Principal Directions of Synthetic Exact Filters for Robust Real-Time Eye Localization

Vitomir Štruc[1,2], Jerneja Žganec Gros[1], and Nikola Pavešić[2]

[1] Alpineon Ltd, Ulica Iga Grudna 15, SI-1000 Ljubljana, Slovenia
{vitomir.struc,jerneja.gros}@alpineon.com
[2] Faculty of Electrical Engineering, University of Ljubljana, Tržaška cesta 25,
SI-1000 Ljubljana, Slovenia
{vitomir.struc,nikola.pavesic}@fe.uni-lj.si

Abstract. The alignment of the facial region with a predefined canonical form is one of the most crucial steps in a face recognition system. Most of the existing alignment techniques rely on the position of the eyes and, hence, require an efficient and reliable eye localization procedure. In this paper we propose a novel technique for this purpose, which exploits a new class of correlation filters called Prinicpal directions of Synthetic Exact Filters (PSEFs). The proposed filters represent a generalization of the recently proposed Average of Synthetic Exact Filters (ASEFs) and exhibit desirable properties, such as relatively short training times, computational simplicity, high localization rates and real time capabilities. We present the theory of PSEF filter construction, elaborate on their characteristics and finally develop an efficient procedure for eye localization using several PSEF filters. We demonstrate the effectiveness of the proposed class of correlation filters for the task of eye localization on facial images from the FERET database and show that for the tested task they outperform the established Haar cascade object detector as well as the ASEF correlation filters.

Keywords: Biometrics, eye localization, advanced correlation filters.

1 Introduction

Advanced correlation filters have been receiving increasing attention in recent years because of their desirable properties such as mathematical simplicity, computational efficiency and robustness to distortions [8]. They have successfully been applied to various problems ranging from pattern recognition tasks such as face and palmprint recognition to basic computer vision problems related to object detection and tracking.

Correlation filters exhibit a high degree of similarity with templates and correlation-based template matching techniques, where patterns of interest in images are searched for by cross-correlating the input image with one or more example templates and examining the resulting correlation plane for large values - also known as correlation peaks. With properly designed templates, these correlation peaks can be exploited to determine the presence and/or location

of patterns of interest in the given input image [8]. Early template matching techniques relied on rather primitive templates, computed, for example, through simple averaging of the available training images. Contemporary methods, on the other hand, use correlation templates (also referred to as *correlation filters*) that are constructed by optimizing specific performance criteria [7], [8], [1]. Popular examples of these advanced correlation filters include Synthetic Discriminant Function (SDF) filters [4], Minimum Average Correlation Energy (MACE) filters [9], Distance Classifier Correlation Filters (DCCF) [10], Maximum Average Correlation Height (MACH) filters [11], Optimal Tradeoff Filters (OTF) [13], Unconstrained Minimum Average Correlation Energy (UMACE) filters [14], and Average of Synthetic Exact Filters (ASEF) [1].

In this paper we introduce a new class of correlation filters named Principal directions of Synthetic Exact Filters (PSEFs). These filters extend the recently proposed class of advanced correlation filters called Average of Synthetic Exact Filters (ASEF) [1]. Instead of only relying on the average of a set of Synthetic Exact Filters (SEFs), as it is the case with the ASEF filters, we employ the eigenvectors of the correlation matrix of the SEFs as correlation templates (or filters). Hence, the name PSEFs. We apply the proposed filters to the task of eye localization and demonstrate their effectiveness in comparison with ASEF filters as well as the established Haar cascade classifier proposed in [17].

2   Principal Directions of Synthetic Exact Filters

2.1   Review of ASEF Filters

ASEF filters represent a class of recently proposed correlation filters that have already been successfully applied to the tasks of eye localization and pedestrian detection [1], [2]. As with all correlation filters, a pattern of interest in an image is detected with an ASEF filter by cross-correlating the input image with the computed filter and examining the correlation plane for possible correlation peaks. While ASEF filters are deployed much in the same way as other existing correlation filters, they differ from most other filters in the way they are constructed.

