FACE DETECTION IN MOBILE PHONES USING CO-OCCURRENCE OF ADJACENT LOCAL BINARY PATTERNS

Jeaff Wang, Dohyoung Lee, K.N. Plataniotis

Multimedia Lab, The Edward S. Rogers Department of Electrical and Computer Engineering,
University of Toronto, 10 King's College Road, Toronto, Canada

ABSTRACT

This paper investigates a novel combination of Co-occurrence of adjacent Local Binary Patterns histogram and Local Binary Patterns feature extraction methods for face detection in mobile phone applications. In particular, Co-occurrence of adjacent Local Binary Patterns histogram feature extraction provides exceptionally high discriminative power in face/non-face classification and hence is used to ensure the high accuracy of the proposed face detector. Local Binary Patterns feature extraction has low computation complexity and is thus used to reduce the overall processing speed. In the conducted face detection experiments, the proposed face detector yields comparable or better performance as well as faster computation speed than the existing best methods.

***Index Terms*—** Frontal face detection, Feature extraction, Local Binary Patterns, Mobile Phone

1. INTRODUCTION

Due to a growing market, face detection in mobile phone applications has gained much attention in both industry and research studies. Although many face detection algorithms exist, only a few were dedicated to the challenges introduced by mobile phones. Face detection under mobile phone environment is different from general cases [1]. In summary, it has the following advantages: (1) The face images are mostly frontal with less pose variations. (2) The face images are mostly in "close-range" with higher resolution. Thus, additional information is available and significantly less detection frames need to be processed. Conversely, the following challenges are introduced: (1) The mobility ensures that lighting conditions are highly variable, causing performance degradations to many existing face detection algorithms. (2) Compared to high performance devices such as PC and workstation, the computational resources on mobile phones are limited. However, smart-phones nowadays are better equipped compared to conventional mobile phones, allowing more sophisticated face detection algorithms to be developed.

Recently, one of the most popular face detection algorithms: Viola-Jones face detector [2] was successfully implemented on mobile phones [1, 3, 4, 5, 6], demonstrating accurate face detection in real-time. However, the simplicity of Haar-like features used in Viola-Jones face detector causes the performance to be limited under many complications, such as illumination variations. Alternatively, a significant amount of attention has been paid to Local Binary Patterns (LBP) feature. Compared to Haar-like features, LBP features exhibit higher discriminative power from texture patterns and more robustness against illumination variations; in addition, fewer arithmetics are required to extract LBP features.

In [7], Hadid et al had first successfully applied LBP histogram features in face detection, showing exceptionally high discriminative power in face/non-face classification. Many variants of LBP features were also considered for face detection. For example, Improved LBP was proposed by Jin et al [8] to capture more information using 9-bit codes by threshold on the mean calculated from 3×3 pixels neighbourhood. In [9], Zhang et al extracted texture patterns from neighbouring rectangular blocks in different scales to capture larger scale structures. Co-occurrence of multiple LBPs were considered in [10] to increase discriminative power for frontal face detection. Many other variations of LBP features were proposed in [11, 12, 13, 14, 15, 16]. Although LBP features have been widely studied for general purpose face detection, very few work has focused on its applications in mobile phones.

In this research study, we propose a novel combination of two LBP feature extraction variants: Co-occurrence of Adjacent Local Binary Patterns and Local Binary Patterns to enhance the face detection performance for mobile phone applications. The proposed face detector is compared to other state-of-the-art face detectors through face detection experiments.

The remaining of the paper is organized as follows: the proposed face detector design is discussed in Section 2; face detection experiments are presented in Section 3; computation analysis is discussed in Section 4; lastly, conclusion in Section 5.

2. THE PROPOSED FACE DETECTOR

The overall proposed face detector system is shown in Figure 1. To ensure processing speed, Boosting learning and Attentional Cascade structure based classification is used. LBP features are extracted for the following reasons: (1) Simplicity in computation for fast processing and to avoid complicated arithmetics. (2) High discriminative power ensures that only a small number of features are required for accurate face detection which results in faster processing and training. (3) Higher robustness against illumination variations than many other simple features due to threshold scheme.

Specifically, two types of complementary LBP features are extracted: original LBP features that are simple to extract for fast processing, as well as Co-occurrence of Adjacent LBP histogram features that are more expensive to extract, but have higher discrimina-

Acknowledges that this work is partially funded by MITACS Accelerate and Qualcomm Canada through a collaborative research project and was partially supported by the engineering research group from Qualcomm Canada.

