Learning from Partially Occluded Faces

Fares Al-Qunaieer[1] and Mohamed Alkanhal[2]
[1]*The National Center for Computation Technology and Applied Mathematics, King AbdulAziz City for Science and Technology (KACST), Riyadh, Saudi Arabia*
[2]*Communication and Information Technology Research Institute, King AbdulAziz City for Science and Technology (KACST), Riyadh, Saudi Arabia*

Keywords: Face Recognition, Partial Occlusion, Correlation Filters.

Abstract: Although face recognition methods in controlled environments have achieved high accuracy results, there are still problems in real-life situations. Some of the challenges include changes in face expressions, pose, lighting conditions or presence of occlusion. There were several efforts for tackling the occlusion problem, mainly by learning discriminating features from non-occluded faces for occluded faces recognition. In this paper, we propose the reversed process, to learn from the occluded faces for the purpose of non-occluded faces recognition. This process has several useful applications, such as in suspects identification and person re-identification. Correlation filters are constructed from training images (occluded faces) images of each person, which are used later for the classification of input images (non-occluded faces). In addition, the use of skin masks with the correlation filters is investigated.

1 INTRODUCTION

Biometric methods for authentication and identification have become a part of our daily life. There are many available approaches for biometric systems, face recognition is considered an important approach among them. Face images can be captured in many ways using standard cameras. Furthermore, efficient algorithms have been used for face recognition. This make the identification and verification of people by their faces very accessible. However, many of the proposed algorithms were designed for controlled settings. Changes in face expressions, pose, lighting conditions or presence of occlusion can dramatically affect the results (Li and Jain, 2011).

There has been much research conducted to solve the problem of face occlusion. The used techniques span a wide variation of concepts, such as Principle Component Analysis (PCA) (Sharma et al., 2013)(Rama et al., 2008), feature-based learning (Sharma et al., 2013)(Zhang et al., 2007), correlation filters (Kumar et al., 2006), sparse representation (Wright et al., 2009)(Zhou et al., 2009)(Liao et al., 2013), and face completion (Deng et al., 2009).

All these works mainly performed by learning discriminating features from non-occluded faces for the purpose of recognizing occluded faces. An interesting question is what about learning from occluded faces to recognize non-occluded ones? There are several applications that can benefit from such setting. For instance, it can be used for person re-identification purposes, in which a person with occluded face can be tracked, even when the occlusion is eliminated. Another useful application is to identify suspects in public or private places (e.g, banks, airports). The top n suspects can be identified for further investigations. Also, the system can be trained from the occluded faces and set to actively monitor people. An alert can be issued if a face matched the trained one (i.e., the occluded face). In addition to these applications, designing the system this way increases the computation and storage efficiency as will be described.

The main contribution of this paper is the introduction of a new paradigm, where the goal is to identify non-occluded faces by learning from occluded ones. To the best of our knowledge, no previous research was conducted on the learning from occluded faces as described here.

In this paper, Optimal Trade-off Maximum Average Correlation Height (OT-MACH) correlation filter was used. In addition, masked OT-MACH was investigated, in which, a mask is constructed based on skin color, and the correlation filter is built based on the skin region. This mask is averaged for all images of a person and stored along the constructed filter. An input image is multiplied with the skin mask, then cor-

Al-Qunaieer, F. and Alkanhal, M.
Learning from Partially Occluded Faces.
DOI: 10.5220/0005665605340539
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 534-539
ISBN: 978-989-758-173-1

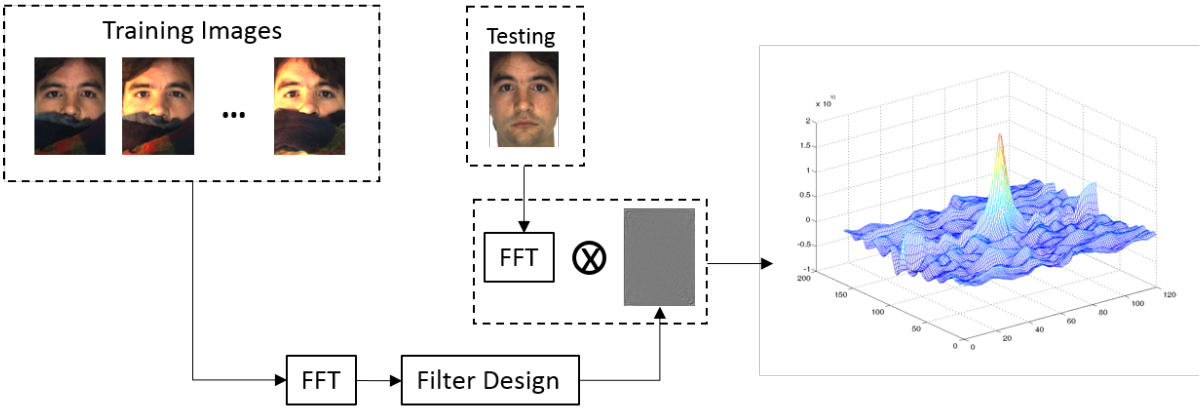

Figure 1: Block diagram of the proposed approach.

related with the filter.