Unlike the majority of existing correlation filters, which specify only a single correlation value per training image, ASEF filters define the entire correlation plane for each available training image. As stated by Bolme et al. [1], this correlation plane commonly features only a high peak centered at the pattern of interest and (near) zeros at all other image locations (Fig. 1 - middle image). Such a synthetic correlation output results in so-called synthetic exact filters (SEFs) (Fig. 1 - right image) that can be used to locate the pattern of interest in the training image, from which they were constructed.

Unfortunately, these SEF filters do not pride themselves with broad generalization capabilities, instead they produce distinct peaks only for the images that were used for their construction. To overcome this shortcoming Bolme et al. [1] proposed to compute a new filter by averaging all of the synthetic exact filters corresponding to a specific pattern of interest. By doing so, the authors ensured

greater generalization capabilities of the computed ASEF filters and elegantly avoided an important problem of many existing correlation filters, namely, over-fitting.

Formally, the presented procedure of ASEF filter construction can be described as follows. Consider a set of n training images $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$ and n corresponding image locations of our pattern of interest[1] $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$. The first step towards computing an ASEF filter for our pattern of interest is the construction of the desired correlation outputs $\mathbf{y}_1, \mathbf{y}_2, ..., \mathbf{y}_n$ for all n training images, i.e.,

$$\mathbf{y}_i(x,y) = e^{-\frac{(x-x_i)^2 + (y-y_i)^2}{\sigma^2}}, \quad \text{for} \quad i = 1, 2, ..., n, \tag{1}$$

where σ denotes the standard deviation of the Gaussian-shaped correlation output, which controls the balance of the robustness of the filters against noise and the sharpness of the correlation peaks, and (x_i, y_i) represents the coordinates pair corresponding to the location of the pattern of interest in the i-th training image.

Once the correlation outputs have been determined, a SEF is calculated for each of the n pairs $(\mathbf{x}_i, \mathbf{y}_i)$ as follows:

$$H_i^* = \frac{Y_i \odot X_i^*}{X_i \odot X_i^* + \epsilon}, \quad \text{for} \quad i = 1, 2, ..., n, \tag{2}$$

where, $X_i = \mathcal{F}(\mathbf{x}_i)$ and $Y_i = \mathcal{F}(\mathbf{y}_i)$ denote the Fourier transforms of the i-th training image and its corresponding synthetic correlation output, $H_i = \mathcal{F}(\mathbf{h}_i)$ stands for the Fourier transform of the i-th SEF filter $\mathbf{h}_i$, ϵ denotes a small constant that prevents divisions by zero, $\odot$ stands for the Schur product and " * " represents the conjugate operator. It has to be noted that the division in Eq. (2) must be performed element-wise.

In the final step, all n SEFs are simply averaged to produce an ASEF filter (see left image of Fig. 4 for a visual example) that can be used to locate the pattern of interest in a given input image. Here, the ASEF filter in the frequency domain is defined as [1]:

$$H^* = \frac{1}{n} \sum_{i=1}^{n} H_i^*, \tag{3}$$

or equivalently in the spatial domain

$$\mathbf{h} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{h}_i, \quad \text{where} \quad \mathbf{h}_i = \mathcal{F}^{-1}(H_i). \tag{4}$$

An example of the filter construction procedure up to the averaging step is also visualized in Fig. 1. Here, the left image depicts an sample face image, which has been transformed into the log domain, normalized to zero mean and unit variance and finally weighted with a cosine window. The second image shows

---

[1] In our case the image locations correspond to the location of the left eye in all n training images.
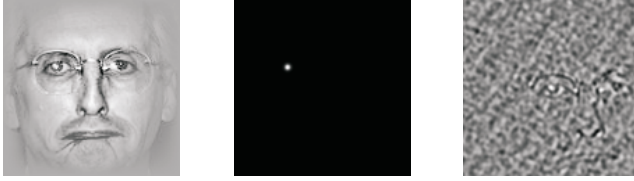
Fig. 1. Construction of a synthetic exact filter (SEF): normalized input image multiplied with a cosine window (left), the synthetic correlation output plane (middle), the synthetic exact filter corresponding to the training image on the left (right)

the visual appearance of a synthetic correlation output with the desired peak response centered at the location of the left eye. Finally, the last image in Fig. 1 represents the SEF filter computed based on the first two images.

