respectively.

band, a database of images was captured using a Goodrich SU640 camera. The SU640 is an Indium Gallium Arsenide (InGaAs) video camera featuring high-sensitivity and wide dynamic range. The model used has a 640x512 FPA with 25m pixel pitch, and >99% pixel operability. The spectral sensitivity of the SU640 ranges uniformly from 700-1700nm wavelength. The response falls rapidly before 700nm and after 1700nm. In order to capture a wide spectrum of wavelengths, band pass filters with a 50nm range centered at 1150, 1250, 1350, 1450, and 1550nm respectively were used to collect individual wavelengths. For the database used, 135 subjects were collected with 10 samples per subject per wavelength, resulting in 6,750 images. Each subject was asked to look forward with no variability in pose, such as in mugshots, and was captured under ideal lighting conditions for each wavelength. All pupils were assumed to be in the middle of the eye, allowing for the pupil points to be used as geometric normalization points. More challenging datasets were not considered in this work, however, have been studied previously [12]. After collection, both the left and right eye were then

manually marked for their ground truth locations to use later for evaluation purposes. All experiments were performed on a 64-bit Windows 7 machine with 12GB of RAM running Intel Core i7 CPU @ 3.2Ghz using MATLAB R2012b.

In all experiments, the normalized error is used. This error, indicating the error obtained by the worse eye estimation, is used as the accuracy measure for the found eye locations. Proposed by [11], the normalized error is described as:

$$e = \frac{\max(d_{left}, d_{right})}{w} \quad (6)$$

where d_{left} and d_{right} is the Euclidean distance between the found left and right eye centers with the manually annotated ground truth and w is the Euclidean distance between the eyes in the ground truth. In the normalized error, $e \leq 0.25$ (or 25% of the interocular distance) roughly corresponds to the width of the eye (i.e. corner to corner), $e \leq 0.10$ roughly corresponds to the diameter of the iris, and $e \leq 0.05$ roughly corresponds to the diameter of the pupil. As a note, the authors were only able to optimize the algorithms used to compare the proposed method to the best of their

abilities with the data available. It is not the focus of this paper to fully optimize all existing techniques but only to use the available algorithms as comparison tools.

4.1. Effect of Summation Range Filter

To show the effectiveness of adopting the summation range filter after eye region localization, our algorithm was performed with and without the use of the summation range filter for three different wavelengths of the SWIR band. When the summation range filter was not used, the center of the localized eye region was used as the final location. The results of this experiment can be seen in Figure 6. It can be seen that the use of the summation range filter significantly increases the accuracy of the proposed algorithm across all three bands. At $e = 0.05$, the summation range filter increases the accuracy of our original approach by 25.3%, 28.5% and 16.4% for the 1150, 1350, and 1550nm bands respectively.

4.2. Photometric Normalization

Because each individual wavelength has different contrast and illumination properties, a study on multiple photometric normalization techniques was performed. In this study, seven scenarios were considered and tested against each other to determine the photometric normalization that has the highest eye detection performance overall. The following photometric normalization techniques were used:

- **Contrast Limited Adaptive Histogram Equalization (CLAHE):** CLAHE [35] operates on small local regions (8x8 for our experiments) in the image and applies histogram equalization on each individual region (in contrast to the entire image in regular histogram equalization). Mathematically,

it is described as:

$$f(n) = \frac{N-1}{M} \times \sum_{k=0}^n h(k), \quad (7)$$

where M and N are the number of pixels and grey level bins in each sub-region, respectively, and h is the histogram of each sub-region. In order to increase the contrast while decreasing the amount of noise, CLAHE redistributes each histogram so that the height of each bin falls below a predetermined threshold (0.1 in reported experiments). Specifically, gray level counts above the threshold are uniformly distributed among the gray levels below it. Finally, the patches are subsequently combined using bilinear interpolation.