The remaining of the paper is organized as follows: next section will describe OT-MACH correlation filters. After that, the proposed method is described in Section 3. The experiments and results are discussed in Section 4. Finally, conclusions are presented in Section 5.

2 CORRELATION FILTERS

Correlation filters have been successfully used in several applications, such as biometrics and object detection and recognition. The basic idea is to design filters through learning, which gives high correlation peaks for objects of interest and low peaks otherwise.

In this section, we represent in the frequency domain an image $x(m,n)$ of size $d \times d$ as a $d^2 \times d^2$ matrix $\mathbf{X}$ with the elements of $\mathbf{x}$ along its diagonal. The superscripts * and $^+$ represent the conjugate and conjugate transpose, respectively.

Maximum Average Correlation Hight (MACH) filter is a class of correlation filters designed to maximize the correlation peak intensity as a response to the average training images. This is performed by using a metric known as the Average Correlation Height (ACH) expressed as

$$ACH_x = |\mathbf{h}^+\mathbf{m}_x|^2, \tag{1}$$

where $\mathbf{m}_x$ is the average of N training images from class Ω_x in the frequency domain. The column vector $\mathbf{h}$ represents the correlation filter. In MACH filter design, a metric known as the Average Similarity Measure (ASM) is minimized to maximize the distortion tolerance. The ASM is defined as

$$ASM_x = \mathbf{h}^+\mathbf{S}_x\mathbf{h}, \tag{2}$$

where

$$\mathbf{S}_x = \frac{1}{N}\Sigma_{i=1}^{N}(\mathbf{X}_i - \mathbf{M}_x)^*(\mathbf{X}_i - \mathbf{M}_x), \tag{3}$$

and $\mathbf{M}_x$ is a diagonal matrix containing $\mathbf{m}_x$.

The MACH filter is designed to maximize the ratio $\frac{ACH_x}{ASM_x}$. This leads to the following form for the MACH filter

$$\mathbf{h} = \mathbf{S}_x^{-1}\mathbf{m}_x. \tag{4}$$

The MACH filter can be extended to the Optimal Trade-Off (OT)-MACH filter, in which there is a trade-off among distortion tolerance, discrimination ability and noise stability. The OT-MACH filter can be written in the following form (Kumar et al., 1994)

$$\mathbf{h} = (\alpha\mathbf{D}_x + \beta\mathbf{S}_x + \gamma\mathbf{C})^{-1}\mathbf{m}_x, \tag{5}$$

where $\mathbf{C}$ is a diagonal matrix modelling the power spectral of the noise, which is usually considered as white (*i.e.*, $\mathbf{C} = \mathbf{I}$), and $\mathbf{D}_x$ is a diagonal matrix containing the average power spectrum of the N training images. The parameters α, β and γ are scalers that control the importance of the three different terms.

3 PROPOSED METHOD

Unlike the usual setting, where the learning is conducted on known people, our approach will learn from unknown people (occluded faces) and try to find the best match from non-occluded faces. In this paper, we constructed OT-MACH correlation filter from occluded faces as described in Section 2 and illustracted in Figure 1. The goal is to obtain high correlation peak if an image of the same person with non-occluded face is correlated with the filter.