Before we turn our attention to the proposed extension of the ASEF filters, let us say a few more words on their characteristics. To ensure adequate generalization capabilities of the ASEF filters, a large number of training images must be used in their construction. Alternatively, a moderate number of training images may be used, however, in this case the SEF filter must be constructed using only the largest Fourier coefficients that contain 95% of the total energy [2]. This alternative approach is also used in our experiments.

2.2   ASEF Filters for Localization

As we have already indicated several times in the paper, ASEF filters can among others also be used for facial landmark localization. In this setting the input image in simply cross-correlated with the ASEF filter corresponding to the desired pattern of interest and the correlation output is then examined for its maximum. The location of the maximum is then declared the location of the pattern of interest. For efficiency reasons all computations are performed in the frequency domain using simple element-wise multiplications:

$$Y = X_t \odot H^*, \tag{5}$$

where Y denotes the correlation output in the frequency domain, $X_t = \mathcal{F}(\mathbf{x}_t)$ denotes the Fourier transform of a test image $\mathbf{x}_t$, H stands for the ASEF filter in the frequency domain and $\odot$ again represents the Schur (i.e., element-wise) product. The procedure is also shown in Fig. 2.

2.3   Beyond Averaging

The filter construction procedure presented in Section 2.1 ensures high generalization capabilities of the ASEF filters by averaging the individual synthetic exact filters. However, this procedure implicitly assumes that the SEF filters represent a random variable drawn from a uni-modal symmetric distribution and, thus, that their distribution is adequately described by their sample mean.

Fig. 2. Visualization of the facial landmark localization procedure using ASEF filters (from left to right): the modified input image, the ASEF filter for the right eye (with shifted quadrants), the correlation output, the input image with the detected correlation maximum

For our derivation, presented in the remainder of this section, we will make a similar assumption and presume that the SEF filters are drawn from a multivariate Gaussian distribution. Under this assumption, we are able to extend the concept of ASEF filters to a more general form, namely, to Principal directions of Synthetic Exact Filters (PSEFs). The basic reasoning for our generalization stems from the fact that the first eigenvector of the correlation matrix of some sample data corresponds to the data's mean (or average), while the remaining eigenvectors encode the variance of the sample data in directions orthogonal to the data's average. By using more than only the first eigenvector (note that the first eigenvector is actually the ASEF filter) of the SEF correlation matrix for the localization procedure, we should be able to further improve upon the localization performance of the original ASEF filters. Similarly as we have done for the ASEF filters, let us now formalize the procedure for PSEF filter construction. Again consider a set of n training images $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n$, for which we have already computed n corresponding SEFs for some pattern of interest, i.e., $\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_n$, in accordance with the procedure presented in Section 2.1. Furthermore, assume that the SEFs reside in a d-dimensional space and that they are arranged into the columns of some matrix $\boldsymbol{\zeta} \in \mathbb{R}^{d \times n}$. Instead of simply averaging the SEFs to produce an ASEF filter with high generalization capabilities, we compute the sample correlation matrix $\boldsymbol{\Sigma}$ of the SEFs:

$$\boldsymbol{\Sigma} = \boldsymbol{\zeta}\boldsymbol{\zeta}^T \in \mathbb{R}^{d \times d}, \tag{6}$$

where ^{T} denotes the transpose operator, and use its leading eigenvectors as our PSEF filters, i.e.:

$$\boldsymbol{\Sigma}\mathbf{f}_j = \lambda_j \mathbf{f}_j, \quad \text{where} \quad j = 1, 2, ..., \min(d, n) \quad \text{and} \quad \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{\min(d,n)}. \tag{7}$$

The presented procedure very much resembles the commonly used principal component analysis [16], [15] with the only difference that the SEF filters are not centered around their global mean.

