Proceedings of the International Conference on Computer and Information Science and Technology
Ottawa, Ontario, Canada, May 11 – 12, 2015
Paper No. 82

Design of State Invariant Target Classifiers

Kaveh Heidary
Department of Electrical Engineering and Computer Science
Alabama A&M University
PO Box 702, Normal, Alabama 35762, USA
kaveh.heidary@aamu.edu

Abstract- This paper presents a novel and computationally efficient algorithm for the design of robust state invariant target classifiers. The objective is to obtain a classifier for imagery based target detection, classification, recognition, identification, authentication, and tracking with potential applications in autonomous and biometric systems. The resultant classifier is a filter bank comprised of a small number of correlation filters and the corresponding threshold values. Each constituent of the filter bank classifier, consisting of the coupled-duet of the correlation filter and the respective threshold, is obtained by fusion of an appropriately clustered set of training images. The algorithm first creates a suitable partitioning of the training set of images, which represent the target class of interest in various states including view angle, range (scale), lighting, shadow effects, and partial obscuration, into the trainer clusters. Each set of images corresponding to a particular training set cluster are subsequently fused into a composite synthetic image from which the enhanced matched filter (EMF) and the respective threshold are obtained. The algorithms for optimal partitioning of the training set of images into cohesive image clusters and the amalgamation of each cluster into the respective EMF are explained. The classifier performance is assessed using realistic test scenarios, and the effects of various parameter settings on filter performance are investigated.

Keywords: Image Processing, Classification, Target Recognition, Enhanced Matched Filter

1. Introduction

This paper presents a novel and computationally efficient algorithm for training robust image classifiers. Application areas include a wide range of defense, surveillance, security, and intelligence systems including imagery based automatic target detection, classification, recognition, identification, and tracking. Other application areas include biometrics, industrial automation and inspection, self aware and intelligent systems where the classifier may be utilized for performing various tasks including authentication, navigation, and situational awareness (Davies, 2012).

The fundamental principles underlying the classifier as well as the basic theory and training algorithms developed in this paper are applicable to various classes of signals. In principle, the algorithms developed here can be applied to signals produced by sensors with diverse modalities including radar, Lidar, synthetic aperture radar (SAR), FLIR, hyperspectral, IR and EO including RGB and grayscale imagery. In order to clarify the primary approach and explain the salient features of the algorithms, the mathematical formulation of the classifier is cast in terms of grayscale imagery data. The grayscale image characterizes the two dimensional (2D) signature of the object located in the sensor field of view (FOV), which represents the filtered optical energy incident upon the focal plane array of the sensor.

The classifier presented in this paper involves a bank of synthetic images and the corresponding thresholds. Each coupled pair of the synthesized image and the respective threshold is called an enhanced matched filter (EMF). The complete set of EMFs, namely, the enhanced matched filter bank (EMFB) collectively constitute the classifier (Heidary, 2005 and 2009). The EMFB classifier, is obtained from a training set of images representing the target class of interest under varied conditions and states such as view angle, including in plane and out of plane rotation, range (scale), lighting, shadowing, and potential partial obscuration.

The algorithm involves a learning procedure (Belongie,2002), where the available set of labeled target class images are distilled into the classifier EMFB which contains a significantly smaller number of synthetic images than the number of images in the training set (Heidary, 2009). The construction of EMFB is entirely based on target class exemplar images and the training process does not require any non-target class images. In the operation phase, the EMFB is applied to the sensor image by computing the correlation surface of each constituent EMF with respect to the sensor output. If the height of correlation surface of any one of the EMFB component filters exceeds the respective threshold presence of target at the indicated locations are declared, otherwise the sensor image is deemed to be void of the target of interest.

Section two provides background material and literature review. Section three contains a succinct description of the algorithm and a concise mathematical formulation of EMFB. Section four provides some illustrative examples and assessment of filter performance. Section five includes concluding remarks and suggestions for extending the theory.