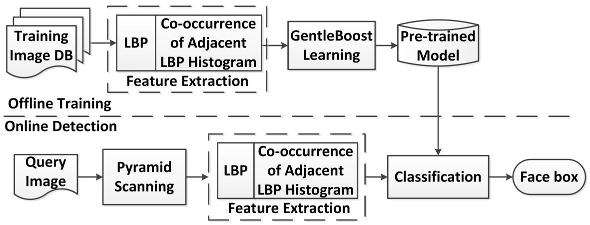


Fig. 1. Block diagram of the proposed face detector

tive power. Previous studies have explored such an approach; using dual LBP feature extractions to increase face detection performance was first proposed in [14] for surveillance applications.

2.1. Local Binary Patterns Feature Extraction

Local Binary Patterns can be extracted according to Equation 1 and only a few simple arithmetics are required.

$$LBP_{P,R}(x,y) = \sum_{p=0}^{P-1} \sigma(i_p - i_c)2^p \qquad (1)$$

where i_c is the center pixel at location (x, y) of the image, i_p are the eight surrounding pixels and $\sigma(x)$ is a threshold function.

The pattern extracted from every pixel is a feature and multi-branch decision tree is used for classification as in [9]. The desired number of features are selected using Gentle AdaBoost algorithm [9] to construct accurate face classifiers. Therefore, only a small amount of features are extracted for online face detection. Conversely, since the maximum number of LBP features is limited by the image's dimension, the discriminative power of LBP features is also limited.

2.2. Co-occurrence of Adjacent Local Binary Patterns Histogram Feature Extraction

Compared to the label based LBP features, LBP histogram features [7] demonstrate higher discriminative power for face/non-face classification. However, packing LBPs into histogram discards the spatial information between LBPs. In [17], Nosaka et al had demonstrated that by considering spatial relationships between LBP pairs, Co-occurrence of Adjacent LBP histogram is able to accurately express a richer set of texture patterns, hence providing even higher discriminative power compared to original LBP histogram features.

To extract Co-occurrence of Adjacent LBP, the encoding method proposed in [17] is used: Simplified $LBP(+)$ operator as illustrated in Equation 2 where i_p is the sampling points at radius S away from the center pixel i_c and $\sigma(x)$ is a threshold function.

$$LBP_S^+(x_c, y_c) = \sum_{p \in \{(\pm S,0)^T,(0,\pm S)^T\}} \sigma(i_p - i_c)2^p \qquad (2)$$

Four 2-D histograms are used to capture both local texture patterns and correlation between spatially adjacent features in four directions: $\{0, \Delta R\}^T, \{\Delta R, 0\}^T, \{\Delta R, \Delta R\}^T, \{-\Delta R, \Delta R\}^T$. The 2-D histogram is defined in Equation 3 where I is the image, r is every pixel in I and R is the distance between two adjacent LBPs. The final feature vector is produced by concatenating the four histograms where each histogram has length 16×16.

$$\begin{aligned} H_{i,j}(R) &= \sum_{r \in I} f_i(r) f_j(r+R)^T \\ f_i(r) &= \begin{cases} 1 & if\ LBP(r) = i \\ 0 & otherwise \end{cases} \end{aligned} \qquad (3)$$

Parameter values S and R in Equation 2 and 3 will affect stability and discriminative power of the extracted features. Relatively large values are able to target texture patterns in large scale, but they also decrease the amount of extracted texture patterns, causing instability to histogram. A range of values is suggested in [17] targeting the size of facial features: $S = \{1, \cdots, 5\}$ and $R = \{1, \cdots, 20\}$. Through face/non-face classification experiments, we have tested all parameter combinations and the most effective pair ($S = 1, R = 2$) is selected for the following face detection experiments.

For the proposed face detector, Co-occurrence of Adjacent LBP histogram features are extracted from 88×88 pixels since the face images are captured in ”close-range” for this particular application. Features are extracted in two configurations: 1. Global histogram extracted from the entire image; 2. Regional histograms extracted from nine 44×44 pixels regions overlapping 22×22 pixels between regions (proportionally same as in [7]).