One problem arising from the presented derivation of the PSEF filters $\mathbf{f}_j$ is the sign ambiguity of the eigenvectors. Since the computed PSEF filters can be multiplied by -1 and still represent valid eigenvectors of $\boldsymbol{\Sigma}$, we have to alleviate this sign ambiguity to be able to use our PSEF filters for localization purposes.
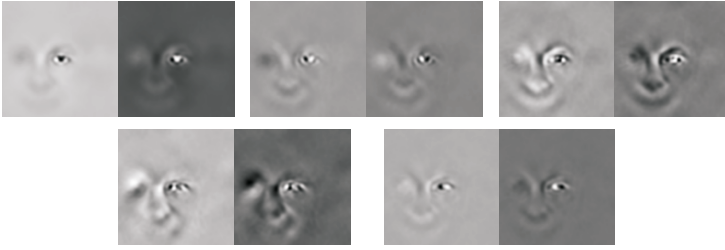
Fig. 3. Visual appearance of the first five PSEFs. The upper row depicts (from left to right) the PSEFs corresponding the largest three eigenvalues of the SEF correlation matrix and the lower row depicts the PSEFs corresponding to the next two eigenvalues, i.e., the fourth and fifth largest eigenvalues. In each image pair the left image represents the computed PSEF multiplied with +1 and the right image represents the computed PSEF multiplied with -1.

In the experimental section we will try to solve the sign ambiguity of our filters through some preliminary experiments. For the moment let us just take a look at the visual appearance of the first five PSEF filters (corresponding to the five largest eigenvalues of $\boldsymbol{\Sigma}$) shown in Fig. 3. Note that the visual appearance of the first PSEF filter (first image in the upper row of Fig. 3) is identical to the appearance of the ASEF filter (Fig. 2 - second image from the left).

2.4 Exploiting Linearity

Similarly to the ASEF filters, PSEF filters can also be exploited for the localization of facial landmarks. The procedure is identical the one presented in Section 2.2, except for the fact that we have more than a single filter at our disposal and, hence, obtain more than one correlation output:

$$Y_j = X_t \odot F_j^*, \quad \text{for} \ \ j \in \{1, 2, ..., \min(d, n)\}, \tag{8}$$

where $X_t = \mathcal{F}(\mathbf{x}_t)$ again denotes the Fourier transform of the given test image $\mathbf{x}_t$, F_j denotes the Fourier transform of the j-th PSEF filter $\mathbf{f}_j$ and Y_j refers to the j-th correlation output in the Fourier domain.

To determine the location of our pattern of interest in the given input image, we obviously have to examine all correlation outputs Y_j for maxima and somehow combine all of the obtained information. A straight forward way of doing this is to examine only the linear combination of all correlation outputs for its maximum and use the location of the detected maximum as the location of our pattern of interest. Thus, we have to examine the following combined correlation output:

$$\mathbf{y}_c = \sum_{i=1}^{k} w_i \mathbf{y}_i, \tag{9}$$

where $\mathbf{y}_i$ denotes the correlation output (int the spatial domain) of the i-th PSEF filter, w_i denotes the weighting coefficient of the i-th correlation output,

Fig. 4. Comparison of the visual appearance of an ASEF filter (left) and the combined PSEF filter (right). Both images show "right eye" filters with shifted quadrants.

$\mathbf{y}_c$ denotes the combined correlation output, and k stands for the number of PSEF filters used ($1 \leq k \leq \min(d, n)$). From the above equation we can see that if $k = 1$ the combined correlation output is identical to the correlation output of the ASEF filter. On the other hand if $k > 1$ we add additional information to the combined correlation output by including additional PSEF filters into the localization procedure. The presented procedure requires one filtering operation for each PSEF filter used. However, the computation can be speeded up by exploiting the linearity of Eq. 9. Instead of combining the correlation outputs, we simply combine all employed PSEF filters into one single filter with hopefully enhanced localization capabilities, i.e.:

$$\mathbf{y}_c = \sum_{i=1}^{k} w_i \mathbf{y}_i = \sum_{i=1}^{k} w_i (\mathbf{f}_i \otimes \mathbf{x}_t) = \left(\sum_{i=1}^{k} w_i \mathbf{f}_i\right) \otimes \mathbf{x}_t = \mathbf{f}_c \otimes \mathbf{x}_t, \qquad (10)$$

where $\mathbf{f}_c = \sum_{i=1}^{k} w_i \mathbf{f}_i$, and $\sum_{i=1}^{k} w_i = 1$. In the presented equations $\mathbf{f}_c$ stands for the combined PSEF filter and $\otimes$ denotes the convolution operator. We can see that instead of using k PSEF filters and produce k correlation outputs that are linearly combined, we simply combine the k employed filters into a single filter and, hence, perform only a single filtering operation. The localization procedure has, therefore, exactly the same computational complexity as the procedure relying on ASEF filters regardless of the number of PSEF filters selected for the localization of our pattern of interest.

