MASKED CORRELATION FILTERS FOR PARTIALLY OCCLUDED FACE RECOGNITION

Eric J. He[1], Joseph A. Fernandez[1], B. V. K. Vijaya Kumar[1], Mohamed Alkanhal[2]

[1]Dept. of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, USA
[2]King Abdulaziz City for Science and Technology (KACST), Saudi Arabia

ABSTRACT

Face recognition is widely used for a variety of applications, such as identifying people for security purposes, as well as photo album organization. A challenge is to perform accurate face recognition when there exist partial occlusions of the face such as scarves or sunglasses. Correlation Filters (CFs) are an occlusion-tolerant object recognition method, potentially suited to deal with partial occlusions. In this paper, we introduce a new class of correlation filters called *Masked Correlation Filters* (MCFs), that are designed specifically to handle partial occlusions in face images. The benefits of using MCFs are illustrated using well-known face image data sets.

***Index Terms*—** Correlation Filters, Masked Correlation Filters, Face Recognition, Partial Occlusion.

1. INTRODUCTION

Face recognition (FR) can be useful in a wide variety of security and commercial applications. One of the challenges for real-world FR is partial occlusions of the face, such as scarves and sunglasses. Since the work of Martinez [1], multiple occlusion-tolerant FR methods have been proposed. Yang et al. [2] used Gabor features for FR tolerant to partial occlusions, while Luan et al. [3] used Robust Principal Component Analysis to perform FR under varying occlusion and illumination.

Kumar et al. demonstrated that Correlation Filters (CFs) are a useful method with attractive properties for FR [4]. CFs also exhibit graceful degradation, meaning that as parts of the image are occluded, the CF's performance degrades slowly. Recently, a new CF design known as Zero Aliasing Correlation Filters (ZACFs) [5, 6] has been proposed to address the circular correlation effects in CF implementations and it has been shown that ZACFs lead to enhanced performance in object recognition problems, including FR.

In this paper, we introduce *Masked Correlation Filters* (MCFs), which are a new type of CF inspired by ZACFs, aimed at the problem of image occlusion. MCFs use prior knowledge of potentially occluded regions (e.g., facial regions that will be occluded by scarves, sunglasses, etc. as shown in Fig. 1), in order to not use the unreliable face image information in the occluded regions. This is done by forcing the occluded regions to zero in the CF template during the filter design stage, which allows the filter design to instead optimize over the unoccluded regions. In contrast, traditional CF formulations yield templates that are not necessarily zero in the occluded regions, which results in increased noise sensitivity.

Fig. 1. Examples of Occluded Faces

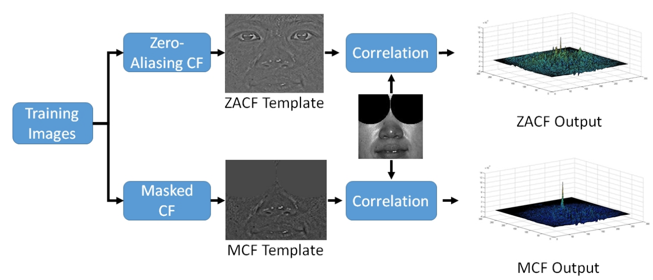


Fig. 2. Overview of our proposed Masked Correlation Filters (MCFs) approach. Conventional CF designs (upper part) result in templates that are non-zero for all values within the image region, making it more susceptible to noise in occluded regions. In MCFs (bottom part), constraining the occluded regions to zero minimizes the noise sensitivity in the occluded regions, resulting in a less noisy correlation output.

The rest of the paper is organized as follows. In Section 2, we briefly review CFs and ZACFs. In Section 3, we introduce MCFs and show how they can be used for occlusion-tolerant recognition. In Section 4, we present experimental results for partially occluded FR. Section 5 has our conclusions.

2. CORRELATION FILTERS

We provide a brief review of CFs and use the Minimum Average Correlation Energy (MACE) [7] filter as an illustrative example. The MACE filter is designed to produce sharply-peaked correlation outputs with pre-specified values at those peaks in response to centered authentic training images.