2. Background

Among the most critical components of autonomous systems are the machine vision modules, which enable the self-directed systems to utilize imagery data in order to assess their operational environments and act accordingly. Other application areas with extensive deployment of machine vision include robotics, automated navigation, industrial automation and inspection, biometrics, target detection and tracking systems (Belongie, et al., 2002). Computers are utilized to process digital imagery data, which is the result of spatial sampling and quantization of filtered light energy emanating from the scene and impinging on the focal plane array of the sensor. A fundamentally essential element of many advanced machine vision systems is the ability to recognize and locate the object classes of interest in the sensor image reliably, with low latency, and moderate expenditure of power, storage and processing resources.

Addressed in this paper is a supervised learning algorithm for imagery based classification of objects (Davies, 2012). The classifier is constructed using a set of training images that represent the object of interest under assorted object states and viewing conditions such as range and perspective. The two concurrent requirements of locating ability and computational efficiency of the desired classifier point in the direction of Fourier filtering, which is inherently predisposed to parallel processing (Heidary, 2005). The Fourier filter achieves the assessment of each sensor image in a single swoop without the need for image segmentation and piecemeal examination of each partition. The optimal method for detection of a known signal in background noise is matched filtering, which results in maximum signal to noise power ratio (SNR) at the sensor output (Turin, 1960). Matched filtering is reduced to template matching when the spectral power density of noise is flat.

In target detection and tracking applications the target of interest, in general, is described by many images that represent the target under different states and view conditions (Georghiades, 2002). As described by Landeau and Dagobert (2006), and cited by Johnson and Heidary (2011), the target dictionary, in practical scenarios, may potentially consist of hundreds or even thousands of image templates. The standard template matching techniques, which require examination of sensor images against numerous target templates, are in most cases not practicable. Processing the output of fast frame-rate sensors with massive banks of templates corresponding to multiple classes of targets of interest, in real time, presents an insurmountable storage and computational barrier.

Many investigators including Mahalanobis (1987), Caulfield (1992), Casasent (1992), Kumar (1996), and Heidary (2013) have developed distortion-tolerant filters and synthetic discriminant functions (SDF) in order to reduce the arduous computational burden of storing and processing vast image dictionaries. As shown by Heidary (20013, 2014), combining the training image ensemble that represent the target class of interest under a controlled range of states and view conditions into an enhanced matched filter (EMF), leads to a correlation filter with a specific zone of effectiveness. This paper advances the EMF theory by providing a systematic mechanism for segmentation of the trainer set of images into a set of cohesive clusters of trainers from which a bank of filters is created with a generalized zone of effectiveness.

3. Problem Formulation

In the analysis presented in this paper the target class of interest is characterized by a set of grayscale intensity images. These images represent the target class in various states and view conditions such as perspective, including in plane and out of plane rotation angles, sensor-to-object distance, lighting and partial obscuration. The entire collection of known target class images constitute the trainer/exemplar set. The objective is to obtain a classifier that can detect and locate potential target occurrences in the sensor images with high reliability, low latency, and minimal storage and processing expenditures. High reliability in this context includes having the capability to generalize well, in the sense of detecting target occurrences that are potentially unlike any of the trainer set elements. It also includes good accuracy in the sense of the ability to guard against misclassification of non-target class objects in the sensor image.

The classifier developed here consists of a number of synthetic images and the respective thresholds. Each synthetic image and the related threshold pair represent an enhanced matched filter (EMF). The enhanced matched filter bank (EMFB) comprised of one or more EMFs constitutes the classifier. The training algorithm consists of two main components, namely, partitioning and synthesis. The partitioning algorithm divides the set of exemplar images into multiple cohesive trainer image clusters based on some user-prescribed parameters. The synthesis algorithm constructs an EMF by merging of the trainer images in each cluster and computing the associated threshold. The assortment of synthesized EMFs constitute the classifier. Each EMF template represents the 2D target class signature that characterizes the target class in a constrained range of states encompassed by the respective cluster.