The last issue to be solved before we turn our attention to the experimental section is the choice of the weighting coefficients w_i, for $i = 1, 2, ..., k$. While an optimization procedure could be exploited to determine the best possible combination of the k filters, we choose in this paper to select the coefficients in accordance with the following expression:

$$w_i = \frac{\lambda_i}{\sum_{i=1}^{k} \lambda_i}, \qquad (11)$$

where λ_i represents the eigenvalue corresponding to the i-th PSEF filter $\mathbf{f}_i$ (see Eq. 7). This procedure is clearly sub-optimal, but it is, nevertheless, enough to demonstrate the usefulness of the proposed filter combination. An example of the visual appearance of the combined PSEF filter obtained with the presented weighting procedure (after the sign ambiguity has been eliminated - see

Section 3) is shown on the right hand side of Fig. 4. For comparison purposes
the left hand side image of Fig. 4 also shows the original ASEF filter.

3   Experiments

To assess the effectiveness of the proposed localization procedure relying on
PSEF filters we adopt two popular face databases, namely, the grey FERET
database [12] and the Labeled Faces in the Wild (LFW) database [5]. We extract
the facial regions from all images of the two databases using the Haar cascade
classifier proposed by Viola and Jones [17]. Here, we rely on the freely available
implementation of the Haar face detector that ships with the OpenCV library [3].
After determining the location of the facial regions in all images, we select 640
images from the LFW database and manually label the locations of the left and
right eye. Next, we produce 40 variations of the facial region of each of the 640
LFW images by randomly shifting the location of the facial regions by up to
± 5 pixels, rotating them by up to ± 15, scaling them by up to 1.0 ± 0.15 and
mirroring them around the y axes. Through these transformations, we augment
the initial set of 640 images to a set of 25600 images (of size 128×128 pixels)
that we employ for training of the ASEF and PSEF filters.

For testing purposes we apply the same random transforms to 3815 images
from the FERET database. Here, we produce only 12 modifications of each facial
region, which results in 45780 facial images being available for our assessment.
Some examples of the 12 modifications of a face image from the FERET database
are also shown in Fig. 5.

Prior to subjecting the face images to the proposed localization procedure, all
face images are transformed into the log-domain and normalized using zero mean
and unit variance normalization. In the last step the images are weighted with a
cosine window to reduce the frequency effects of the edges commonly encountered
when applying the Fourier transform [1]. Once the localization procedure has
been performed, we employ the following criteria to measure the effectiveness of
our approach [6]:

$$\eta_{se} = \frac{\|l_{se} - r_{se}\|}{\|r_{le} - r_{re}\|} \;\; \text{and} \;\; \eta_{te} = \frac{\max \left(\|l_{le} - r_{le}\|, \|l_{re} - r_{re}\| \right)}{\|r_{le} - r_{re}\|}, \qquad (12)$$

where η_{se} and η_{te} stand for the "single eye" and "two eye" criterion, respec-
tively; l_{se} denotes the location of the single eye of interest found by the assessed

Fig. 5. Visual examples of a sample face region from the FERET database detected
with the Viola-Jones face detector and its eleven modifications