- **Log Based Single Scale Retinex (LBSSR):** This photometric technique [36] decomposes the image into two components, i.e. illumination $L(x, y)$ (the amount of light falling on the targeted object) and reflectance $R(x, y)$ (the amount of light reflecting off the targeted object). The illumination component is estimated as a low-pass version of the original image, while the reflectance component is obtained by dividing the original image from the illumination image. Mathematically, this can be described as:

$$I(x, y) = L(x, y) \times R(x, y) \quad (8)$$

$$L(x, y) = I(x, y) * G_{\sigma}(x, y) \quad (9)$$

where G_{σ} is a Gaussian of scale σ and $*$ denotes the convolution between the image and the kernel. Finally, the reflectance image is estimated as:

$$R(x, y) = \log_{10} \left(\frac{I(x, y)}{L(x, y)} \right) \quad (10)$$

Figure 7: Sample success (top row) and failed (bottom row) eye detection locations on the 1150nm database. Red dots are manual locations while blue dots are found locations.

Figure 8: Examples of photometric normalizations used: (a) Original Image (b) CLAHE (c) LBSSR (d) CLAHE LBSSR (e) TBSSR (f) CLAHE TBSSR (g) TT

Table 2: Normalized errors (maximum error for both left and right eye) using the proposed method (including summation range filter) across six different photometric normalizations to determine the highest overall accuracy per wavelength. 5% normalized error is approximately within the diameter of the pupil, 10% diameter of the iris, and 25% the diameter of the eye. The bold results represent the highest 5% normalized error and were used for further experimentation.

		Photometric Normalization						
		Original	CLAHE	LBSSR	CLAHE LBSSR	TBSSR	CLAHE TBSSR	TT
1150	5%	42.74	27.70	59.63	41.78	56.52	33.33	51.93
	10%	80.67	65.70	90.81	81.78	90.00	72.96	83.26
	25%	93.19	90.37	98.00	96.89	98.22	95.85	96.96
1250	5%	36.89	25.63	55.93	44.15	53.56	35.85	42.52
	10%	81.48	69.63	91.85	86.30	91.33	79.70	77.33
	25%	90.15	90.81	98.00	95.33	98.37	95.85	86.44
1350	5%	20.37	15.93	41.41	27.63	53.56	23.41	32.30
	10%	52.44	46.81	87.04	71.78	91.33	60.67	75.26
	25%	68.37	81.11	96.07	93.63	98.37	89.70	91.56
1450	5%	8.89	10.00	22.81	28.83	26.81	28.52	12.00
	10%	24.81	30.52	70.30	73.85	70.67	72.37	31.33
	25%	44.07	46.15	84.74	87.85	83.70	86.67	38.15
1550	5%	10.00	11.78	32.15	31.63	35.33	31.48	24.74
	10%	22.96	29.19	80.67	74.89	80.30	73.41	53.56
	25%	46.52	49.85	90.74	89.93	88.81	90.89	59.56

- **CLAHE LBSSR:** A common problem with LBSSR is that images tend to become over saturated or “washed out”. This can have negative effects on eye detection algorithms. Furthermore, “halo” artifacts may be introduced depending on the scene and scale value chosen for the Gaussian smoothing function. Certain modifications to the LBSSR have been introduced, such as the multi-scale retinex approach [18], but at the cost of increased processing speed. Therefore, in this work, we applied the CLAHE approach listed above to the LBSSR image to help compensate for the afore-

mentioned approaches and to increase the contrast of the image.

- **Tangent Based Single Scale Retinex (TBSSR):** By using different non-linear transformations on the LBSSR [36], different image representations can be obtained. Therefore, the log in Equation 10 is replaced with an arc-tangent transformation, resulting in the following:

$$R(x, y) = \text{atan} \left(\frac{I(x, y)}{L(x, y)} \right) \quad (11)$$

- **CLAHE TBSSR:** Just as LBSSR, the TBSSR can cause over saturation and

haloing effects. Therefore, the CLAHE approach was also applied to the TBSSR image to correct the contrast issues mentioned.

