Pre-print of article that appeared at the IEEE BTAS 2009.

The published article can be accessed from:
http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5339040

Difficult Detection: A Comparison of Two Different Approaches to Eye Detection for Unconstrained Environments

Walter J. Scheirer[1,2]   Anderson Rocha[3]   Brian Heflin[2]   Terrance E. Boult[1,2]

[1]University of Colorado at Colorado Springs, Colorado Springs, CO, USA
[2]Securics, Inc. Colorado Springs, CO, USA
[3]Institute of Computing, University of Campinas (Unicamp)
Campinas, SP, Brazil

Abstract— Eye detection is a well studied problem for the *constrained face recognition problem*, where we find controlled distances, lighting, and limited pose variation. A far more difficult scenario for eye detection is the *unconstrained face recognition problem*, where we do not have any control over the environment or the subject. In this paper, we take a look at two different approaches for eye detection under difficult acquisition circumstances, including low-light, distance, pose variation, and blur. A new machine learning approach and several correlation filter approaches, including a new adaptive variant, are compared. We present experimental results on a variety of controlled data sets (derived from FERET and CMU PIE) that have been re-imaged under the difficult conditions of interest with an EMCCD based acquisition system. The results of our experiments show that our new detection approaches are extremely accurate under all tested conditions, and significantly improve detection accuracy compared to a leading commercial detector. This unique evaluation brings us one step closer to a better solution for the unconstrained face recognition problem.

I. INTRODUCTION

Eye detection is a necessary processing step for many face recognition algorithms. For some of these algorithms, the eye coordinates are required for proper geometric normalization before recognition. For others, the eyes serve as reference points to locate other significant features on the face, such as the nose and mouth. The eyes, containing significant discriminative information, can even be used by themselves as features for recognition. Eye detection is a well studied problem for the *constrained face recognition problem*, where we find controlled distances, lighting, and limited pose variation. A far more difficult scenario for eye detection is the *unconstrained face recognition problem*, where we do not have any control over the environment or the subject. In this paper, we will take a look at eye detection for the latter, which encompasses problems of flexible authentication, surveillance, and intelligence collection.

A multitude of problems affect the acquisition of face imagery in unconstrained environments, with major problems related to lighting, distance, motion and pose. Existing work on lighting [12] [13] has focused on algorithmic issues (specifically, normalization), and not the direct impact of acquisition. Under difficult acquisition circumstances, normalization is not enough to produce the best possible recognition results - considerations must be made for image intensification, thermal imagery and electron multiplication. Long distances between the subject and acquisition system present a host of problems, including high f-numbers from very long focal lengths, which significantly reduces the amount of light reaching the sensor, and a smaller amount of pixels on the faces, as a function of distance and sensor resolution. Further, the interplay between motion blur and optics exasperates the lighting problems, as we require faster shutter speeds to compensate for the subjects movement during exposure, which again limits the amount of light reaching the sensor. In general, we'll have to face some level of motion blur in order for the sensor to collect enough light. Pose variation, as is well known, impacts the nature of facial features required for recognition, inducing partial occlusion and orientation variation, which might differ significantly from what a feature detector expects.

Both lighting and distance should influence sensor choice, where non-standard technologies can mitigate some of the problem discussed above. For instance, EMCCD sensors have emerged as an attractive solution for low-light surveillance (where low-light is both conditional, and induced by long-range optics), because they preserve a great bit of detail on the face and can use traditional imagery for the gallery (as opposed to midwave-IR sensors). This makes them very attractive for biometric application as well. However, the noise induced by the cooling of the sensor also presents new challenges for facial feature detection and recognition. In this paper, for the reasons cited above, we use the EMCCD to acquire our difficult imagery under a variety of different conditions, and apply several different eye detectors on the acquired images. To our knowledge, this is the first instance of specific feature detection work with EMCCD imagery.

The rest of this paper is as follows. In Section II we take a brief survey of the existing literature related to difficult detection and recognition problems, as well as the pattern recognition works relevant to the detection techniques discussed in this paper. In Section III we introduce a new machine learning based approach to feature detection for difficult scenarios, with background on the learning and feature approach used. In Section IV we introduce the correlation filter approach for feature detection, including a new adaptive variant. Our experimental protocol is defined

in Section V, followed by a thorough series of experiments to evaluate the detection approaches. Finally, in Section VI, we make some concluding remarks on our examination of algorithms for difficult feature detection.

II. RELATED WORK

On the algorithm front, we find only a few references directly related to difficult facial feature detection and recognition. Super-resolution and deblurring were considered in [16] as techniques to enhance images degraded by long distance acquisition (50m - 300m), and goes further to show recognition performance improvement for images processed with these techniques compared to the original images. The test data set in this paper for outdoor conditions is taken as sequential video under daylight conditions; the super-resolution process considers direct sequences of detected faces from the captured frames. The problem with this approach is that under truly difficult conditions, as opposed to the very controlled settings of [16] (full frontal imagery, with a constant inter-ocular distance), it is likely that a collection of detected faces in a direct temporal sequence will not be possible, thus reducing the potential of such algorithms. The work of [7] is more along the lines of what is explored in this paper, including a thorough discussion of the underlying issues that impact algorithm design, as well as an explanation of how to perform realistic controlled experiments under difficult conditions, and algorithmic issues such as predicting when a recognition algorithm is failing in order to enhance recognition performance.