(a) Results for PSEF 1   (b) Results for PSEF 2   (c) Results for PSEF 3

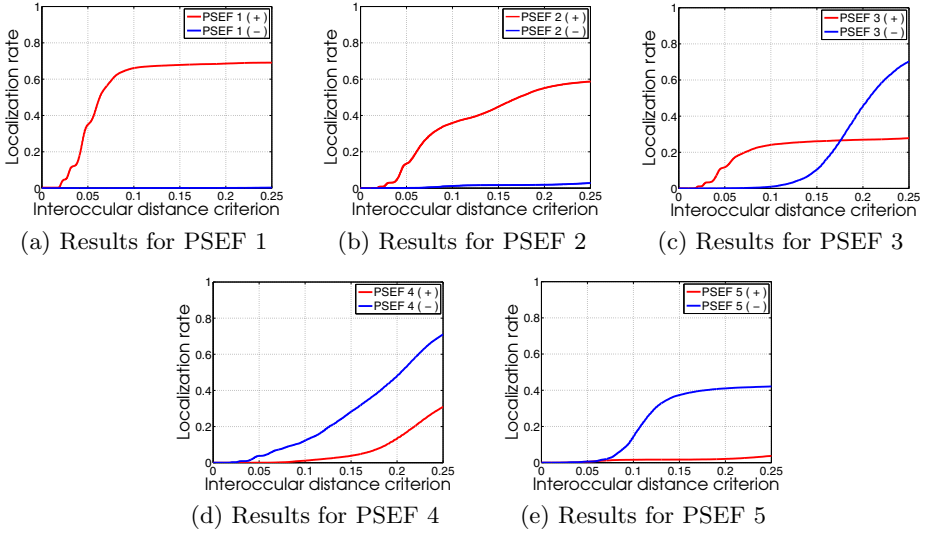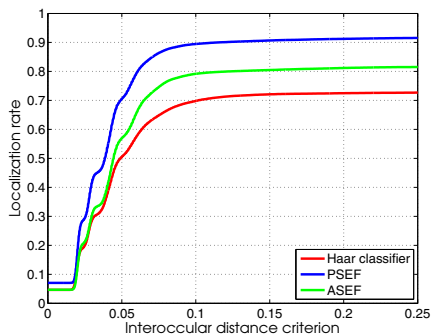(d) Results for PSEF 4   (e) Results for PSEF 5

Fig. 6. Results of preliminary experiments aimed at alleviating the sign ambiguity of the computed PSEFs (using the "two-eye" criterion)
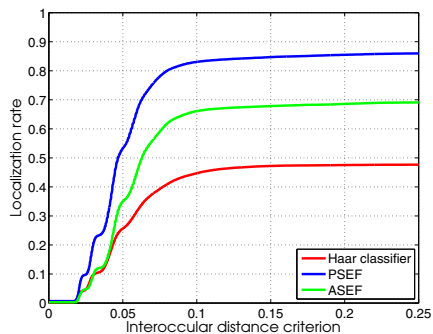
procedure, r_{se} denotes the reference location of the single eye of interest, the expression $\|r_{le} - r_{re}\|$ represents the interoccular L_2 distance, and the subscripts le and re stand for the left and right eye, respectively. We can see that the "two eye" criterion is more restrictive, as it requires both eyes to be near their reference locations for the criterion to have a small value. The "single eye" criterion, on the other hand, requires only the eye of interest to be near the reference location. For our assessment we observe the correct localization rate for different operating points, i.e., $\eta_{se}, \eta_{te} < \Delta \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$.

Our first series of experiments aims at alleviating the sing ambiguity of the computed PSEF filters. To this end, we compute 5 PSEF filters (corresponding to the 5 largest, non-zero eigenvalues of Eq. 7), derive two filters from each of the 5 PSEF filters by multiplying them with $+1$ and -1, and normalizing the result to zero mean and unit variance. With the 5 computed filter pairs, we conduct localization experiments with the 45780 face images of the FERET database and plot the results in form of graphs as shown in Fig. 6. We select the "two eye" criterion with $\Delta = 0.25$ as the relevant operating point and based on this value determine the appropriate sign of each of the five PSEF filters. Note here that more (or less) filters than 5 could be used for our experiments, the presented results, however, are enough to show the feasibility of our approach.