3. 1. Partitioning Algorithm

This algorithm partitions the set of exemplar images into multiple cohesive clusters of training images for subsequent processing by the EMFB synthesis algorithm. The hierarchical partitioning process involves formation of one cluster at a time, where a single cluster of training images is formed from the existing exemplar set; the images of the constructed cluster are taken out of the trainer set; the remaining set of trainers are used to construct the next cluster; and so on. The process is terminated when one of two conditions are met, namely, the trainer set is empty or the number of clusters has reached the user prescribed limit. The original set of exemplar images is given below.

$$\mathrm{s}_p(m,n)\,|\,1\le p\le \mathrm{P}, 0\le m\le \mathrm{M}-1, 0\le n\le \mathrm{N}-1\ . \tag{1}$$

Where, P is the number of trainers and $M' N$ denote the image dimensions along two spatial directions. It is assumed that all images in Eq. (1) have the same spatial dimensions and are normalized such that, for each image, the sum of pixels is equal to zero and the sum of pixel squares is equal to one. In case the spatial dimensions across the trainer set are not uniform, the images are appended with zero rows and/or columns in order to make all images as large as the one with the greatest spatial dimensions. Equation (2) shows the process of normalizing the training images, where $\hat{\mathrm{s}}_p, \mathrm{s}_p, \mathrm{w}_p$ denote, respectively, a typical raw trainer, the respective normalized trainer, and the related binary mask.

$$\mathrm{s}_p(m,n)=\frac{\hat{\mathrm{s}}_p(m,n)-\dfrac{\mathrm{w}_p(m,n)}{MN}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\hat{\mathrm{s}}_p(m,n)}{\sqrt{\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\left[\hat{\mathrm{s}}_p(m,n)-\dfrac{\mathrm{w}_p(m,n)}{MN}\sum_{n=0}^{N-1}\sum_{m=0}^{M-1}\hat{\mathrm{s}}_p(m,n)\right]^2}};\ \mathrm{w}_p(m,n)=\begin{cases}1; \hat{\mathrm{s}}_p(m,n)\neq 0\\ 0; \hat{\mathrm{s}}_p(m,n)=0\end{cases}. \tag{2}$$

The trainer set distortion matrix is a zero-diagonal $P' P$-matrix, with off diagonal terms characterizing the mutual distortions between corresponding pairs of normalized trainers, as shown below.

$$\mathbf{D} = \left[d_{p,p'} \mid 1 \le p, p' \le P, d_{p',p'} = 0 \right]; d_{p,p'} = 1 - \max_{m',n'} \left\{ \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} s_p(m,n) s_{p'}((m-m'),(n-n')) \right\}. \tag{3}$$

Where, $d_{p,p'}$ denotes the mutual distortion between a typical pair of trainers, and $\mathbf{D}$ is the trainer set distortion matrix. The cluster formation process is commenced with finding the pair of trainer images with the greatest mutual distortion. If there are multiple pairs with such property, one pair is chosen randomly. A latent cluster is evolved about each one of the chosen trainers by adding one trainer at a time to one of the two clusters. This is done by assigning to each remaining trainer two distortion coefficients, one associated with each latent cluster. The distortion coefficient of a trainer with respect to a latent cluster is computed by provisionally assigning the trainer to the cluster and obtaining the distortion factor of the provisionally formed cluster.

$$\Delta_q = \max_{p,p'} \left[d_{p,p'} \mid p, p' \in I_q \right]; I_q \subseteq 1,2,...,P \ . \tag{4}$$

Where, D_q, I_q denote, respectively, the distortion factor of a typical cluster (provisional, latent or otherwise), and the set of indices associated with trainers that belong to the cluster. The trainer with the lowest distortion coefficient is attached to the respective latent cluster. The process is repeated for the remaining trainers by assigning two distortion coefficients to each, and attaching the lowest coefficient trainer to the respective latent cluster. As the two latent clusters are evolved, their distortion factors are recomputed. If the distortion factor of a latent cluster reaches the user prescribed maximum distortion, its evolution process stops. The remaining trainers are then assigned a single distortion coefficient each, corresponding to the provisional assignment of the trainer to the still evolving latent cluster. The process of assigning one remaining trainer at a time to the still evolving latent cluster continues until the distortion factor reaches the limit or the trainer set is empty. At the conclusion of this process, there are two latent clusters, and the one with the greatest number of trainers is added to the set of actual clusters. The trainers that belong to the actual cluster are removed from the trainer set; the process continues using the remaining set of trainers by evolving two latent clusters; the latent cluster with the greatest number of trainers is added to the set of actual clusters; and so on.