Let us assume that we train a filter using Q training images of size $N_{x,1} \times N_{x,2}$. Typically, we determine the filter $\bar{\mathbf{h}}$ in the frequency domain using the vectorized 2D DFTs (of size $N_1 \times N_2$,$N_1 \geq N_{x,1}$and $N_2 \geq N_{x,2}$) of the training images, $\bar{\mathbf{x}}_q$, for $q = 1, \ldots, Q$. In our notation, a symbol with an overbar represents a frequency domain quantity and a symbol with no overbar denotes a space domain quantity. The solution to the MACE filter is given by [7]

$$\bar{\mathbf{h}} = \mathbf{D}^{-1}\bar{\mathbf{X}}\left(\bar{\mathbf{X}}^{+}\mathbf{D}^{-1}\bar{\mathbf{X}}\right)^{-1}\mathbf{u} \tag{1}$$

where superscript $^+$ denotes the complex conjugate transpose, $\bar{\mathbf{X}}$ is a matrix with Q columns with its q-th column containing the vectorized version of the 2D DFT of the q-th training image. The diagonal matrix $\mathbf{D} = \frac{1}{N_1 N_2 Q}\sum_{q=1}^{Q}\bar{\mathbf{X}}_q\bar{\mathbf{X}}_q^{+}$ where the diagonal matrix $\bar{\mathbf{X}}_q$ contains the vectorized 2D DFT of the q-th training image along its diagonal, and $\mathbf{u}$ is the correlation peak constraint vector, whose elements are typically set to 1 for positive class training images and to 0 for negative class training images (if used). The MACE filter and other conventional CF implementations have an important limitation: as the correlation output is obtained as the inverse DFT of the element-wise product of two DFTs, a circular correlation will be obtained in the space domain. Circular correlation is an aliased version of the desired linear correlation, and thus degrades the correlation peaks. In order to solve this problem, a change in CF template design known as ZACF was recently introduced [5, 6].

The Zero Aliasing MACE (ZAMACE) filter formulation, is similar to the MACE filter formulation except that we introduce zero aliasing (ZA) constraints that force the template's tail to zero. For example, for a 1D template $h(n)$, we can express these constraints as $h(n) = 0$ for $n \geq N_x$. The expression for ZAMACE filter and its performance results are well summarized in [6].

Note that the Optimal Trade-off Synthetic Discriminant Function (OTSDF) filter is a CF that is similar to the MACE filter. In the OTSDF formulation, the matrix $\mathbf{D}$ is replaced by $\mathbf{T} = \mathbf{D} + \delta\mathbf{I}$, where $\mathbf{I}$ is an identity matrix and $\delta > 0$. The inclusion of the identity matrix is to improve noise tolerance. The Unconstrained OTSDF (UOTSDF) filter is a CF in which we remove the peak constraints, making the filter computation less computationally complex.

3. MASKED CORRELATION FILTER (MCF)

We propose a new CF formulation, in order to create a filter which is more tolerant to partial occlusions, assuming prior knowledge of what regions will be occluded. Fig. 2 illustrates the concepts of MCFs as compared with the basic concept of CFs for FR. MCFs are best used when there are unoccluded training images, and partially occluded test images in which the regions that would be occluded are known a priori. Also, a separate detector (e.g., sunglasses detectors or mouth scarf detectors) can be employed prior to the recognition step. In MCFs, we constrain regions of the CF template that would correspond to the occluded region to zero. This is done so that the CF design will not use image pixels that will be unavailable (due to partial occlusion) during testing. We again use the MACE filter as an example, even though the MCF formulation is general and can be applied to all CF formulations. In this case, let us assume that we have prior knowledge that the set of M spatial pixels in the region R will be occluded during testing. Then we can modify the filter design formulation to include these new constraints.

$$h(n) = \frac{1}{N}\sum_{k=0}^{N-1} H(k)e^{j2\pi kn/N} = 0 \qquad \text{for } n \in R \tag{2}$$