If we take a look at the presented results in Fig. 6, we can see that in our case the best localization results are obtained with the first two filters being multiplied with $+1$ and the remaining filters being multiplied with -1. Furthermore, we can notice, that the best localization performance is obtained with the first PSEF filter, which in fact corresponds to an ASEF filter, while the remaining filters perform worse. Nevertheless, they hopefully contain complementary information
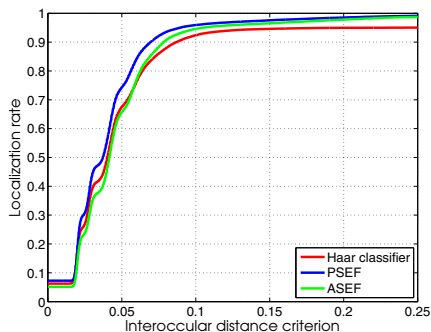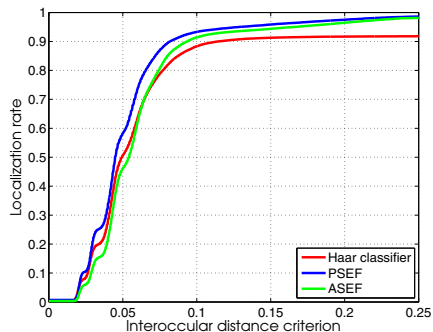
(a) Localization results for the left eye   (b) Localization results for both eyes

Fig. 7. Comparison of the eye localization performance of different localization techniques using the: (a) "single eye" and (b) "two eye" criterion. In the experiments the entire 128×128 face region of the test images was searched for the eyes.

(a) Localization results for the left eye   (b) Localization results for both eyes

Fig. 8. Comparison of the eye localization performance of different localization techniques using the: (a) "single eye" and (b) "two eye" criterion. In the experiments only the upper left quadrants of the 128×128 face regions were searched for the left eye and the upper right quadrants for the right eye.

to the first PSEF filter. Our second series of experiments comprises two types of tests. The first type uses no a priori knowledge about the locations of the left and right eye, while the second type relies on a priori knowledge about the eye locations and, hence, looks for the left eye only in the upper left quadrant of the test images and for the right eye only in the upper right quadrant of the test images. This setup is identical to the experimental setup adopted in [1] and is used here to allow for a comparison of the localization performance with previously published results.

   The results for the first type of experiments are shown in Fig. 7, while the results of the second type of experiments are shown in Fig. 8. Some numerical results for different values of Δ are also summarized in Table 1. Note that the proposed PSEF filters outperform both tested alternatives to eye localization,

Table 1. Localization rates (in %) at different criterion thresholds. Note that for the localization rates corresponding to the "Left eye" columns the "single eye" criterion was used, while for the localization rates corresponding to the "Both eyes" columns the "two eye" criterion was adopted

Criterion	Unconstrained search space						Constrained search space					
	Left eye			Both eyes			Left eye			Both eyes		
	Haar	ASEF	PSEF	Haar	ASEF	PSEF	Haar	ASEF	PSEF	Haar	ASEF	PSEF
0.05	50.5	56.9	70.5	25.6	35.0	53.0	67.5	65.6	74.3	50.6	46.1	58.2
0.10	69.8	79.2	89.5	44.7	66.1	83.0	92.4	94.6	95.9	88.3	91.4	93.3
0.15	71.1	80.5	90.7	47.2	67.8	84.7	94.6	96.5	97.6	91.3	94.4	95.8
0.20	72.5	81.2	91.2	47.5	68.6	85.5	95.0	97.8	98.5	91.7	96.5	97.5
0.25	72.7	81.5	91.5	47.7	69.1	86.0	95.0	98.7	99.1	91.8	98.1	98.6

Table 2. Best average time needed for the localization procedure

Face part	Unconstrained search space		Constrained search space	
	Haar classifier	Correlation filter	Haar classifier	Correlation filter
Left eye	21.6 ms	0.65 ms	11.5 ms	0.66 ms
Right eye	24.8 ms	0.35 ms	13.6 ms	0.35 ms
Both eyes	46.4 ms	1.00 ms	25.1 ms	1.01 ms

namely, ASEF filters as well as the Haar cascade classifier. The proposed filters perform best for both criteria, i.e., the "single eye" and the "two eye" criterion, and both types of conducted experiments.

If we look at the execution times in Table 2 needed for the localization procedure, we can see that the correlation filters require significantly less time for the localization of both eyes than the Haar cascade classifier. Moreover, we can see that the localization time with the Haar classifier for each of the two eyes is more or less identical, while the correlation filters require approximately half the time for the second eye due to the fact that the test image only needs to be transformed into the Fourier domain once. Thus, when looking for the right eye, we already have the frequency representation of the test image at our disposal. It should be noted here that all durations presented in Table 2 represent the best average duration of the localization procedure we have measured in our experiments.