In the more general pattern recognition work, we do find several learning techniques applied to standard data sets for eye detection. Many different learning techniques have been shown to be quite effective for the eye detection problem. The work of [1] is most closely related to the learning technique presented in this work in a feature sense, with PCA features derived from the eyes used as input to a neural network learning system. Using a data set of 240 images of 40 different full frontal faces, this technique is shown to be as accurate as several other popular eye detection algorithms. [18] uses normalized eye images projected onto weighted eigenspace terrain features, as features for an SVM learning system. [17] uses a recursive non-parametric discriminant feature as input to an AdaBoost learning system.

For recognition, a very large volume of work exists for correlation, but we find some important work on feature detection as well. Correlation filters [19] [20] are a family of approaches that are tolerant to variations in pose and expression, making them quite attractive for detection and recognition problems. Excellent face recognition results have been reported for the PIE data set [10], and the FRGC data set [22]. For the specific problem of eye detection, [21] first demonstrated the feasibility of correlation filters, while [11] introduced a more sophisticated class of filters that are more insensitive to over-fitting during training, more flexible towards training data selection, and more robust to structured backgrounds. All of these approaches have been tested on standard well-known data sets, and not the more difficult imagery we consider in this paper. We discuss correlation in detail in Section IV.

Of course we should reduce the impact of difficult conditions using better sensors and optics, which is why we choose to use EMCCD sensors to allow faster shutter speeds. For the optics, one possibility gaining attention is the use of advanced Adaptive Optics (AO) models [14], which have proved effective for astronomy, though most do not apply to biometric systems. Astronomy has natural and easily added artificial "point sources" for guiding the AO process. Secondly, astronomical imaging is vertical, which changes the type and spatial character of distortions. More significantly, they have near point sources for guides, allowing for specialized algorithms for estimation of the distortions. Horizontal terrestrial atmospheric turbulence is much larger and spatially more complex making it much more difficult to address. To date, no published papers discuss an effective AO system for outdoor biometrics. While companies such as AOptix[1] have made interesting claims, public demonstrations to date have been stationary targets indoor < 20m, where there is no atmospherics and minimal motion blur.

A critical limiting question for adaptive optics, beyond the lack of an obvious reference point sources in biometric data, is the assumption of wave-front distortion and measurement. For visible and NIR light, the isoplanatic angle is about 2 arc seconds (0.00027 degrees or motion of about 0.08mm at 50m). Motion outside the isoplanatic violates the wave-front model needed for AO correction [15]. An AO system may be able to compensate for micro-motion on stationary targets, where a wave-front isoplanatic compensation AO correction approach would be a first-order isoplanatic approximation to small motions, but it's unclear how it could apply to walking motions that are not well modeled as a wave-front error.

III. THE MACHINE LEARNING APPROACH

The core concept of our machine learning approach for detection is to use a sliding window search for the object feature, using image features extracted from the window and applying a classifier to those features. For different difficult environments we can learn different classifiers. We first review the learning and image features used.

A. Learning Techniques

Supervised learning is a machine learning approach that aims to estimate a classification function f from a *training data set*. Such training data set consists of pairs of input values X and its desired outcomes Y [4]. Observed values in X are denoted by x_i, i.e., x_i is the ith observation in X. Often, x is as simple as a sequence of numbers that represent some observed features. The number of variables or features in each $x \in X$ is p. Therefore, X is formed by N input examples (vectors) and each input example is composed by p features or variables.

The commonest output of the function f is a label (class indicator) of the input object under analysis. The learning
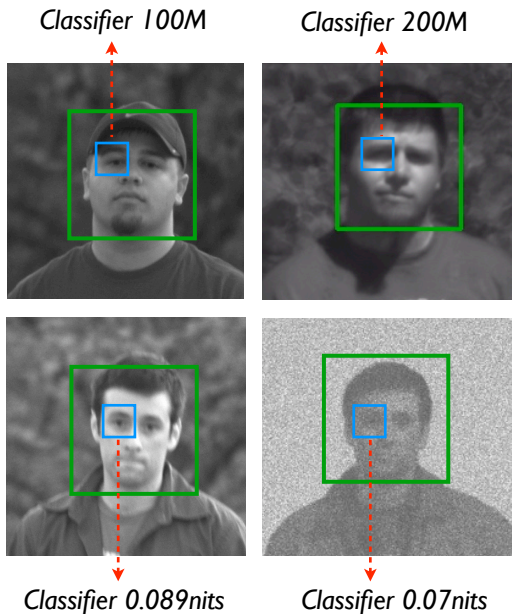
Fig. 1. The approach: build classifiers for different conditions, such as distance and illumination. While general enough for any feature that can be represented by a window of pixels, the eye is shown here, and in subsequent figures, as an example.
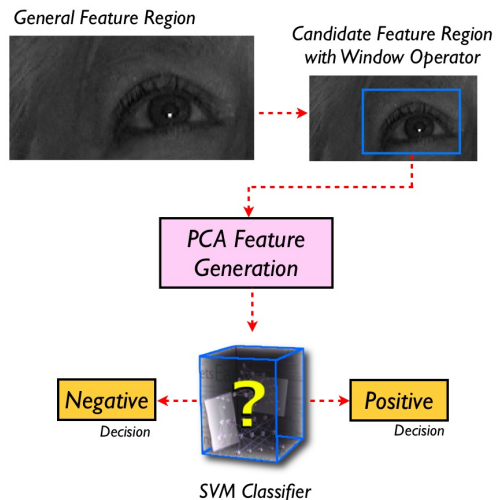

Fig. 2. The basic algorithm is straightforward. First, a feature region is isolated (using pre-set coordinates) from the face region returned by the face detector (separate from the feature detection). Next, using a pre-defined sliding window over the feature region, candidate pixels are collected. PCA feature generation is then performed using the pixels in the window. Finally, the PCA coefficients are treated as feature vectors for an SVM learning system, which produces the positive or negative detection result.

task is to predict the function outcome of any valid input object after having seen a sufficient number of training examples.