At the conclusion of the clustering process, we arrive at one or multiple trainer clusters whose distortion factors are all below the user prescribed limit. In certain cases, the user may place a limit on the number of clusters. In such cases some trainers may potentially be left out of the clusters, and these outlier trainers are not utilized in the EMF synthesis process.

$$\bigcup_{q=1}^{Q} c_q \subseteq s_p(m,n); 1 \le p \le P \ ; c_q = s_{i_q}(m,n) \mid I_q \subset 1,2,...,P \ . \tag{5}$$

Where, c_q denotes a typical trainer cluster, and Q is the number of clusters. At the conclusion of the clustering process, the set of exemplar images has been partitioned into one or more trainer clusters and potentially a small number of outlier trainers. If the distortion factor of the exemplar set is less than the user prescribed maximum distortion, the entire exemplar set is lumped into a single cluster.

In an alternative implementation of the algorithm, if there are any outlier trainers, which are left out of the clusters after the initial clustering phase, the outliers are subsequently absorbed by the clusters as

follows. A set of distortion factors are assigned to each outlier trainer, corresponding to the provisional attachment of the outlier to each cluster. The outlier with the smallest distortion factor is then assigned to the respective cluster. The trainer is removed from the set of outliers, and the process continues by assigning one remaining trainer to one cluster at a time, until no outliers are left. This process leads to some clusters whose distortion factors exceed the user prescribed maximum. This implementation of the partitioning algorithm is used in all the simulations presented here.

3. 2. Synthesis Algorithm

The trainers of each cluster are fused together to form a synthetic image target model, which characterizes the target class in the range of target states incorporated by the cluster. The combination of the resultant synthetic image and the corresponding threshold constitute the EMF. The set of EMFs associated with all the clusters collectively constitute the EMFB. The EMF synthesis process starts by determining the anchor trainer, which is defined as the image that has the smallest maximum distortion with respect to all the other cluster trainers.

$$\exists k \in 1,2,...,P_q \ : \forall k' \in 1,2,...,P_q \qquad \max_{1\le l\le P_q} d_{lk} \le \max_{1\le l\le P_q} d_{lk'} \ . \tag{6}$$

Where, q, P_q denote, respectively the cluster index and the number of exemplar images in the cluster, d_{lk} is the mutual distortion between a typical image pair in the cluster, and k denotes the anchor image index. It is noted that in Eq. (6) the cluster image indices have been reassigned and are different from the respective image indices in the original set of exemplars as provided in Eq. (1). All the cluster images are subjected to proper spatial shifts with respect to the anchor image and are then added pixel-wise. Each image is spatially shifted such that its peak correlation with respect to the anchor image is at pixel (0, 0).

$$\exists (m_l, n_l) : \forall \ \ 0 \le i \le M-1, 0 \le j \le N-1 \ \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} s_k(m,n)\, s_l((m-m_l),(n-n_l)) \ge \sum_{n=0}^{N-1}\sum_{m=0}^{M-1} s_k(m,n)\, s_l((m-i),(n-j)). \tag{7}$$

$$\hat{f}_q(m,n) = \sum_{l=1}^{P_q} s_l((m-m_l),(n-n_l)); \ \ f_q(m,n) = \frac{\hat{f}_q(m,n)}{\sqrt{\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} \hat{f}_q(m,n)^2}}. \tag{8}$$