- **Tan and Triggs (TT):** This photometric normalization, proposed by [21], incorporates a series of algorithmic steps that allow for the reduction of illumination variations, local shadowing and highlights, while still preserving the essential elements of visual appearance. These steps include gamma correction (raising each pixel value to a certain value, in this case 2), difference of Gaussian filtering (subtraction of an original image from a blurred version of the same image), and contrast equalization (suppressing larger intensities while maintaining lower intensities).

Sample images for each photometric normalization can be seen in Figure 8. To evaluate the performance, the entire system (including face and eye detection) was tested on all photometric normalizations as well as the original image. Average face and eye templates were created for each specific photometric normalization and used to determine the location of the eye. Table 2 shows the results of the study using the proposed algorithm (including the summation range filter) while Figure 7 has example success and failure cases. It can be seen that the LBSSR has the best normalized error for 1150, 1250, and 1350 while CLAHE LBSSR and TBSSR performs best for 1450 and 1550, respectively. All further experiments on each database are conducted on the respective photometric normalizations that resulted in the highest normalized error.

4.3. Effects of Variable Average Templates

In order to justify the use of multiple templates while performing the proposed eye detection, a study was conducted to show the effects different templates have on accurately detecting eyes. Here, 5 average templates were created from different groups of 7 subjects chosen at random, creating unique templates for each eye, as shown in Figure 9. Assuming the face is known, the proposed method (taking the best match from all five templates) is compared to the performance of each individual template as well as the average of all 5 templates on both the 1150 and 1550 database. The normalized error for all 5 individual methods as well as the proposed method can be seen in Figure 10. Notice that each average

Figure 9: Different average eye templates for (a) the right eye and (b) the left eye in the 1550nm database

template has varying accuracies, especially when $e \geq 0.10$. Also notice that the proposed method, which uses the best match from all 5 templates, achieves much higher accuracy than any one individual template and the average of all 5 templates.

4.4. Comparison with State of the Art

Because there is little to no work done in SWIR eye detection, it is hard to compare the obtained results with the state of the art. In order to do this, algorithms that are commercially or academically available for testing must be used. Therefore, 3 different methods have been tested for comparison. First, the Cascaded Adaboost eye detection

Figure 10: Effects different templates have on eye detection accuracy. (a) effects on 1150 database and (b) effects on 1550nm database.

Figure 11: Our proposed method (red) compared to multiple available eye detection algorithms, including a commercial software (green), Viola and Jones (blue), and the state of the art (black) for 1150nm, 1350nm, and 1550nm assuming the faces are known. The top row shows results up to 25% of the interocular distance while the bottom row shows up to 10% of the interocular distance.

Table 3: Comparison of proposed and tested algorithm with manual and automatic face detection as well as face detection results. VJ - Viola and Jones; SOA - State of the Art.

	1150	1250	1350	1450	1550
Face Detection Rate	99.26	99.70	98.67	83.93	91.78
Proposed Eye Detection w/ Manual Faces	94.44	93.78	92.22	89.41	92.07
Proposed Eye Detection w/ Detected Faces	94.52	94.07	91.33	78.37	86.37
VJ Eye Detection w/ Manual Faces	35.33	46.07	72.15	79.11	82.90
VJ Eye Detection w/ Detected Faces	34.59	43.85	69.85	66.30	76.38
SOA Eye Detection w/ Manual Faces	1.56	2.52	3.93	6.67	5.78
SOA Eye Detection w/ Detected Faces	1.85	1.11	2.52	0.74	1.04
G8 Eye Detection w/ Detected Faces	97.26	93.44	79.33	24.81	31.04

classifier proposed by Viola and Jones [27] is available through the Computer Vision System Toolbox in MATLAB R2012b¹. This method uses weak classifiers and Haar features to encode details and uses a decision stump to determine the eye region. If an eye was detected, the final location used for accuracy was deemed to be the middle of the detected region. If no eye was detected, the location was set to [0,0]. Secondly, the state of the art eye detection algorithm, proposed by Valenti et al. [8] and published in PAMI in 2012, was obtained through the authors website² and tested using all optimized parameters listed in the publication. Finally, a commercial algorithm, namely G8³ is used. G8 is a black box algorithm that takes as input a raw image and outputs both left and