In the literature, there are many different approaches for supervised learning such as Linear Discriminant Analysis, Support Vector Machines (SVMs), Classification Trees, and Neural Networks. We focus on an SVM-based solution.

B. PCA Features

Principle Components Analysis [2], that battle-worn method of statistics, is well suited reduction of dimensionally reduction of image data. Mathematically defined, PCA is an orthogonal linear transformation, which after transforming data leaves the greatest variance by any projection of data on the first coordinate (the *principal component*), and each subsequent level of variance on the following coordinates. For a data matrix X^T, after mean subtraction, the PCA transformation is given as

$$Y^T = X^T W = V\Sigma \tag{1}$$

where $V\Sigma W^T$ is the singular value decomposition of X^T. In essence, for feature detection, PCA provides a series of coefficients that become a feature vector for machine learning. Varying numbers of coefficients can be retained, depending on the energy level that provides the best detection resolution.

C. PCA + Learning Algorithm

A learning based feature detection approach allows us to learn over features gathered in the appropriate scenarios in which a recognition system will operate, including

illumination, distance, pose, and weather (Fig. 1). By projecting a set of candidate pixels against a pre-computed PCA subspace for a particular condition, and classifying the resulting coefficients using a machine learning system yields an extremely powerful detection approach. The basic algorithm, depicted in Figure 1, begins with the results of the Viola-Jones face detector [3], implemented to return a face region that is symmetrical. With the assumption of symmetry, the face can be separated into feature regions, which will be scanned by a sliding window of a pre-defined size $w \times h$. Each positive marginal distance returned by an SVM classifier is compared against a saved maximum, with new maximums and corresponding x and y coordinates being saved. When all valid window positions are exhausted, the maximum marginal value indicates the candidate feature coordinate with the highest confidence. While for this work we are only interested in the eyes, we do note that the generic nature of the proposed approach allows for the detection of any defined feature.

Speed of such a sliding window approach is of concern. If we assume the window slides one pixel at a time, a 50×45 window (a window size suitable for eye detection on 160 pixels faces) in an 80×60 potential feature region, 496 locations must be scanned. One approach to enhancing speed is through the use of multiple resolutions of feature regions. Figure 3 depicts this, with the full feature region scaled down by 1/4 as the lowest resolution region considered by the detector. The best positive window (if any) then determines a point to center around for the second (1/2 resolution) scale's search, with a more limited bounding box defined around this point for the search. The process then repeats again for the highest resolution. Presuming a strong classifier, the positive

windows will cluster tightly around the correct eye region. A further enhancement to the algorithm is to determine the best positive window for the first row of the feature region where positive detections have occurred. From the x coordinate of this best window x_{best}, the scanning procedure can reduce the search space to $(x_{best} + c) - (x_{best} - c) + 1$ windows per row of the feature region, where c is some pixel constant set to ensure flexibility for the search region. This approach does come with a drawback - the space requirement for PCA subspaces and SVM classifiers increases by number of features $\times$ number of scales.
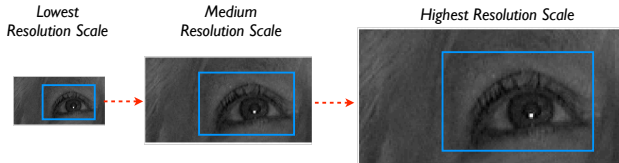

Fig. 3. The speed of sliding window approaches is always a concern. To increase computational performance, a multi-resolution approach can be used to reduce the area that must be scanned. While reducing time, this does increase the space requirement for PCA subspaces and SVM classifiers (number of features x number of scales).

IV. THE CORRELATION FILTER APPROACH

Existing correlation filters as considered in this work consist of Minimum Average Correlation Energy (MACE) filters [19] and Unconstrained Minimum Average Correlation Energy (UMACE) filters [20]. Both of these approaches produce a single correlation filter for a set of training images. For feature detection, these techniques produce a sharp correlation peak after filtering in the positive case, from which the correct coordinates for the feature can be derived (an example of this is shown in Figure 4). The variations between MACE, UMACE, and our own AACE are described below.

Synthesis of the baseline MACE filter begins with cropping out regions of size 40x32 from our training data with the eye centered at coordinates (21,19). The MACE filter specifies a single correlation value per input image, which is the value that should be returned when the filter is centered upon the training image. Unfortunately when more than 4-6 training images are used this leads to over fitting of the training data and decreases accuracy in eye detection. After cropping the eye region, it is transformed to the frequency domain using a 2D Fourier transform. Next the average of the power spectrum of all of the training images is obtained. Then MACE filter is synthesized using the following formula:

$$h = D^{-1}X(X'D^{-1}X)^{-1}u \qquad (2)$$

where D is the average power spectrum of the N training images, X is a matrix containing the 2D Fourier transform of the N training images, and u is the desired filter output.

Finally, after the normalized cross correlation operation is performed the global maximum or peak location is chosen

as the detected eye location in the original image with the appropriate offsets.