The MCF constraints can be rewritten as

$$\mathbf{C}^{+}\bar{\mathbf{h}} = \mathbf{0}_M \tag{3}$$

where $\mathbf{C}^{+}$ is the $M \times N$ IDFT matrix corresponding to the spatial pixels in R, and $\mathbf{0}_M$ is a zero vector of length M. To solve for the Masked ZAMACE filter, we can combine these new constraints with the original ZA constraints, as well as the traditional CF peak constraints to obtain

$$\mathbf{G}^{+}\bar{\mathbf{h}} = \mathbf{k} \tag{4}$$

where $\mathbf{G}^{+} = \begin{bmatrix} \bar{\mathbf{X}}^{+} & \mathbf{A}^{+} & \mathbf{C}^{+} \end{bmatrix}^T$, $\mathbf{k} = \begin{bmatrix} \mathbf{u} & \mathbf{0}_L \end{bmatrix}^T$, $\mathbf{A}^{+}$ is the $(N - N_x) \times N$ IDFT matrix corresponding to the tail pixels in the ZACF and $\mathbf{0}_L$ is the zero vector of length $L = N - N_x + M$. This leads to the following MCF-ZAMACE

$$\bar{\mathbf{h}} = \mathbf{D}^{-1}\mathbf{G}\left(\mathbf{G}^{+}\mathbf{D}^{-1}\mathbf{G}\right)^{-1}\mathbf{k} \tag{5}$$

4. NUMERICAL EXPERIMENTS

We conducted numerical experiments to test the performance of MCFs compared with other filters. To investigate the performance of MCFs, we tested CFs with three face image databases: the CMU-PIE database [8], the AR database [9], and the KACST database. The training sets were unoccluded face images from the respective databases. There are two classes of occlusions for which we performed testing: sunglasses, and scarves. For each, we create a mask, constraining regions to zero centered at either the eyes, or the bottom portion of the face, as shown in Fig. 3. We vary the percentage occluded for the scarf occlusion from 0% to 60%. For the simulated sunglasses occlusion, we vary the radius of occlusion from 0 pixels to 50 pixels.

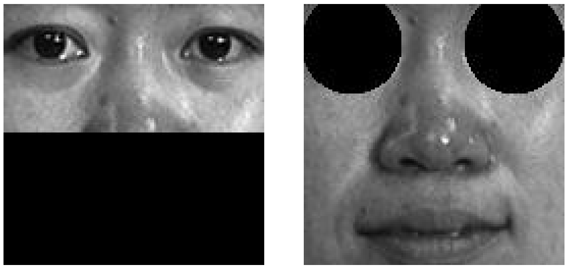


Fig. 3. Occlusion Mask Examples

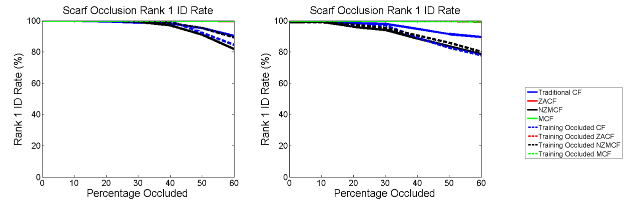
We use the UOTSDF to investigate four types of filter design: traditional CF, ZACF, MCF, and Non Zero-Aliased Masked Correlation Filter (NZMCF). NZMCF is the same formulation as MCF, except $\mathbf{G}^+ = \begin{bmatrix} \bar{\mathbf{X}}^+ & \mathbf{C}^+ \end{bmatrix}^T$. In addition to using the base training images, we also trained the CFs using artificially occluded version of the training images. This is done by setting the pixel values at the assumed occluded regions to 0 in the training images. The results for these are designated using the "Training Occluded" prefix for the appropriate CFs.

During the experimental procedure, we created one CF per subject, and then applied each CF to each test image. The correlation output is analyzed to obtain a Peak-to-Correlation Energy ratio (PCE) score, which is defined as the ratio of the square of the peak value of the correlation output to its total energy. The CF corresponding to the highest PCE is classified as the identified subject. We evaluate the CF methods based on their rank-1 ID rate, which is defined as the ratio of the number of correctly classified test images to the total number of testing images.

4.1. CMU-PIE Database

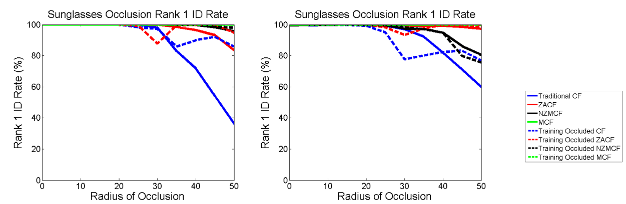
The CMU-PIE database [8] contains face images with different pose, expression, and illumination variations. In this study, we used frontal images of neutral expressions with varying illuminations. Both the PIE-lights and PIE-nolights subsets were used where ambient lights are on and off, respectively. The PIE-lights database consists of 68 classes with 24 images per class, while the PIE-nolights database consists of 66 classes with 21 images per class. These face images were converted to grayscale, cropped and resized to resolution 128x128.