right eye locations. The normalized errors (assuming the face was known) for 1150nm, 1350nm, and 1550nm can be seen in Figure 11. Also, in order to show the effectiveness of the face detection step, each of the compared algorithms and the proposed algorithm are performed using manual face annotations and automatic face annotations (except in the case of G8 where manual face annotation was not an option). The results across all wavelengths for $e = 0.15$ can be seen in Table 3. In order to evaluate the accuracy of our face detection step, the middle pixel of the found face is compared to the middle pixel of the manually annotated face. If the found face was $e \leq 0.25$, then it was deemed to be a found face. The accuracy of our face detection step can be seen in the top row of Table 3.

Results for the face detection experiments show that in 1150, 1250, and 1350, where face detection has a high success rate, eye localization accuracy does not signifi-

¹<http://www.mathworks.com/help/vision/ref/vision.cascadeobjectdetectorclass.html>

²<http://staff.science.uva.nl/~rvalenti/index.php?content=EyeAPI>

³Provided by L1 Industries

cantly decrease. In some cases (e.g. 1150 and 1250), accuracy improves if only very slightly in the proposed approach. However, when the face detection accuracy is low, as in 1450 and 1550, eye localization accuracy is effected more greatly, dropping the accuracy by a significant amount (i.e. more than 12% for the proposed algorithm).

Notice that only in 1150nm does any of the algorithms outperform the proposed method, however, only at an error of $e \geq 0.10$. In order to have a fair comparison between our algorithm and the Viola and Jones algorithm, retraining the classifier with SWIR images was necessary. Therefore, OpenCV’s Cascaded Adaboost training algorithm, using default Haar feature parameters, was used to build new classifiers using strictly 1550nm SWIR images. 1,425 positive training samples (for both right and left eyes) and 13,590 negative samples, including 2,990 “hard” negative samples, were used for the training process. “Hard” negative samples include eye images that have extreme tilt from $+45^\circ$ to -45° in increments of 5° . To further increase the fairness of the comparison, the tested faces were broken into four quadrants and the left eye classifier was only applied to the top left quadrant while the right eye classifier was applied to the top right quadrant, as is done in the proposed methodology. If multiple eye candidates were found, the one candidate that was closest to the ground truth was kept as the final eye location. It can be seen that even after retraining the classifiers, the proposed method still outperforms the Viola and Jones in all wavelengths. The authors recognize that to obtain a better understanding of how the Viola and Jones algorithm performs in lower wavelengths, specific classifiers should be made for the comparison, however, lack of

data and time restraints made producing these results challenging.

The performance of the state of the art eye detection algorithm obtained from the authors website performs extremely poorly in comparison to the proposed method. With $e \leq 0.25$, the performance rate for 1150nm, 1350nm, and 1550nm is 6.37%, 9.04% and 14.74% respectively when given the face region only. After empirically testing the state of the art software, higher accuracies were achieved when limiting the search space to only the eye regions, however these results were not reported due to the authors publication and software in which it was stated that only the face region is necessary. In conclusion, using the proposed method, when the normalized error to $e \leq 0.09$ (or less then the estimated width of the iris), our algorithm outperforms all tested algorithms by more than 0.67%, 22.44%, and 73.63% for 1150nm, 1350nm, and 1550nm respectively.

4.5. Robustness of Approach

In order to test the robustness of our approach, three studies were conducted, i.e. robustness to scale, robustness to image compression, and robustness to image blurring. Each experiment was conducted on all images from 1150nm and 1550nm databases assuming the face is known.

- **Scale:** In order to study the effect of changing the scale (i.e. spatial resolution), the test images and average templates were down sampled from 100% (130x130px) to 10% (13x13px) in 10% increments. The results from this experiment can be seen in Figure 12. Note that in 1150nm, downscaling from 100% to 80% does not significantly affect the results, while in 1550nm, the

(a)

(b)

Figure 12: Effect of eye detection on downsampling images from 100% to 10% in 10% increments in a)1150nm and b)1550nm.

images can be down sampled to more than 60% without a significant change.