1) UMACE and AACE Filters for Feature Detection: Synthesis of our Adaptive Average Correlation Energy (AACE) filter is based on a UMACE filter. We start the filter design start by cropping out regions of size 64x64 for the training data, with the eye centered at coordinates (32,32). O After the eyes are cropped, each cropped eye region is transformed to the frequency domain using a 2D Fourier transform. Next, the average training images and the average of the power spectrum is calculated. The base UMACE filter for our AACE filter is synthesized using the following formula:

$$h = D^{-1}m \qquad (3)$$

where D is the average power spectrum of the N training images, and m is the 2D Fourier transform of the average training image. The UMACE filter is stored in its frequency domain representation to eliminate another 2D Fourier transform before the correlation.

One advantage of the UMACE filter over the MACE filter is that over-fitting of the training data is avoided by averaging the training images. Furthermore, we found that training data taken under ideal lighting conditions performed well for difficult detection scenarios. Whenever the UMACE filter was trained on EMCCD or low light images the UMACE filter did not work as well. An unseen advantage of the UMACE filter not performing well when training on EMCCD or low light images and performing well when training on images taken under ideal lighting conditions is that this allows us to build an extremely robust filter that can operate in a wide array of lighting conditions, instead of requiring different training data for different lighting levels, as was the case with the machine learning based detector.

The concept of the AACE filter is to take the UMACE filter, trained on good data, and adapt it, per image, for the environmental degradations using estimates of blur and noise. The UMACE filter has to be preprocessed by a Hamming window to help reduce the edge effects and impact of high frequency noise that is prevalent in low-light EMCCD imagery, in the spectrum. Our testing showed that windowing both the filter and input image decreased the accuracy of the eye detector. For the second AACE extension, a motion blur estimate can be convolved into a UMACE filter using only a point wise multiply of the motion blur Optical Transfer Function (OTF) and the UMACE filter. Finally, after the correlation operation is performed the global maximum or peak location is chosen as the detected eye location in the original image with the appropriate offsets. Figure 4 shows an example correlation output with the detected eye centered at coordinates (40,36).

V. EXPERIMENTS

A. The Photo-head Testing Protocol

Performing controlled experiments for difficult acquisition circumstances is a challenge, due to high degrees of variability. [7] introduce a specialized experimental setup called "Photo-head". In the setup described in those papers,
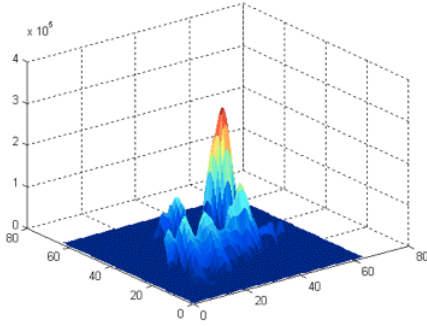
Fig. 4. Example Correlation Output with the Detected Eye Centered at Coordinates (40,36).

two cameras were placed 94ft and 182ft from a weather-proof LCD panel in an outdoor setting. The FERET data set was displayed on the panel at various points throughout the day, where it was re-imaged by the cameras over the course of several years. This unique re-imaging model is well suited to biometric experiments, as we can control for distance, lighting and pose, as well as capture statistically meaningful samples in a timely fashion. Further, it allows for reproducible experiments by use of standard data sets that are re-imaged.

In our own setup, instead of imaging an LCD panel, we used a Mitsubishi PK10 LCD pocket projector, which has a resolution of 800x600 pixels and outputs 25 ANSI Lumens, to project images onto a blank screen. The experimental apparatus was contained in a sealed room, where lighting could be directly controlled via the application of polarization filters to the projector. The camera used for acquisition was a SI-1M30-EM low-light EMCCD unit from Salvador Imaging. At its core, this camera utilizes the TI TX285SPD-B0 EMCCD sensor, with a declared resolution of 1004x1002 (the effective resolution is actually 1008x1010). To simulate distance, all collected faces were roughly 160 pixels in width (from our own work in long distance acquisition, this is typical of what we'd find at 100M, with the appropriate optics). Photo-head images can be seen in Figure 7.

In order to assess and adjust the light levels of the photo-head imagery, we directly measure the light leaving the projected face in the direction of the sensor - its *luminance*. The candela per square meter ($\frac{cd}{m^2}$) is the SI unit of luminance; nit is a common non-SI name also used for this unit (and used throughout the rest of this paper). Luminance is valuable because it describes the "brightness" of the face and does not vary with distance. For our experiments, luminance is the better measure to assess how well a particular target can be viewed - what is most important for biometric acquisition. More details on this issue of light face and face acquisition can be found in [6].

B. Evaluation of Machine Learning Approach

In order to assess the viability of the detector described in Section III-C, a series experiments under very difficult

conditions was devised. First, using the photo-head methodology of Section V-A, a subset of the CMU PIE [5] data set was re-imaged in a controlled (face sizes at approximately the same width as what we'd collect at 100M), dark indoor setting (0.043 - 0.017 face nits). Defined feature points are the eyes, with window size of 45×35 pixels. For SVM training, the base positive set consisted of 250 images $\times$ (8 1-pixel offsets from the ground-truth + ground-truth point), for each feature. The base negative set consisted of 250 images $\times$ 9 pre-defined negative regions around the ground-truth, for each feature. The testing set consisted of 150 images per feature. The actual data used to train the PCA subspaces and SVM classifiers varies by feature, and was determined experimentally based on performance. For the left eye, 1000 training samples were provided for subspace training, and for the right eye, 1200 samples were provided. The experiments presented in this section are tailored to assess accuracy of the base technique, and are performed at the highest resolution possible, with the window sliding 1 pixel at a time.