For classification, each bin of the Co-occurrence of Adjacent LBP histogram is a feature and each weak classifier consists of only a single feature. Gentle AdaBoost algorithm [18] is used to select the best performing weak classifiers to construct strong classifier. Among the variants of Boosting algorithms, Gentle AdaBoost is proved to have higher performance for face detection [18].

A simple experiment is conducted to evaluate Co-occurrence of Adjacent LBP histogram feature in Boosting based face classification. A large face image dataset is used for this experiment containing variations including costumes, facial expression, pose and illumination conditions. Specifically, $9,916$ face images were collected by combining Viola and Jones dataset and Ole Jensen dataset [19]; as well as more than $100,000$ negative training images extracted from $2,000$ high resolution images collected from the Internet which contain no face. All the images are resized to 88×88 pixels. From the entire dataset, $7,916$ face images and $10,000$ non-face images were randomly selected for classifier training and another independent $2,000$ face images and $10,000$ non-face images were randomly selected for validation. The classification performance is evaluated using Performance Rate (PR) on the validation dataset. PR is defined to be the accuracy in the ROC space [20]: $PR = \frac{TP+TN}{N}$ where TP is the number of correctly classified faces, TN is the number of correctly classified non-faces and N is the total number of images.

Strong classifiers are trained using increasing numbers of weak classifiers, and the performance is compared between Co-occurrence of Adjacent LBP histogram feature and LBP histogram feature [7]. As shown in Figure 2, Co-occurrence of Adjacent LBP histogram feature has demonstrated very high discriminative power.

2.3. Cascade Classifiers

To ensure processing speed, face classifiers with different complexities are grouped in cascade stages, as shown in Figure 3. Each stage consists of a strong classifier trained using Gentle AdaBoost algorithm. A candidate window passes stage i if the classification output by classifier i is greater than the stage threshold Th_i. A detection window becomes face candidate if it passes all N stages. The

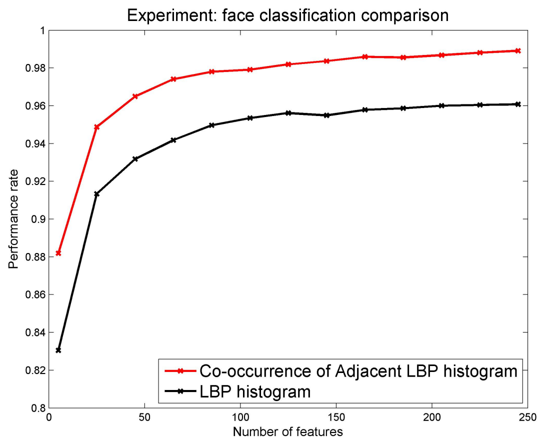


Fig. 2. Face classification experiment result

stages are arranged in increasing complexity, as the first few stages are trained using LBP feature with a limited number of weak classifiers, and increasing number of weak classifiers are used in succeeding stages. Co-occurrence of Adjacent LBP histogram feature is used in the last stage classifier, since it is very expensive to extract.

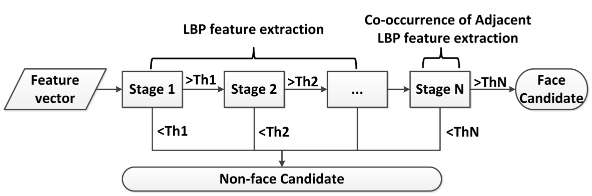


Fig. 3. Face classifiers in cascade

3. FACE DETECTION EXPERIMENTS

3.1. Face Detector Training

To train face detector, the same dataset and partitioning method as described in the face classification experiment in Section 2.2 is used. The baseline image size (24×24 pixels) is applied directly to LBP feature based classifiers training and the images are resized to 88×88 pixels for Co-occurrence of Adjacent LBP histogram feature based classifier training.

For high detection accuracy, the LBP feature based classifiers are trained to have 99.7% true positive rate (TPR) evaluated on validation dataset. The number of features and threshold for each stage is determined based on a pre-assigned false positive rate (FPR). For fast processing, a relatively high FPR is allowed in the initial stage by using only a small number of features. Each succeeding stage use increased number of features to reduce FPR by around 10% from its prior stage. However, decreasing FPR by 10% after the 4^{th} stage requires a large number of features and it was not practical. Hence, later stages are set to keep FPR at around 20%. The number of stages is increased until the desired overall performance is achieved.