Where, $s_l(m,n)$ is a typical trainer in the cluster, (m_l, n_l) is the spatial shift of the trainer with respect to the anchor, and $\hat{f}_q(m,n)$ denotes the composite image. The resultant synthetic image $f_q(m,n)$ in Eq. (8) is normalized such that the sum of its pixels, and the sum of pixel squares are equal to zero and one, respectively. The EMF threshold is set equal to the minimum peak correlation of the filter in Eq. (8) with respect to all the trainer images within the cluster.

$$T_q = \min_{1\le l\le P_q} \left\{ \max_{i,j} \left[\sum_{n=0}^{N-1}\sum_{m=0}^{M-1} f_q(m,n)\, s_l((m-i),(n-j)) \right] \right\}. \tag{9}$$

The synthetic image represented by $f_q(m,n)$ of Eq. (8) and the threshold T_q of Eq. (9) constitute the EMF associated with cluster-q, and EMF's of all clusters comprise the EMFB.

In an alternative implementation of this algorithm the user prescribes a minimum threshold T_0. Following the EMFB computation phase, each EMF whose threshold falls below T_0 is modified by dropping the trainer with the lowest peak correlation with respect to the EMF from the respective cluster and re-computing the corresponding EMF and threshold. This process continues by peeling away one trainer at a time as long as the EMF threshold is below T_0.

4. Filter Performance

This section provides the filter performance assessment results pertaining to the ability of the EMFB to distinguish between the images of target and non-target classes of objects. A large number of experiments were performed utilizing the images obtained from the Columbia Object Image Library (COIL) database (Nene, 1996). In order to illustrate the EMFB operation and assess the effects of parameter settings on the filter performance a representative subset of the experimental results are presented here. The COIL dataset contains color images of one-hundred objects whose geometry and surface reflectance cover wide ranges of variability. The dataset includes seventy-two 128×128- pixel images of each object taken against black background at uniform range and equally spaced view angles, where consecutive view angles are five degrees apart.

Twelve of the COIL objects including *tank, fire-engine, truck, pickup,* and eight *cars* were used in the experiments reported here. Each color image was converted to grayscale intensity image and was down sampled to 64×64- pixels. Each image was then rescaled into five sizes resulting in 360 images for each object. The experimental dataset, therefore, comprises 4320 images, including twelve objects, seventy-two view angles, and five scales. At each view angle, image scales for each object vary from 58×58 to 70×70-pixels. Figure 1, on the left shows the images of twelve objects which are used in our experiments, at 0° view angles and uniform scales. Shown on the right side of Figure 1 are images of six of the experimental objects, each shown in six different view angles at uniform scales.

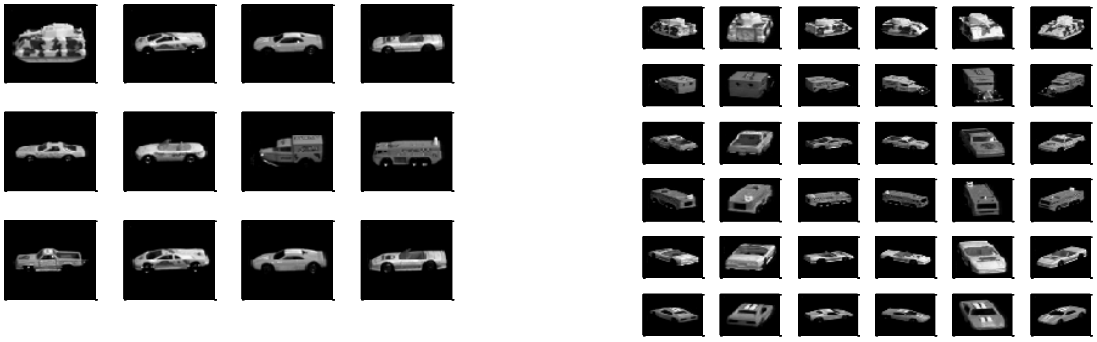

Fig. 1. Images of twelve COIL objects used in the experiments are shown on the left. Images of *tank, truck, pickup, fire-engine,* and two *cars* at different view angles are shown on the right, where consecutive images in each row are fifty degrees apart. All the images shown here are 64×64-pixels.