The results for eye detection are shown in Figures 5 and 6. On each plot, the x axis represents the pixel tolerance as a function of distance from the ground-truth for detection, and the y axis represents the detection percentage at each tolerance. The proposed detection approach shows excellent performance for the photohead imagery. For comparison, the performance of a leading commercial detector (chosen for its inclusion in a face recognition suite that scored at or near the top of every test in FRVT 2006), is also plotted. The proposed detection approach clearly outperforms it till both approaches start converge after the pixel tolerance of 10. Examples of the detected feature points from the eye comparison experiment are shown in Figure 7.
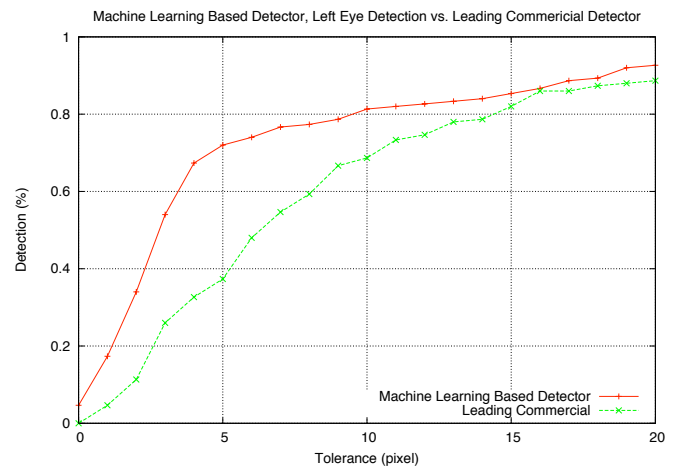

Fig. 5. Performance of the proposed machine learning based detector against a leading commercial detector for the left eye. The machine learning based detector clearly outperforms the commercial detector.

Even more extreme conditions were of interest for this research. Another photohead set was collected based on the FERET [8] dataset between 0.0108 - 0.002 nits. For an eye feature (left is shown here), a window of 50×45 was defined. The Gallery subset was used for training, with a subspace
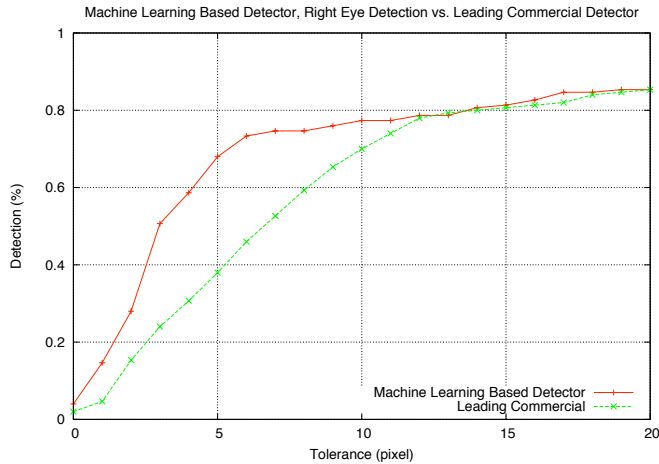
Fig. 6. Performance of the proposed machine learning based detector against a leading commercial detector for the right eye. Results are similar to that of the left eye in figure 5
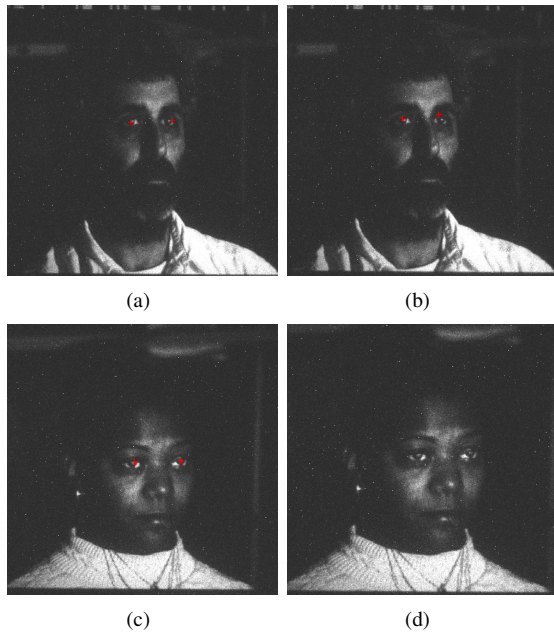

(a) (b)

(c) (d)

Fig. 7. Qualitative results for the proposed machine learning based detector (left) and a leading commercial detector (right) for comparison. The commercial detector is not able to find any eyes in image d.

of 1100 training samples, and a classifier composed of 4200 training samples (with additional images generated using the perturbation protocol above). For testing, all of the FAFC subset was submitted to the detector. The results of this experiment are shown in Figure 8; the commercial detector is once again used for comparison. From the plot, we can see the commercial detector failing nearly outright at these very difficult conditions, while the proposed detector performs rather well.