The final stage is trained with Co-occurrence of Adjacent LBP histogram features to achieve minimum FPR by using minimum number of features. Specifically, 400 Co-occurrence of Adjacent LBP histogram features are selected to achieve around 99% TPR and 1% FPR. Overall, the face detector achieves 97.2% TPR and 5.76×10^{-6} FPR on validation dataset, as summarized in Table 1.

3.2. Face Detection Performance

The proposed face detector is evaluated with two face databases: BioID database [21] and the Extended Yale Face Database B [22]. The BioID database contains $1,521$ gray scaled images in 384×286 pixels. $1,521$ frontal "close-range" faces are recorded from 23 human subjects under "real world" conditions, having large variations in face expression, pose and costumes. The frontal pose subset of the Extended Yale Face Database B is used to evaluate the face detector under illumination complications. This subset contains $1,820$ gray scaled images in 640×480 pixels. Frontal "close-range" face images are captured from 28 human subjects under 64 controlled illumination conditions and other variations are minimized.

The correctness of face detection is evaluated on pre-determined ground truth. The ground truth is generated based on manually marked eye positions, as illustrated in Figure 5. A detection hypothesis is considered to be correct if it satisfies the following 2 criterions [11]: (1) The Euclidian distance between center of hypothesis face box and center of ground truth face box is less than 30% of the width of ground truth face box. (2) The width of hypothesis face box is within 50% of the width of ground truth face box.

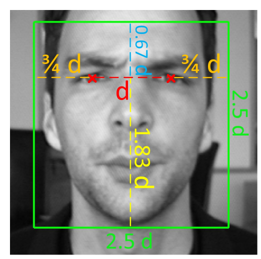


Fig. 5. Example of ground truth face box generation

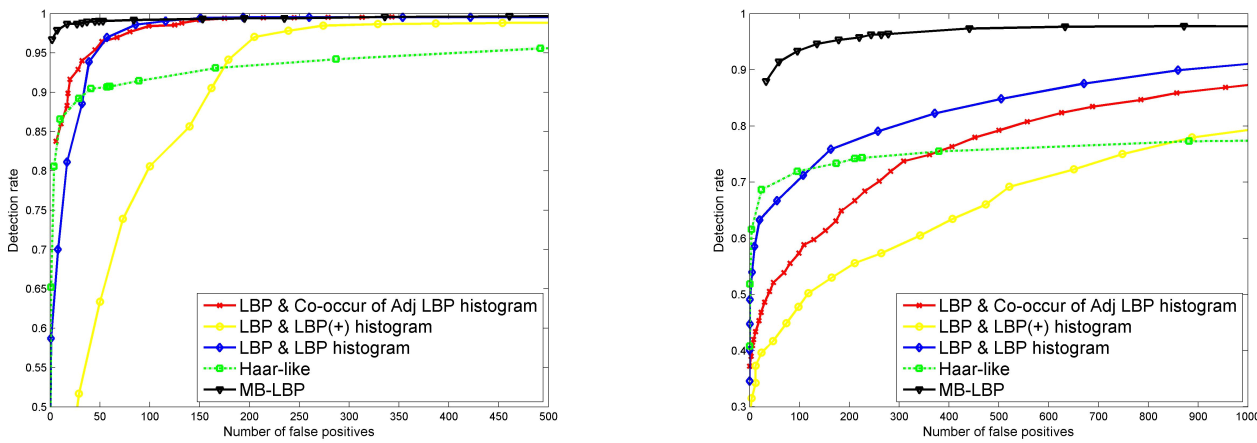
The proposed face detector (LBP & Co-occurrence of Adjacent LBP histogram features) is compared to other popular or best performing face detectors: MB-LBP face detector [9], Haar-like face detector [18] ($haarcascade\_frontalface\_alt.xml$) and LBP histogram face detector [7] (LBP & LBP histogram features). In addition, it is also compared to a simplified version without considering co-occurrence of LBPs (LBP & $LBP(+)$ histogram features).

Free Receiver-Operating-Characteristics (FROC) curves are obtained on BioID database to evaluate face detection under normal conditions, as shown in Figure 4 (a). In general, the LBP face detectors show better performance than the Haar-like face detector, suggesting LBP has exceptionally high discriminative power in face detection. The proposed face detector shows better performance than $LBP(+)$ histogram face detector, demonstrating higher accuracy by co-occurrence of adjacent LBPs. However, the proposed detector shows backward performance compared to MB-LBP face detector.