In the first experiment the *tank* was declared as the target class object. A randomly selected subset of the 360-image *tank* dataset was used in the filter training process. The remaining *tank* images, which were not used in the training process, and all the 3960 images of eleven non-target class objects were used as the source from which test images were drawn. The filter training algorithm of Section 3 was utilized to construct the EMFB using the randomly selected *tank* trainer images. The resultant EMFB was subsequently utilized as a binary filter in order to assign one of two labels, namely, target and non-target to each one of the test images. The effect of the number of filters comprising the EMFB on the filter performance was assessed in this experiment. There are two types of error, namely, assignment of non-

target label to a *tank* test image, and assignment of target label to a test image pertaining to one of the eleven non-target objects. The number of target class *tank* images that were randomly selected and were used as the training set of images was set at ninety. The test set of images consists of 540 unlabeled images including 270 non-trained-on *tank* images and 270 non-target images randomly selected from the 3960-image set. The trainer set was utilized to build three sets of EMFB, comprised of six, nine, and twelve filters. The experiment was repeated 200 times for each setting of the number of EMFB filters, and the results were averaged for each setting across all the experiment runs. For each run of the experiment a different set of randomly selected *tank* images were used as trainers, and the non-trained-on tank images were combined with 270 randomly picked non-target images to test the EMFB. Figure 2, on the left shows the classifier performance results by plotting the receiver operating characteristics (ROC) of three EMFBs, consisting of six, nine, and twelve filters. Detection and false alarm rates denote, respectively, the percentage of target class test images that are correctly labeled by the EMFB, and the percentage of non-target class test images that are mislabeled. As expected, the classifier performance is improved as the number of filters is increased. The right hand side of Figure 2 shows all the correlation filters of the resultant nine-filter EMFB classifier for one instantiation of the experiment, with the respective nominal threshold values listed in the figure caption. The number of outliers in the particular instantiation of the experiment that resulted in the EMFB shown here was seventeen, which means 73 trainers were utilized in the synthesis of this filter.

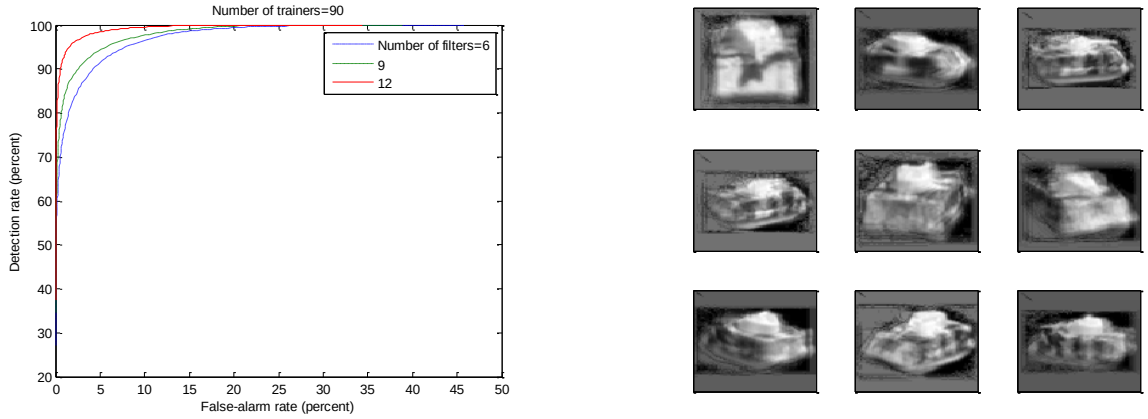

Fig. 2. Left shows the *tank*-EMFB performance results with the number of filters (EMFs) as the parameter. Shown on right are the correlation filters for a nine-filter EMFB. Nominal thresholds for EMF one through nine are 0.6805, 0.7261, 0.6818, 0.6788, 0.7930, 0.6714, 0.8822, 0.6808, 0.6795, where filter orders increase from left to right and top to bottom. All the images on the right are $70' \; 70-$ pixels.