Blur is another difficult scenario that we have looked at. For this set of repeatable experiments, we produced a subset of images from the FERET data set (including the

ba, bj, and bk subsets) for three different uniform linear motion models of blur: blur length of 15 pixels, at an angle of 122 degrees; blur length of 17 pixels, at an angle of 59 degrees; blur length of 20 pixels, at an angle of 52 degrees. Sample images from each of these sets are shown in Figure 9. The classifier for detection was trained using 2000 base images of the left eye (split evenly between positive and negative training samples), derived from 112 base images (again, additional images were generated using the perturbation protocol above) at the blur length of 20 pixels, at an angle of 52 degrees. The subspace was trained with the 1000 positive images, with the same blur model. The testing set consisted of 150 images, with the left eye as the feature target for each of the three blur models. The results for this experiment are shown in Figure 10. From this figure, we can see that the machine learning based approach has a slight advantage over the commercial detector for the blur length of 20 pixels - the blur model it was trained with. For testing with the other blur models, performance is acceptable, but drops noticeably. Thus, we conclude that incorrect blur estimations can negatively impact this detection approach.
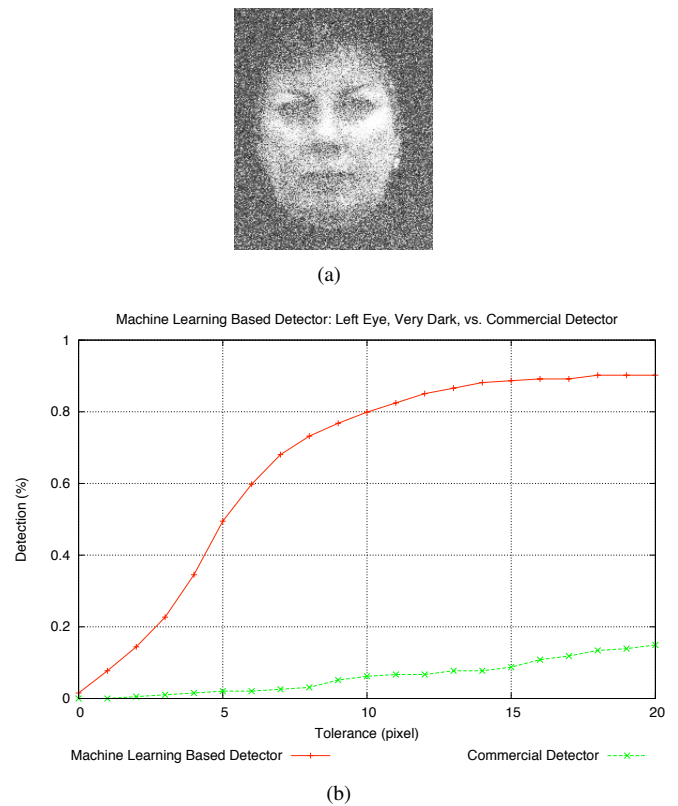

(a)

(b)

Fig. 8. Results comparing the machine learning based detector to a leading commercial detector, for the left eye, with very dark imagery (0.0108 - 0.002 nits). A sample image (a) is provided to signify the difficulty of this test (histogram equalized here to show "detail"). The commercial detector fails regularly under these conditions.

C. Evaluation of Correlation Approach

The experiments performed for the correlation approach are identical to the ones we performed for the machine
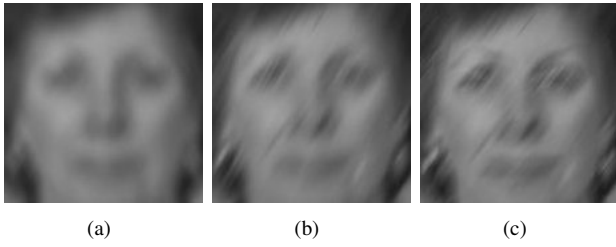
(a)   (b)   (c)

Fig. 9. Examples of blurred imagery for three different blur models used for experimentation. (a) Blur length of 15 pixels, at an angle of 122 degrees (b) Blur length of 17 pixels, at an angle of 59 degrees (c) Blur length of 20 pixels, at an angle of 52 degrees
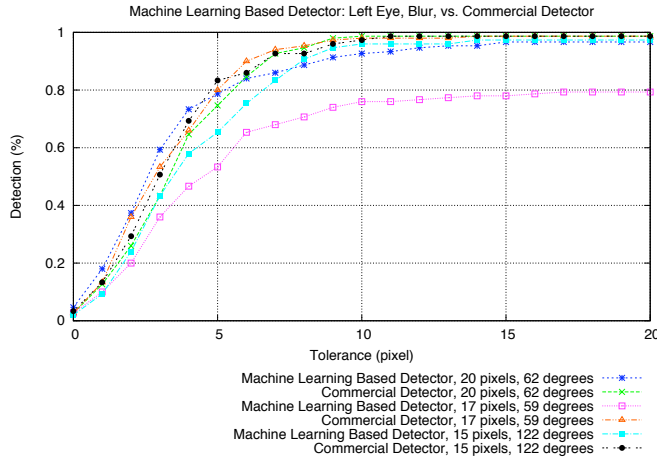

Fig. 10. Results comparing the machine learning based detector to a leading commercial detector, for the left eye, with three varying degrees of blur. The detector was trained using the blur length of 20 pixels, at an angle of 52 degrees.

learning approach, with the following training details. The MACE filters used in Figures 11 & 12 were trained with 6 eye images, while the MACE filter for Figure 13 used 4 training images (these values were determined experimentally, and yield the best performance). For the experiments of Figure 11 & 12, the AACE filter was synthesized with 266 images, for the experiment of Figure 13, the filter was synthesized with 588 images. For the AACE filter used in the experiment of Figure 14, the filter was synthesized with 1500 images, incorporating the exact same blur model as the machine learning experiments into the convolution operator. Furthermore, the training data for the experiments in Figures 11, 12 & 13 used images taken under ideal lighting conditions.