Also, we investigated the use of skin mask in the process of designing and applying OT-MACH filters, as shown in Figure 2. Statistical color models for skin and non-skin (Jones and Rehg, 1999) was used to detect the most probable skin location, resulting in a skin mask image (e.g., skin=1, non-skin=0). Because

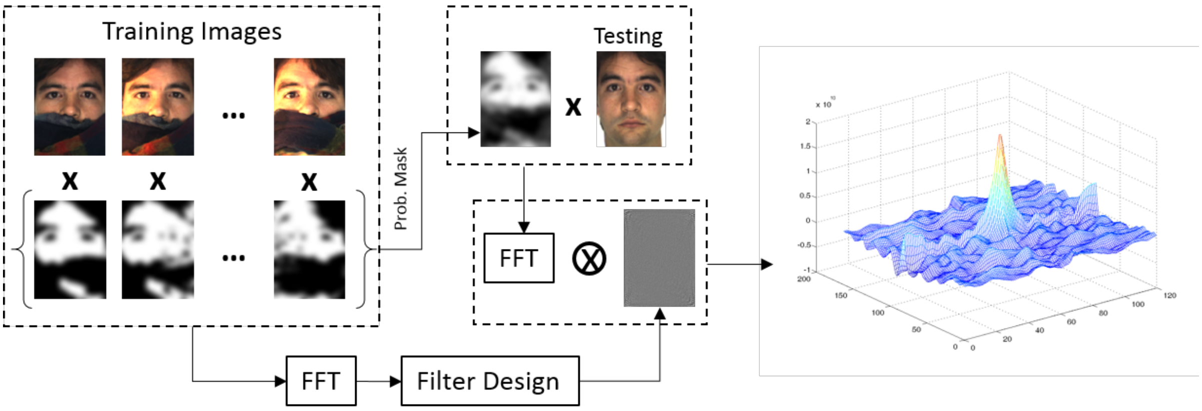

Figure 2: Block diagram of the proposed approach with skin mask.

Figure 3: Occlusion types. Top: images with scarf, bottom: images with sun glasses.

the presence of some errors in skin detection, the detected skin locations were smoothed by a Gaussian filter. For each person, all these skin locations are averaged to construct an averaged skin mask. OT-MACH filter is created for each person from the masked training images as described in Section 2. Each input image is multiplied with the skin mask. This will ensure that the filtering will be performed on the same parts of training, which will only work with the assumption that all training and testing images are properly warped to the same locations as in the cropped AR faces (Martinez and Kak, 2001). After masking the input image, it is correlated with the constructed filter.

The correlation output is evaluated by the sharpness and hight of the resulting peaks. This can be quantified by the Peak-to-Sidelope Ratio (PSR) as follows (Kumar et al., 2006)

$$PSR = \frac{p - \mu}{\sigma}, \tag{6}$$

where p is the peak of the correlation output and μ and σ are the mean and standard deviation of the correlation values, respectively. Here, PSR is computed with the exclusion of a small window of size 5×5 centered at the peak. An image with PSR above a specified threshold is classified as genuine, while a one below the threshold is classified as imposter.

4 EXPERIMENTS AND RESULTS

In this research, the cropped version (Martinez and Kak, 2001) of AR face database (Martinez and Benavente, 1998) was used to verify the proposed approach. It consists of a total of 2600 images, 26 images per person for 100 people. Each person has images taken in different expressions, lightings, and occlusions. The images have been taken in two sessions.

In the experiments, the correlation filters were constructed using only occluded images. The parameters of OT-MACH were empirically selected to be α = 1.4 and β = 1.0. Images in the training and test stages are converted to gray-scales. There are two types of occlusions in the dataset, scarf and sunglasses (6 images of each, with different lighting directions). Therefore, two correlation filters were constructed for each person. Figure 3 illustrates both occlusion types for one person.

Table 1: AUC of classifiers trained with faces occluded with Scarf and Sun Glasses, and tested on *Neutral* and *Expressions* faces.

	OT-MACH	OT-MACH with Mask
Scarf (Neutral)	0.95	0.95
Scarf (Expressions)	0.91	0.90
Sun Glasses (Neutral)	0.97	0.87
Sun Glasses (Expressions)	0.92	0.80