In the PIE database, we induced artificial occlusions to simulate scarves and sunglasses in test images. This is done using the same mask used to artificially occlude the training images for Training Occluded CFs, shown in Fig. 3. The 3 training images used were indices 3, 7 and 16 (corresponding to left, center, and right illuminations). Results for simulated scarf occlusion are shown in Fig. 4, where we vary the percentage occluded from 0% to 60%. For simulated sunglasses occlusion, we vary the radius of occlusion from 0 pixels to 50



(a) PIE-lights (b) PIE-nolights

Fig. 4. Simulated Scarf Rank 1 ID Rate of Traditional CFs, ZACFs, NZMCFs, and MCFs on CMU PIE-lights and PIE-nolights.



(a) PIE-lights (b) PIE-nolights

Fig. 5. Simulated Sunglasses Rank 1 ID rate of Traditional CFs, ZACFs, NZMCFs, and MCFs on CMU PIE-lights and PIE-nolights

pixels, as shown in Fig. 5.

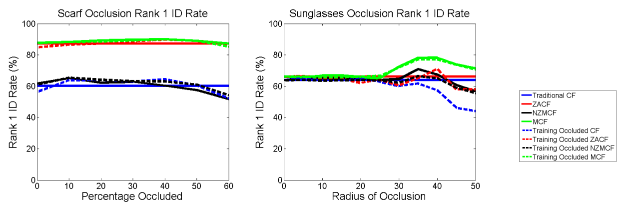
Fig. 4 shows the results of artificial scarf occlusions. MCFs are the most robust to increasing levels of occlusion, whereas traditional CFs are the least robust at higher levels of occlusion, for both types of occlusion. ZACFs are significantly more robust to occlusion than traditional CFs, but in the sunglasses case, we see a small drop in performance at the highest levels of occlusion. It is also interesting that using artificially occluded training images does not improve results.

4.2. AR Database

The AR database [9] consists of over 4,000 face images from 126 subjects (70 men and 56 women). For each subject 26 images are taken of the subject over 2 sessions 2 weeks apart. In each session, images were taken with varying facial expressions (neutral, smile, anger, scream), varying illuminations (left light on, right light on, both side lights on), and 2 different types of occlusion (scarf and sunglasses). For our experiments, we utilized a subset of the database for which each subject had the full set of images, consisting of 119 subjects (65 males and 54 females). The greyscale images were cropped and resized to resolution 128x128. Examples of both unoccluded images and occluded images from the AR database can be seen in Fig. 6.

Training images are of neutral expression with varying illumination. Testing images are varying illumination while wearing either sunglasses or scarves. Fig. 7 shows

Fig. 6. An example of AR database face images



(a) Scarf (b) Sunglasses

Fig. 7. Simulated Scarf and Sunglasses Rank 1 ID Rate of Traditional CFs, ZACFs, NZMCFs, and MCFs on the AR Database

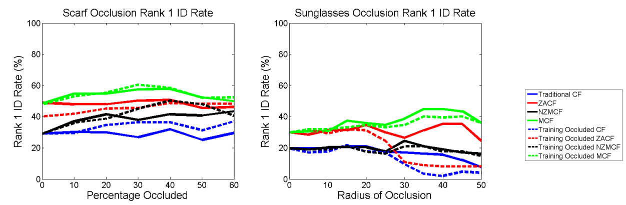
that Masked CFs have the highest performance for both the scarves and sunglasses cases, at 90.0% Rank 1 ID Rate for scarves and 78.2% for sunglasses. This is in comparison to ZACFs (87.2% and 66.1% respectively), and traditional CFs (60.3% and 63.8% respectively). We also find that using artificially occluded training images does not seem to significantly improve results.