Comparing the AACE approach to the machine learning approach, the new correlation filter detector shows a significant performance gain over the learning based detector on blurry imagery (Figure 10 vs. Figure 14). What can also be seen from our experiments is a stronger tolerance for incorrect blur estimation, with the blur length of 17 pixels, 59 degrees performing just as well as the training blur model; this was not the case with the machine learning based detector. In all other experiments, the AACE filter detector

produced a modest performance gain over the machine learning based detector. The performance of MACE was poor for all test instances that it was applied to.
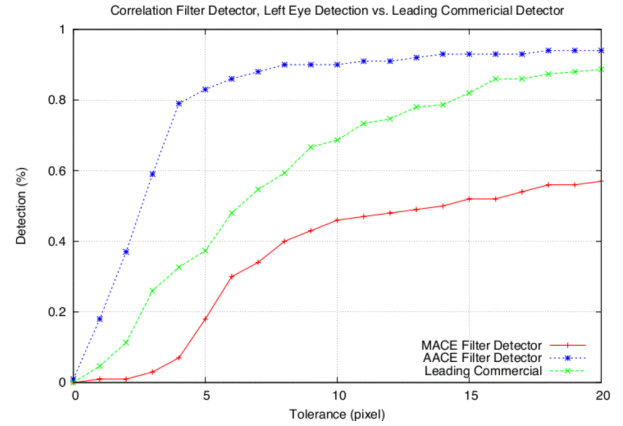

Fig. 11. Performance of the correlation filter detectors against a leading commercial detector for the left eye.
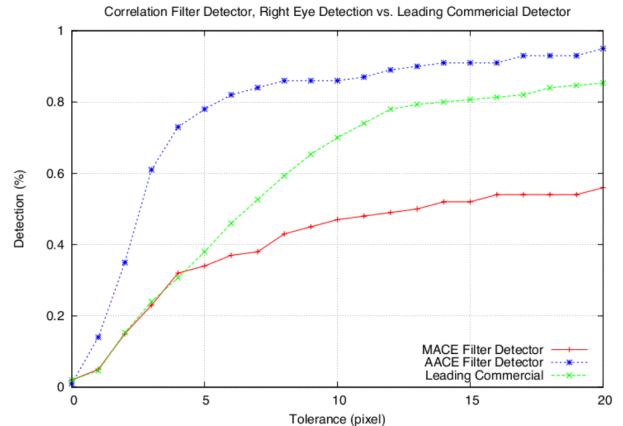

Fig. 12. Performance of correlation filter detectors against a leading commercial detector for the right eye.

VI. CONCLUSIONS

As face recognition moves forward, difficult imagery becomes a primary concern. But before we can even attempt face recognition, we often need to perform some necessary pre-processing steps, including geometric normalization and facial feature localization, with the eyes providing the necessary reference points. Thus, in this paper, we have concentrated on the eye detection problem for unconstrained environments. First, we introduced an EMCCD approach for low-light acquisition, and subsequently described an experimental protocol for simulating low-light conditions, distance, pose variation and motion blur. Next, we described two different detection algorithms: a novel machine learning based algorithm and a novel adaptive correlation filter based algorithm. Finally, using the data generated by our testing protocol, we performed a thorough series of experiments incorporating the aforementioned conditions. Both approaches
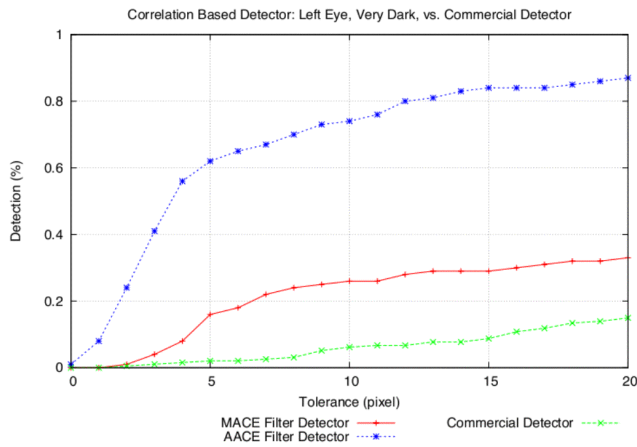
Fig. 13. Results for the correlation filter detectors for the left eye, with very dark imagery (0.0108 - 0.002 nits).
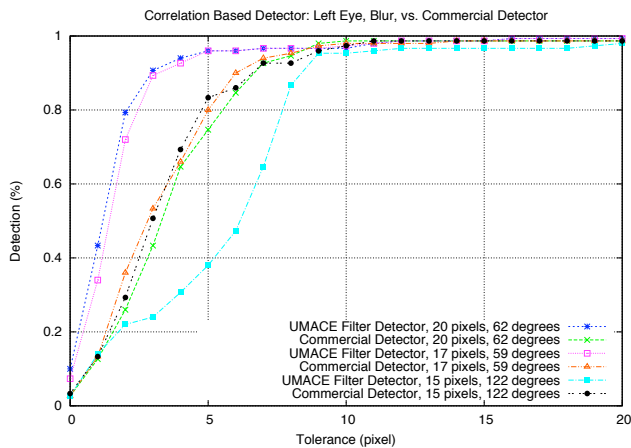