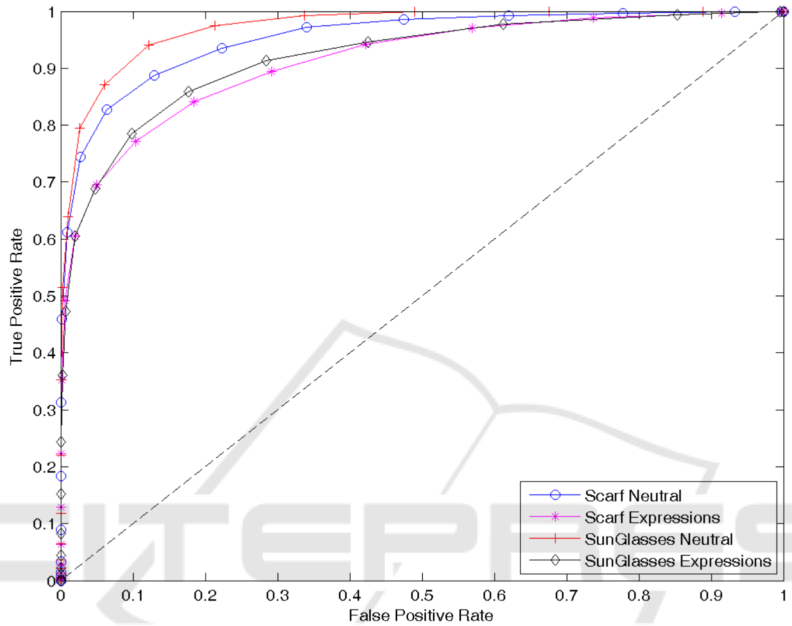

Figure 4: ROC curves for OT-MACH filters trained with faces occluded by Scarf and Sun Glasses, and tested on *Neutral* faces and *Expressions* faces.

For each person, the two correlation filters of scarf and sun glasses were correlated with the images of all the people in the dataset. This is performed in two ways, the first is to use only the *Neutral* faces (no expressions), the other is to use all images except the occluded ones (*Expressions*). Receiver Operating Characteristics (ROC) graph is created for all the results by varying the threshold from 1 to 25. ROC curve is created by plotting True Positive Rate (TPR) against False Positive Rate (FPR) defined as

$$TPR = \frac{TP}{TP+FN} \tag{7}$$

and

$$FPR = \frac{FP}{FP+TN}, \tag{8}$$

where TP, FN, FP and TN are true positive, false negative, false positive and true negative, respectively. In addition, the Area Under the ROC Curve (AUC) is calculated for each experiment as a summary performance measure. These experiments were performed for each person, and the average performance is calculated. Figures 4 and 5 show the ROC graphs for the four experiments for OT-MACH correlation filters with and without skin masks, and their AUC are presented in Table 1.

Figures 4 and 5 show the trade-off between TPR and FPR with respect to the selected threshold, which facilitate selecting the most appropriate threshold for certain applications. It is clear that using skin masks has reduced the accuracy of OT-MACH results, especially when trained with sun glasses occlusion. This might be due to loss of information when using skin masks. In addition, it can be noticed that both scarf and sun glasses have better accuracy with Neutral images than Expressions. This is because the high variability in Expressions faces set. OT-MACH correlation filters results are robust against illumination variation presented in the used dataset.

Figure 6 illustrates the response of the correlation filters of both scarf and sun glasses images for a genuine person (positive) and an imposter one (negative).

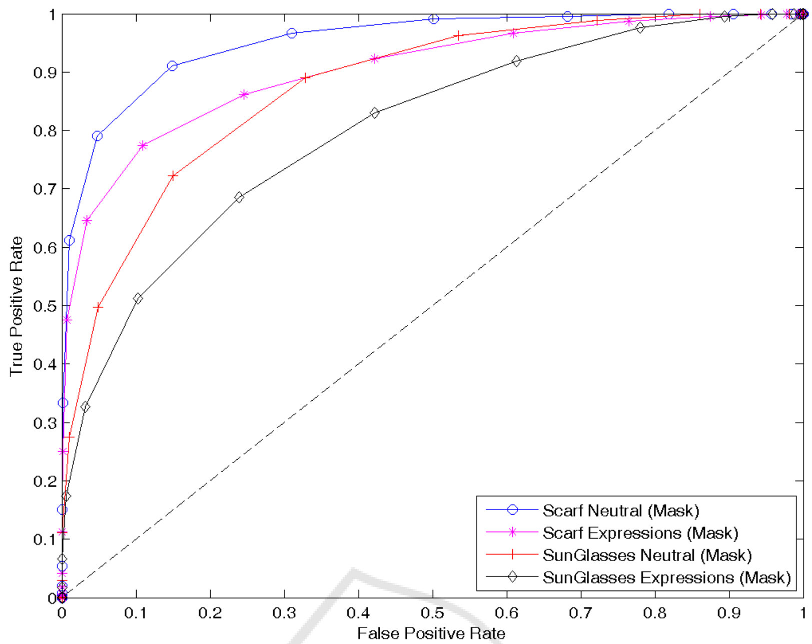

Figure 5: ROC curves for OT-MACH filters using skin masks trained with faces occluded by Scarf and Sun Glasses, and tested on *Neutral* faces and *Expressions* faces.