Figure 13: Sample JPEG compressed images from 1150nm with a) 100% b) 50% and c) 10% compression rates.

- **JPEG Compression:** Due to the high number of pixels many cameras have available today, storage size of a single image can be an issue. In order to decrease the amount of storage needed to save an image, compression can be used. However, different levels of compression can cause different artifacts in the images. Therefore, a study on the effect that image compression has on our approach was studied.

Here, JPEG compression was applied to all images ranging from 100% (11.3 KB) to 10% (5.22 KB) in 10% intervals. Example images with 100%, 50% and 10% JPEG compression can be seen in Figure 13 while the results for the study can be seen in Figure 14. Note that JPEG compression of up to 50% does not significantly affect the results in 1150nm while in 1550nm compression of up to 20% has minimal effect.

Figure 15: Sample images with Gaussian blur from 1150nm with a) $\sigma = 0.5$ b) $\sigma = 1.0$ and c) $\sigma = 1.5$.

- **Blur:** Because images can easily be taken when the capturing lens is out of focus, a study on image blurring was

(a)

(b)

Figure 14: Effect of eye detection on JPEG compression from 100% to 10% in 10% increments in a)1150nm and b)1550nm.

(a)

(b)

Figure 16: Effect of eye detection on Gaussian blur from $\sigma = 0.5$ to $\sigma = 1.5$ in 0.1 increments in a)1150nm and b)1550nm.

considered. In order to simulate a camera being out of focus, Gaussian blur was added to all images with a window size of 7 and σ ranging from 0.5 to 1.5 in 0.1 increments. Sample images with Gaussian blur can be seen in Figure 15 while the results of the study can be seen in Figure 16. Note that as σ increases, the accuracy has a vertical shift downwards in both 1150nm and 1550nm. This can be explained by the fact that as σ increases, the blurring effect increases, which in turn decreases the range within the eye region. This causes higher errors when determining the center of the eye.

4.6. Face Recognition Study

The main purpose of automatic eye detection is to assist in a FR's systems ability to geometrically normalize the face. If the locations of the eyes are found incorrectly, geometric normalization will not be performed correctly (i.e. lining up the eyes in the same location for all subjects) and consequently facial recognition performance will suffer. This is why correct and accurate eye locations are needed. In order to test the accuracy of the proposed method, three different face recognition algorithms were used, i.e. Local Binary Patterns (LBP), Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA). In these experiments, 1 sample from every subject (gallery image) was geometrically normalized using the manually annotated positions. This guaranteed that gallery had one geometrically correct aligned face image. Then, this image was matched to 1 sample from each subject (probe image) that was geometrically normalized using the found eye locations from each respective method (i.e. manual anno-

tation, proposed method, commercial, state of the art, and Viola and Jones). For both the LDA and PCA algorithms, 30 subjects with three samples per subject were chosen at random from the manually annotated geometrically aligned face images to perform training. Then the remaining 105 subjects were used for testing in LDA, PCA, and LBP. In order to obtain a fair comparison, the subjects used for training the LDA and PCA were left out during the testing phase of LBP, even though LBP requires no training. The Receiver Operating Characteristic (ROC) curve and the Cumulative Match Characteristic (CMC) curve are used to describe the performance of the methods. Figure 17 shows the performance of each method using the LDA, PCA, and LBP algorithms in the 1550nm databases (the most challenging database). Note that in all cases (LDA, PCA, and LBP), the proposed method produces better ROC curves and higher Rank-1 identification rates over all the other tested eye detection algorithms, except for the LBP algorithm where the proposed method has almost identical rates to the Viola and Jones method. Please also note that while all SWIR bands are useful for FR, the more suitable band is at 1550nm due to the fact that it is safest for the eyes. We have conducted a lot of experiments in the past on FR in the SWIR band. More details can be found at [2][5].