Similarly, FROC curves are obtained on The Extended Yale Face Database B to evaluate face detection under controlled illumination

	stage 1	stage 2	stage 3	stage 4	stage 5	stage 6	stage 7	Overall
Number of features	10	20	35	60	120	180	400	-
True Positive Rate	99.7%	99.7%	99.7%	99.7%	99.7%	99.7%	99.0%	97.2%
False Positive Rate	56.7%	40.1%	33.6%	18.6%	19.1%	23.6%	0.9%	5.76×10^{-6}

Table 1. Face detector stages evaluated on validation dataset. True Positive Rate $TPR = \frac{TruePositive}{N_f}$ where true positive is correctly classified face images and N_f is the number of face images in the dataset. False Positive Rate $FPR = \frac{FalsePositive}{N_{nf}}$ where false positive is misclassified non-face images as face and N_{nf} is the number of non-face images in the dataset.



(a) FROC curve using BioID face database

(b) FROC curve using Extended Yale Face Database B

Fig. 4. Face detection results; Detection Rate is $\frac{TruePositives}{N}$ where True Positives is the number of correct positive hypotheses and N is the total number of ground truth; False Positive is the number of incorrect positive hypotheses

variations, as shown in Figure 4 (b). The LBP face detectors show higher performance than the Haar-like face detector, demonstrating the robustness of LBP feature under illumination variations. Again, the proposed face detector shows better performance than $LBP(+)$ histogram face detector by considering co-occurrence of adjacent LBPs. However, the proposed face detector shows backward performance compared to the MB-LBP face detector and LBP histogram face detector by using simplified $LBP(+)$ operator. On the other hand, the simplified $LBP(+)$ operator contributes to fast processing speed of the proposed face detector.

4. COMPUTATION TIME AND MEMORY SIZE

To compare computation speed of face detectors, we first model the average number of calculations (additions & comparisons) per scanning window required by each face detector. The average is calculated based on the complexity of each stage weighted by the cumulative false positive rate of that stage (as indicated by Table 1). To verify the above model in a "real" implementation, we also measure the average time to process each image from $2,000$ images in 640×480 pixels using MATLAB on a 2.67 GHz Intel Core i7 M620 PC system with 8 GB RAM. Although the MATLAB implementations are slow due to software overheads, the obtained results are proportional to the calculations from the above model, as shown in Table 2. From the results, the proposed face detector has the fastest speed and it is about three times faster than the Haar-like face detector.

	Average number of operations	**MATLAB** (unit: seconds)
Proposed	792	5.61
LBP Histogram	20,292	135.86
Haar-like	2,308	16.92
MB-LBP	1,616	13.66

Table 2. Computation time comparison between face detectors

5. CONCLUSION

In this paper, we have introduced a novel combination of LBP and Co-occurrence of Adjacent LBP histogram features in duality for face detection under mobile phone applications. The proposed method shows fast processing and accurate detection in conducted experiments. Compared to Haar-like and other LBP based methods, the proposed method shows higher or comparable performance under normal and challenging illumination. Although MB-LBP method shows higher accuracy, the proposed method has faster speed by extracting simple LBPs. Thus, the proposed method could be more suitable for low-powered devices such as mobile phones. In addition, MB-LBP and the proposed method enhance the original LBP from different approaches, which may combine to produce better results and hence it could be investigated in the future.

6. REFERENCES

[1] A. Hadid, J.Y. Heikkila, O. Silven, and M. Pietikainen, "Face and eye detection for person authentication in mobile phones," in *Proceedings of the First ACM/IEEE International Conference on Distributed Smart Cameras (ICDSC-07)*, 2007, pp. 101–108.

[2] Paul Viola and Michael Jones, "Robust real-time object detection," in *International Journal of Computer Vision*, 2001.

[3] Sung-Uk Jung, Yun-Su Chung, Jang-Hee Yoo, and Ki-Young Moon, "Real-time face verification for mobile platforms," in *Proceedings of the 4th International Symposium on Advances in Visual Computing, Part II*, Berlin, Heidelberg, 2008, pp. 823–832.

[4] Bin Chen, Jie Shen, and Helei Sun, "A fast face recognition system on mobile phone," in *Proceedings of the 2012 International Conference on Systems and Informatics (ICSAI)*, may 2012, pp. 1783 –1786.