Fig. 14. Results comparing the AACE correlation filter detector to a leading commercial detector, for the left eye, with three varying degrees of blur. A blur length of 20 pixels, at an angle of 62 degrees was used to modify the input data. As can be seen in the plot, blur estimates closer to the test data perform well.

show significant performance improvement over a leading commercial eye detector. Comparing both approaches, our new AACE correlation filter detector shows a significant performance gain over the learning based detector on blurry imagery, and a modest performance gain on low-light imagery. Based on the presented results, we conclude that both approaches are suitable for the problem at hand - the choice of one as a solution can be based upon implementation requirements. As far as we know, this is the first study of feature detection under a multitude of difficult acquisition circumstances, and its results give us confidence for tackling the next steps for unconstrained face recognition.

VII. ACKNOWLEDGEMENTS

REFERENCES

[1] B. Leite, E. Pereira, H. Gomes, L. Veloso, C. Santos and J. Carvalho, "A Learning-based Eye Detector Coupled with Eye Candidate Filtering and PCA Features," In Proc. of the *XX Brazilian Symposium on Computer Graphics and Image Processing*, 2007

[2] L. Smith, A Tutorial on Principal Components Analysis, http://kybele.psych.cornell.edu/~edelman/Psych-465-Spring-2003/PCA-tutorial, 2002

[3] P. Viola and M. Jones, "Robust Real-time Face Detection", *Int. Journal of Computer Vision*, vol. 57, num. 2, pp. 137-154, 2004

[4] C. M. Bishop, "Pattern Recognition and Machine Learning", Springer, 1st edition, 2006.

[5] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression (PIE) Database", In Proc. of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002

[6] T.E. Boult, W.J. Scheirer, "Long Range Facial Image Acquisition and Quality", in M. Tistarelli, S. Li and R. Chellappa, editors, Biometrics for Surveillance and Security. Springer-Verlag, 2009.

[7] T.E. Boult and W. J. Scheirer and R. Woodworth, "FAAD: Face at a Distance", SPIE Defense and Security Symposium, Orlando FL, April 2008

[8] P.J. Phillips, H. Moon, S. Rizvi and P. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, num. 10, pp. 1090-1104, 2000

[9] T. Vogelsong, T. Boult, D. Gardner, R. Woodworth, R.C. Johnson and B. Heflin, "24/7 Security System: 60-FPS Color EMCCD Camera With Integral Human Recognition", Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VI. Edited by M. Carapezza, *Proceedings of the SPIE 6538* (2007), 65381S

[10] M. Savvides, B. Kumar, and P. Khosla, "Corefaces - Robust Shift Invariant PCA based Correlation Filter for Illumination Tolerant Face Recognition", in Proc. of the *IEEE Conference on Computer Vision and Pattern Recognition*, 2004

[11] D. Bolme, B. Draper, and J. R. Beveridge, "Average of Synthetic Exact Filters", in Proc. of the *IEEE Conference on Computer Vision and Pattern Recognition*, 2009. Available from: `http://www.cs.colostate.edu/~bolme/Bolme2009Asef.pdf`

[12] P.J. Phillips and Y. Vardi, "Efficient Illumination Normalization of Facial Images", *Pattern Recognition Letters* 17 (1996), no. 8, 921-927.

[13] T. Chen, W. Yin, X. Zhou, D. Comaniciu and T. Huang, "Total Variation Models for Variable Lighting Face Recognition", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 9, pp. 1519-1524, 2006.

[14] R. Tyson, "Introduction to adaptive optics", SPIE, The International Society for Optical Engineering, Bellingham, WA, 2000.

[15] J. Carroll, D. Gray, A. Roorda, D. Williams, "Recent Advances in Retinal Imaging with Adaptive Optics, *Optics and Photonics News*, pp. 36-42, 2007.

[16] Y. Yao, B. Abidi, N. Kalka, N. Schmid, M. Abidi, "Improving Long Range and High Magnification Face Recognition: Database Acquisition, Evaluation, and Enhancement", *Comput. Vis. Image Understand.* 2007, doi:10.1016/j.cviu.2007.09.004

[17] L. Jin, X. Yuan, S. Satoh, J. Li, and L. Xia, "A Hybrid Classifier for Precise and Robust Eye Detection", in *Proc. of the Int. Conf. on Pattern Recognition*, vol. 4, pp. 731–735, 2006.

[18] P. Wang, M.B. Green, Q. Ji, and J. Wayman, "Automatic Eye Detection and Validation", in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, 2005.

[19] A. Mahalanobis, B. Kumar, D. Casaent, "Minimum Average Correlation Energy Filters", *Appl. Opt.*, 26(17):3633, 1987.

[20] M. Savvides and B. Kumar, "Efficient Design of Advanced Correlation Filters for Robust Distortion-Tolerant Face Recognition", *AVSS*, pp. 45-52, 2003.

[21] R. Brunelli and T. Poggiot, "Template Matching: Matched Spacial Filters and Beyond", *Pattern Recognition*, vol. 30, num. 5, pp. 751-768, 1997.

[22] M. Savvides, R. Abiantun, J. Jeo, S. Park, C. Xie, B. Kumar, "Partial & Holistic Face Recognition on FRGC-II data using Support Vector Machine Kernel Correlation Feature Analysis", *IEEE Computer Society Workshop on Biometrics*, 2006.