5. Conclusion

In this paper, we have proposed a novel unified eye detection method that can operate in multiple bands (i.e. 1150nm to 1550nm) using normalized correlation coefficients and summation range filters. After determining the location of the face, the eye regions are empirically found and sum-

Figure 17: Face recognition results when geometrically normalizing the face images with the proposed algorithm and the different comparison methods. Both verification (ROC curves - top row) and identification (CMC curves - bottom row) experiments were conducted for three different face recognition algorithms, i.e. a) LBP b) LDA and c) PCA.

mation range filters are used to accurately determine the center of the eye. The use of these simple filters yields low computational time (for 1,350 images, $\mu = 75.7ms/image$ and $std = 7.1ms/image$ using MATLAB R2012b) while allowing for robustness to multiple forms of image degradations, all while still achieving high face recognition rates.

An extensive evaluation of the proposed method was performed, testing it for accurate eye locations across multiple SWIR bands and for robustness to scale, image compression, and image blurring. The comparison with the commercial and academic algorithms and the state of the art suggests that under the general assumptions (i.e. the location of the face is known), our eye detection is able to achieve higher accuracy and can be successfully applied to multiple SWIR bands. With $e \leq 0.9$, our

method outperforms every other method. Also, using the locations found through our method obtained higher face recognition results than almost all other methods across the spectra when geometrically normalizing the face. Given the reported accuracy of the overall system, we believe that the proposed method provides a reliable way to not only localize the face region in SWIR images, but to accurately determine the correct location of the eyes.

As with any system, our approach has drawbacks that can cause failed eye localization. If a person is wearing eye glasses with a large reflection from the illuminating source, this can cause an uncharacteristically high value within the summation range filter, causing the system to believe the eye location is somewhere on the rim of the glasses. In our experiments, however, only 1 subject (or 10 samples per wave-

length) was wearing eye glasses. Another drawback of our system is dealing with uncooperative subjects who heavily squint or have their eyes totally closed. This causes very low range amounts in the range map. Consequently, our approach has a tendency to believe that the eye corners or the eye brows is the correct location. In our system, this subsample equates to $\sim 4\%$ of the database.

We have many plans for future work. First, collecting more data in order to build a larger, more diverse database would allow for further validation of our approach. Secondly, collecting data and performing automatic eye detection and face recognition on long range SWIR images (up to 1km away) at night would show the operational potential that this approach has. Finally, because of the challenges that higher wavelengths have (i.e no iris/pupil data, low contrast, different face features, etc.), altering the approach to work in both MWIR and LWIR face images will be studied to show that this approach works not only in SWIR but across multiple spectra.

Acknowledgments

This work is sponsored through a grant from the Office of Naval Research (N00014-08-1-0895) Distribution A - Approved for Unlimited Distribution. The authors are grateful to the students and staff at West Virginia University, especially Dr. Bojan Cukic, Dr. Jeremy Dawson, as well as WVU students including Nnamdi Osia, Neeru Narang, Jason Ice and Jerin Young, for their assistance in this work.