[5] Qian Tao and Raymond Veldhuis, "Biometric authentication system on mobile personal devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 59, no. 4, pp. 763–773, 2010.

[6] Chee Kiat Ng, Marios Savvides, and Pradeep K. Khosla, "Real-time face verification system on a cell-phone using advanced correlation filters," in *Proceedings of the Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, Washington, DC, USA, 2005, pp. 57–62, IEEE Computer Society.

[7] Abdenour Hadid, Matti Pietikäinen, and Timo Ahonen, "A discriminative feature space for detecting and recognizing faces," in *Proceedings of the 2004 IEEE computer society conference on Computer vision and pattern recognition*, Washington, DC, USA, 2004, CVPR'04, pp. 797–804, IEEE Computer Society.

[8] Hongliang Jin, Qingshan Liu, Hanqing Lu, and Xiaofeng Tong, "Face detection using improved lbp under bayesian framework," in *Proceedings of the Third International Conference on Image and Graphics*, Washington, DC, USA, 2004, ICIG '04, pp. 306–309, IEEE Computer Society.

[9] Lun Zhang, Rufeng Chu, Shiming Xiang, Shengcai Liao, and StanZ. Li, "Face detection based on multi-block lbp representation," in *Advances in Biometrics*, Seong-Whan Lee and StanZ. Li, Eds., vol. 4642 of *Lecture Notes in Computer Science*, pp. 11–18. Springer Berlin Heidelberg, 2007.

[10] Wael Louis and Konstantinos N. Plataniotis, "Co-occurrence of local binary patterns features for frontal face detection in surveillance applications," *EURASIP J. Image and Video Processing*, vol. 2011, 2011.

[11] Zhe Wei, Yuan Dong, Feng Zhao, and Hongliang Bai, "Face detection based on multi-scale enhanced local texture feature sets," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Kyoto, Japan, Mar. 2012.

[12] Bernhard Fröba and Andreas Ernst, "Face detection with the modified census transform," in *Proceedings of the Sixth IEEE international conference on Automatic face and gesture recognition*, Washington, DC, USA, 2004, FGR' 04, pp. 91–96, IEEE Computer Society.

[13] Bongjin Jun and Daijin Kim, "Robust face detection using local gradient patterns and evidence accumulation," *Pattern Recognition*, vol. 45, no. 9, pp. 3304 – 3316, 2012.

[14] Wael Louis and Konstantinos N. Plataniotis, "Frontal face detection for surveillance purposes using dual local binary patterns features," in *Proceedings of the International Conference on Image Processing, ICIP 2010, September 26-29, Hong Kong, China*, 2010, pp. 3809–3812.

[15] Shengye Yan, Shiguang Shan, Xilin Chen, and Wen Gao, "Locally assembled binary (lab) feature with feature-centric cascade for fast and accurate face detection," in *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, june 2008, pp. 1–7.

[16] Anindya Roy and Sbastien Marcel, "Haar local binary pattern feature for fast illumination invariant face detection the proposed framework : Face detection using hlbp features," *British Machine Vision Conference*, vol. 29, pp. 1–12, 2009.

[17] Ryusuke Nosaka, Yasuhiro Ohkawa, and Kazuhiro Fukui, "Feature extraction based on co-occurrence of adjacent local binary patterns," in *Proceedings of the 5th Pacific Rim conference on Advances in Image and Video Technology - Volume Part II*, Berlin, Heidelberg, 2012, PSIVT'11, pp. 82–91, Springer-Verlag.

[18] Rainer Lienhart and Jochen Maydt, "An extended set of haar-like features for rapid object detection," in *Proceedings of the Internaitonal Conference on Image Processing, ICIP 2002*, 2002, pp. 900–903.

[19] Ole H. Jensen, "Implementing the Viola-Jones Face Detection Algorithm," M.S. thesis, Technical University of Denmark, Department of Informatics and Mathematical Modeling, 2008.

[20] Tom Fawcett, "An introduction to roc analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[21] Oliver Jesorsky, Klaus J. Kirchberg, and Robert Frischholz, "Robust face detection using the hausdorff distance," in *Proceedings of the Third International Conference on Audio- and Video-Based Biometric Person Authentication*, London, UK, UK, 2001, AVBPA '01, pp. 90–95, Springer-Verlag.

[22] Kuang chih Lee, Jeffrey Ho, and David Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 684–698, 2005.