The final comment we need to make before concluding the experimental section refers to the time needed to train the eye locators. The ASEF filters typically require only a few minutes to be trained, since the rely only on a simple average of the exact synthetic filters. The PSEF filters require a few hours for their training, as this involves the computation of a large correlation matrix and its decomposition. Finally, the Haar cascade classifier is known to have training times in the order of days or even weeks. While the training is commonly performed off-line, it is nevertheless important that it is as rapid as possible,

as small changes (such as changes in the photometric normalization procedure used) in the systems relying on eye localization procedures often induce the need for retraining of the eye locator.

4   Conclusion

We have presented a new class of correlation filters called Principal directions of Synthetic Exact Filters and applied them to the task of eye localization. We have shown that the filters outperform the recently proposed ASEF filters and the established Haar cascade classifier at this task, and that they exhibit some desirable properties such as extremely low execution times.

Acknowledgements

References

1. Bolme, D.S., Draper, B.A., Beveridge, J.R.: Average of synthetic exact filters. In: Proc. of CVPR 2009, pp. 2105–2112 (2009)
2. Bolme, D.S., Liu, Y.M., Draper, B.A., Beveridge, J.R.: Simple real-time human detection using a single correlation filter. In: Proc. of the 12th Workshop on Performance Evaluation of Tracking and Surveillance, pp. 1–8 (2009)
3. Bradski, G., Kaehler, A.: Learning OpenCV: computer vision with the OpenCV library. O'Reilly Media, Sebastopol (2008)
4. Hester, C.F., Casasent, D.: Mulitvariant technique for multiclass pattern recognition. Applied Optics 19(11), 1758–1761 (1980)
5. Huang, G.B., Ramesh, M., Berg, T., Learned-Miller, E.: Labeled Faces in the Wild: a database for studying face recognition in unconstrained environments. University of Massachusetts, Amherst, Technical Report 07-49 (October 2007)
6. Jesorsky, O., Kirchberg, K.J., Frischholz, R.W.: Robust face detection using the hausdorff distance. In: Bigun, J., Smeraldi, F. (eds.) AVBPA 2001. LNCS, vol. 2091, pp. 90–95. Springer, Heidelberg (2001)
7. Kerekes, R.A., Kumar, B.V.K.V.: Correlation filters with controlled scale response. IEEE Transactions on Image Processing 15(7), 1794–1802 (2006)
8. Kumar, B.V.K.V., Mahalanobis, A., Takessian, A.: Optimal tradeoff circular harmonic function correlation filter methods providing controlled in-plane rotation response. IEEE Transactions on Image Processing 9(6), 1025–1034 (2000)
9. Mahalanobis, A., Kumar, B.V.K.V., Casasent, D.: Minimum average correlation energy filters. Applied Optics 26(17), 3633–3640 (1987)
10. Mahalanobis, A., Kumar, B.V.K.V., Sims, S.R.F.: Distance-classifier correlation filters for multiclass target recognition. Applied Optics 35(17), 3127–3133 (1996)
11. Mahalanobis, A., Kumar, B.V.K.V., Song, S., Sims, S.R.F., Epperson, J.: Unconstrained correlation filters. Applied Optics 33(17), 3751–3759 (1994)

12. Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J.: The FERET evaluation methodology for face-recognition algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(10), 1090–1104 (2000)
13. Refregier, P.: Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency. Optics Letters 16(11), 829–831 (1991)
14. Savvides, M., Kumar, B.V.K.V.: Efficient design of advanced correlation filters for robust distortion-tolerant face recognition. In: Proc. of the IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 45–52 (2003)
15. Štruc, V., Gajšek, R., Pavešić, N.: Principal Gabor filters for face recognition. In: Proc. of BTAS 2009, pp. 1–6 (2009)
16. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience 3(1), 71–86 (1991)
17. Viola, P., Jones, M.J.: Robust real-time face detection. International Journal of Computer Vision 57, 137–154 (2004)