References

References

- [1] P. Grother, G. Quinn, P.J. Phillips, "Report on the Evaluation of 2D Still-Image Face Recognition Algorithms", in *Nat'l Inst. of Standards and Technology* interagency/internal report (NISTIR) 7709, 2010.
- [2] N. Kalka, T. Bourlai, B. Cukic, L. Hornak, "Cross-spectral Face Recognition in Heterogeneous Environments: A Case Study on Matching Visible to Short-wave Infrared Imagery", in *International Joint Conference on Biometrics* (IEEE, IAPR), 2011.
- [3] T. Bourlai, A. Ross, C. Chen, and L. Hornak, "A Study on using Middle-Wave Infrared Images for Face Recognition", in *SPIE Biometric Technology for Human Identification IX*, Baltimore, USA, Apr. 2012.
- [4] N. Osia, T. Bourlai, "Holistic and Partial Face Recognition in the MWIR Band Using Manual and Automatic Detection of Face-based Features", in *IEEE Conf. on Technologies for Homeland Security*, Boston, USA, Nov. 2012.
- [5] C. Whitelam, T. Bourlai, "On Designing SWIR to Visible Face Matching Algorithms", in *Intel Technology Journal, Edition on Biometrics and Authentication*, August, 2014.
- [6] H. Mendez, C. San Martn, J. Kittler, Y. Plasencia, E. Garca-Reyes, "Face recognition with LWIR imagery using local binary patterns", in *Advances in Biometrics*, Springer, Berlin Heidelberg 2009, pp. 327-336.
- [7] P. Wang, G., M.B., Q. Ji, J. Wayman, "Automatic Eye Detection and Its Validation", in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA June 2005, pp. 164-171.
- [8] R. Valenti, T. Gevers. "Accurate Eye Center Location through Invariant Isocentric Patterns", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Sept. 2012, Vol. 34, Issue 9, pp. 1785-1798.
- [9] C. Whitelam, Z. Jafri, T. Bourlai, "Multispectral Eye Detection: A Preliminary Study", in *Proc. IEEE International Conference on Pattern Recognition*, Turkey, Istanbul August 2010, pp. 209-212
- [10] W.M.K. Wan Mohd Khairrosfaizal, A.J. Nor'aini, "Eyes detection in facial images using Circular Hough Transform", in *Proc. 5th*

- International Colloquium on Signal Processing & Its Applications*, March 2009, pp. 238-242.
- [11] O. Jesorsky, K. J. Kirchbergand, R. Frischholz. "Robust face detection using the Hausdorff distance", in *Audio and Video Biom. Pers. Auth.*, 1992, pp 90-95.
 - [12] T. Bourlai, C. Whitelam, I. Kakadiaris, "Pupil detection under lighting and pose variations in the visible and active infrared bands", in *Proc. International Workshop on Information Forensics and Security*, Foz do Iguacu, Brazil, 2011, pp. 1-6.
 - [13] K. Chang, K. Bowyer, P. Flynn, "Face Recognition Using 2D and 3D Facial Data", in *Proc. Workshop on Multimodal User Authentication*, Santa Barbara, CA, 2003. pp. 25-32.
 - [14] M. Shafi, P. Chung, "A Hybrid Method for Eyes Detection in Facial Images", in *Proc. World Academy of Science, Engineering and Technology*, August 2008.
 - [15] G. Feng, P. Yuen, "Multi-cues eye detection on gray intensity image", in *Pattern Recognition*, May 2001, Vol. 34, Issue 5, pp. 1033-1046.
 - [16] S. Kong, J. Heo, B. Abidi, J. Paik, and M. Abidi. "Recent Advances in Visual and Infrared Face Recognition: A Review", in *Computer Vision and Image Understanding*, 2005, Vol. 97, Issue 1, pp. 103-135.
 - [17] H. Lu, W. Zhang, and D. Yang. "Eye detection based on rectangle features and pixel-pattern-based texture features", in *Proc. International Symposium on Intelligent Signal Processing and Communication Systems*, Xiamen, China, 2007, pp. 746-749.
 - [18] Z. Rahman, D. Jobson, and G. Woodell. "Multi-scale retinex for color image enhancement", in *Image Processing, 1996. Proceedings., International Conference on*, volume 3, Sept. 1996, pp. 1003-1006.
 - [19] J. Dowdall, I. Pavlidis, and G. Bebis. "Face detection in the near-ir spectrum", in *Image and Vision Computing*, New Zealand, 2003, Vol. 27, Issue 7, pp. 565-578.
 - [20] Z. Zhu, Q. Ji, K. Fujimura, and Kuangchih Lee. "Combining Kalman Filtering and Mean Shift for Real Time Eye Tracking Under Active IR Illumination", in *Proc. International Conference on Pattern Recognition*, Quebec, Canada, 2002, pp. 318-324.
 - [21] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions", in *Transactions on Image Processing*, 2010, Vol. 19, Issue 6, pp. 1635-1650.
 - [22] T. Bourlai and Z. Jafri, "Eye detection in the Middle-Wave Infrared Spectrum: Towards Recognition in the Dark", in *Proc. International Workshop on Information Forensics and Security*, Foz do Iguacu, Brazil, 2011.
 - [23] C. W. Park, J. M. Kwak, H. Park, and Y. S. Moon, "An Effective Method for Eye Detection Based on Texture Information", in *Proc. International Conference on Communications and Information Technology*, Washington, DC, USA, November 2007, pp. 586-589.
 - [24] Z. Xingming and Z. Huangyuan, "An Illumination Independent Eye Detection Algorithm", in *Proc. International Conference on Pattern Recognition*, Hong Kong, China, August 2006, pp. 392-395.
 - [25] S. Asteriadis, N. Nikolaidis, A. Hajdu, and I. Pitas, "An Eye Detection Algorithm using Pixel to Edge Information", in *Proc. International Symposium on Communications, Control and Signal Processing*, Marrakech, Morocco, March 2006.
 - [26] F. Nicolo, N. Schmid, "Long Range Cross-Spectral Face Recognition: Matching SWIR against Visible Light Images", in *IEEE Transaction of Information Forensics and Security*, 2012, Vol 7, Issue 6, pp. 1717-1726.
 - [27] P. Viola, M. Jones, "Robust Real-Time Face Detection", in *International Journal of Computer Vision*, May, 2004, Vol. 57, Issue 2, pp. 137-154.
 - [28] X. Xie, R. Sudhakar, H. Zhuang, "On improving eye feature extraction using deformable templates", in *Pattern Recognition*, June 1994, Vol. 27, Issue 6, pp. 791-799.
 - [29] W. Huang, R. Mariani, "Face detection and precise eyes location", in *Proc. International Conference on Pattern Recognition*, 2000.
 - [30] J. Huang, H. Wechsler, "Eye Detection Using Optimal Wavelet Packets and Radial Basis Functions (RBFS)", in *International Journal of Pattern Recognition. Artificial Intelligent*, 1999, Vol. 13, Issue 7, pp. 1009-1025.
 - [31] G.C. Feng and P.C. Yuen, "Variance Projection Function and its Application to Eye Detection for Human Face Recognition", in *International Journal of Computer Vision*, 1998, Vol. 19, Issue 9, pp. 899-906.