4.3. KACST Database

The KACST database consists of face images from 150 subjects. For each subject up to 18 types of images in 5 different poses were taken. The pose variation consits of full frontal view, as well as $\pm 15°$ and $\pm 30°$ from frontal view. Of the 18 types, one is of the subject wearing a cap, one with no headwear, 2 side profiles with no headwear, and the 14 others have the subject wearing a shemagh. There are 6 expression images, 3 illumination images, 3 occlusion images, one image with glasses, and one neutral image. Anyone interested in receiving the database should email KFID@kacst.edu.sa. For our experiments, we utilized a subset of the database where each subject has the full set of frontal images, consisting of 147 subjects. The greyscale images were cropped and resized to 128x128 pixels.

For training images, we used neutral expression images with neutral illumination, while the testing images were neutral illumination wearing either sunglasses or scarves. Fig. 9 shows Masked CFs have the highest performance for both scarf and sunglasses cases, at 60.5% Rank 1 ID Rate and 44.9% respectively. ZACFs have 50.6% and 35.4% Rank 1 ID Rates, whereas traditional CFs perform at 32.0% and 21.2% Rank 1 ID Rates. Here we find that in the scarf case using

Fig. 8. Example KACST database face images



(a) Scarf (b) Sunglasses

Fig. 9. Simulated Scarf and Sunglasses Rank 1 ID Rate of Traditional CFs, ZACFs, NZMCFs, and MCFs on the KACST Database

artificially occluded training images improves results.

5. CONCLUSIONS

Masked Correlation Filters (MCFs) are designed to take advantage of prior knowledge of where partial occlusions will be located in test images, as well as the ZACF technique to enhance performance. In the CMU-PIE, AR, and KACST databases, we have found that ZACFs perform better than traditional CFs, even in the presence of occlusion. We also found that masking seems to help performance for both traditional CFs and ZACFs, and that using artificially occluded training images did not seem to significantly improve results.

In summary, we have proposed Masked Correlation Filters (MCFs), a new method of correlation filters, for object recognition under partial occlusion. Using prior knowledge of the occluded region, we can compensate for the occlusion, improving recognition performance. We tested MCFs on 128x128 face images from the CMU-PIE, AR and KACST databases, and found that MCFs can improve the performance of correlation filters in the presence of large, continuous occlusion. While MCF design needs prior knowledge of occluded regions, it may be possible to determine occlusions using a soft-biometric detector (beards, sunglasses, etc.) [10]. MCFs could also be useful in other pattern recognition tasks, such as automatic target recognition.

Acknowledgments

This research is supported in part by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1252522 and by the King Abdulaziz City for Science and Technology (KACST), Saudia Arabia.

6. REFERENCES

[1] A. M. Martinez, "Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 6, pp. 748–763, 2002.

[2] M. Yang, L. Zhang, S. C. K. Shiu, and D. Zhang, "Gabor feature based robust representation and classification for face recognition with gabor occlusion dictionary," *Pattern Recognition*, vol. 46, no. 7, pp. 1865–1878, 2013.

[3] X. Luan, B. Fang, L. Liu, W. Yang, and J. Qian, "Extracting sparse error of robust pca for face recognition in the presence of varying illumination and occlusion," *Pattern Recognition*, vol. 47, no. 2, pp. 495–508, 2014.

[4] B.V.K. Vijaya Kumar, M. Savvides, and C. Xie, "Correlation pattern recognition for face recognition," *Proceedings of the IEEE*, vol. 94, no. 11, pp. 1963–1976, Nov 2006.

[5] J.A. Fernandez and B.V.K. Vijaya Kumar, "Zero-aliasing correlation filters," in *Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on*, 2013, pp. 101–106.

[6] J.A. Fernandez, V.N. Boddeti, A. Rodriguez, and B.V.K. Vijaya Kumar, "Zero-aliasing correlation filters for object recognition," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, in press 2015.

[7] A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," *Appl. Opt.*, vol. 26, no. 17, pp. 3633–3640, 1987.

[8] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database," in *Automatic Face and Gesture Recognition, 2002.*, 2002, pp. 46–51.

[9] A.M. Martinez and R. Benavente, "The AR face database," *Computer Vision Center, Universitat Autónoma de Barcelona, CVC Technical Report*, vol. 24, 1998.

[10] A Dantcheva, C Velardo, A D'angelo, and JL Dugelay, "Bag of soft biometrics for person identification: New trends and challenges, mutimedia tools and applications," *Springer, October*, vol. 10, pp. 2010, 2010.