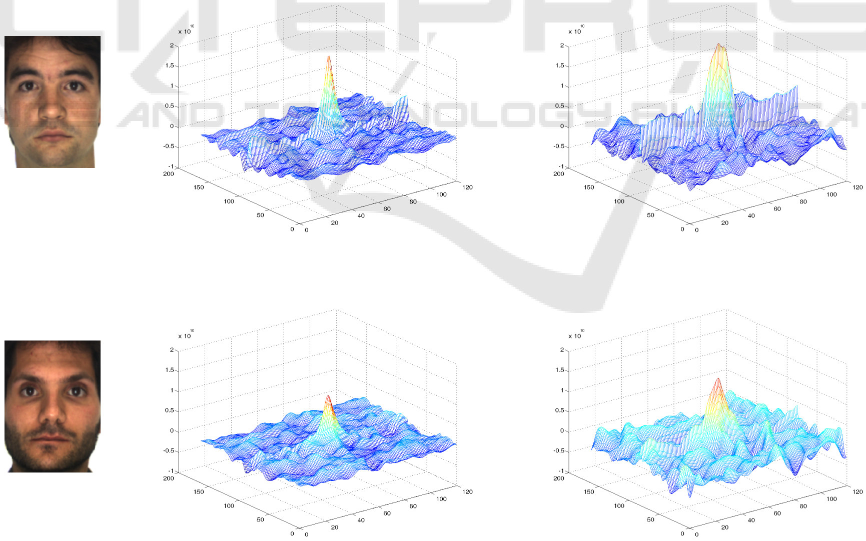

Figure 6: Correlation filters outputs; top: genuine, bottom: imposter. Columns from left to right: test images, outputs of scarf filter, outputs of sun glasses filter.

It can be observed that the peaks of the genuine response is higher than that of the Imposter.

Learning only from occluded faces can enhance the computation and storage efficiency. For instance, in this research, instead of learning form 100 people to identify one person with occluded face, we only learn from the occluded faces then make the comparison to all people. In this way, only one filter is saved instead of 100.

5 CONCLUSIONS

In this paper, we considered the problem of learning from occluded faces for the purpose of recognizing non-occluded ones. OT-MACH correlation filters were used for classification. In addition, the use of skin masks was investigated. Using OT-MACH without skin masks showed better results than that with masks. Also, faces with neutral expressions have exhibited better results compared to faces with variable expressions. The proposed approach can be used in applications that require to identify suspects (e.g., in crime) for further investigation. Also, it can be used for person re-identification purposes. There is still a room for future work, such as accounting for pose variation in both training and testing images.

REFERENCES

Deng, Y., Li, D., Xie, X., Lam, K.-M., and Dai, Q. (2009). Partially occluded face completion and recognition. In *16th IEEE International Conference on Image Processing (ICIP)*.

Jones, M. and Rehg, J. (1999). Statistical color models with application to skin detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*.

Kumar, B., Savvides, M., and Xie, C. (2006). Correlation pattern recognition for face recognition. *Proceedings of the IEEE*.

Kumar, B. V. K. V., Mahalanobis, A., and Carlson, D. W. (1994). Optimal trade-off synthetic discriminant function filters for arbitrary devices. *Opt. Lett.*, 19(19):1556–1558.

Li, S. Z. and Jain, A. K. (2011). *Handbook of Face Recognition*. Springer London, 2nd edition.

Liao, S., Jain, A., and Li, S. (2013). Partial face recognition: Alignment-free approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Martinez, A. and Benavente, R. (1998). The AR face database. Technical Report CVC Technical Report No.24, Universitat Autònoma de Barcelona.

Martinez, A. M. and Kak, A. (2001). Pca versus lda. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Rama, A., Tarres, F., Goldmann, L., and Sikora, T. (2008). More robust face recognition by considering occlusion information. In *8th IEEE International Conference on Automatic Face Gesture Recognition*.

Sharma, M., Prakash, S., and Gupta, P. (2013). An efficient partial occluded face recognition system. *Neurocomputing*, 116:231–241.

Wright, J., Yang, A., Ganesh, A., Sastry, S., and Ma, Y. (2009). Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Zhang, W., Shan, S., Chen, X., and Gao, W. (2007). Local gabor binary patterns based on kullback-leibler divergence for partially occluded face recognition. *Signal Processing Letters, IEEE*.

Zhou, Z., Wagner, A., Mobahi, H., Wright, J., and Ma, Y. (2009). Face recognition with contiguous occlusion using markov random fields. In *IEEE 12th International Conference on Computer Vision*.