- [32] S. Kawato and J. Ohya, "Real-time Detection of Nodding and Head Shaking by Directly Detecting and Tracking the Between-eyes", in *Proc. 4th IEEE Int. Conf. on Automatic Face and Gesture Recognition*, Grenoble, France, 2000, pp. 40-45.
- [33] R. Hsu, M. Abdel-Mottaleb, A. Jain, "Face Detection in Color Images", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, Vol. 24, Issue 5, pp. 696-706.
- [34] C. Chiang, W. Tai, M. Yang, Y. Huang, C. Huang, "A Novel Method for Detecting Lips, Eyes and Faces in Real Time", in *Real-Time Imaging*, 2003, pp. 277-287.
- [35] A.M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement" in *Journal of VLSI signal processing systems for signal, image and video technology*, 2004, Vol. 38, Issue 1, pp. 35-44.
- [36] D. Jobson, Z. Rahman, and G. Woodell. "Properties and performance of a center/surround retinex" in *IEEE Transactions on Image Processing*, March, 1997, Vol. 6, Issue 3, pp. 451-462.
- [37] M. Sha, P. Chung, "Eyes Extraction from Facial Images Using Edge Density", in *Proc. 7th IEEE International Conference on Cybernetic Intelligent Systems*, 2008.
- [38] J. Song, Z. Chi, J. Liu, "A Robust Eye Detection Method Using Combined Binary Edge and Intensity Information", in *Pattern Recognition*, June 2006, Vol. 39, Issue 6, pp. 1110-1125.
- [39] O. Jesorsky, K. Kirchberg, R. Frischholz, "Robust Face Detection using the Hausdorff Distance", in *Audio- and Video-Based Biometric Person Authentication*, Springer, 2001, pp. 90-95.