In the example of Figure 3 *truck* was declared as target class and eleven other types of vehicle were used as non-target class test objects. The number of filters comprising the EMFB was set at nine, and the effect of the number of trainers on the filter performance was assessed. As before, a randomly selected subset of *truck* images was used as the training set. The remaining non-trained-on *truck* images and equal number of non-targets, randomly selected from 3960 images, were used as the test image set. The ROC plots on the left hand side of Figure 3 show the performance of the nine-filter EMFB, where the number of trainers is the control parameter. As before, for each parameter setting the experiment was repeated two-hundred times and the EMFB performance was averaged across all trials of the experiment. It is seen that increasing the number of trainers from 36 to 90 results in a slight improvement of EMFB performance. Increasing the number of trainers to 180 in this experiment, however, results in inferior performance, which seems to be counterintuitive. This result may be attributable to the potentially deleterious effects of overtraining, which happens when the number of trainers is increased beyond some optimal number. The right hand side of Figure 3 shows the correlation filters of the *truck*-EMFB corresponding to one instantiation of the experiment, where nominal filter thresholds are listed in the

caption. The number of outliers in the instantiation of the experiment that resulted in the classifier shown on the right of Figure 3 was 21, which means 69 trainers were used in the synthesis of the EMFB.

The next experiment involves declaration of the *fire-engine* as the target class. The number of training images was set at ninety. The test set of images comprise 540 images including 270 no-trained-on *fire-engine* and 270 images randomly selected from eleven non-target image sets. The left hand side of Figure 4 shows the ROC plots pertaining to EMFBs comprised of three, six and nine filters, averaged across two hundred instantiations of the experiment for each case. Shown on the right are the correlation filters of the nine-filter EMFB that pertain to one instantiation of the experiment.

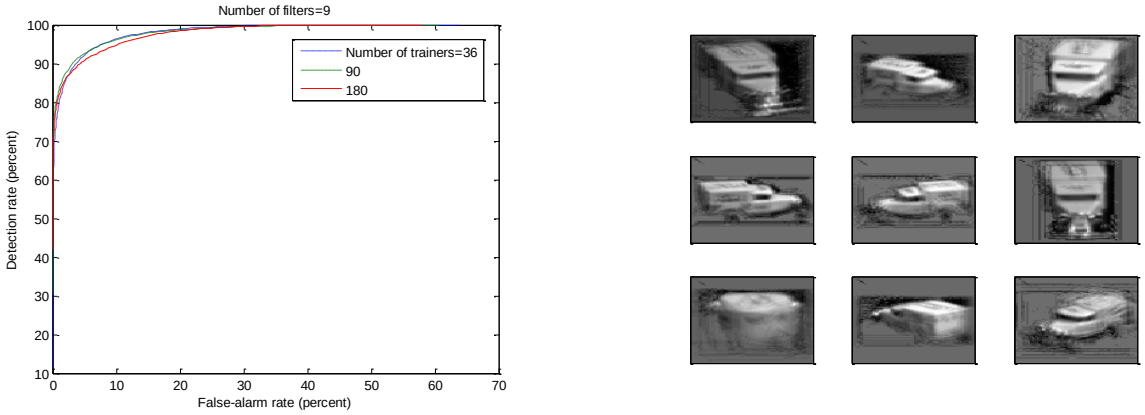

Fig. 3. Left shows the effect of the number of trainers on *truck*-EMFB performance. Shown on right are the correlation filters for the EMFB corresponding to one instantiation of the experiment using 90 trainers. Nominal thresholds for EMF one through nine are 0.7473, 0.7011, 0.7346, 0.7131, 0.6786, 0.8420, 0.6589, 0.7422, 0.6585. All the images on right are 70′ 70- pixel.

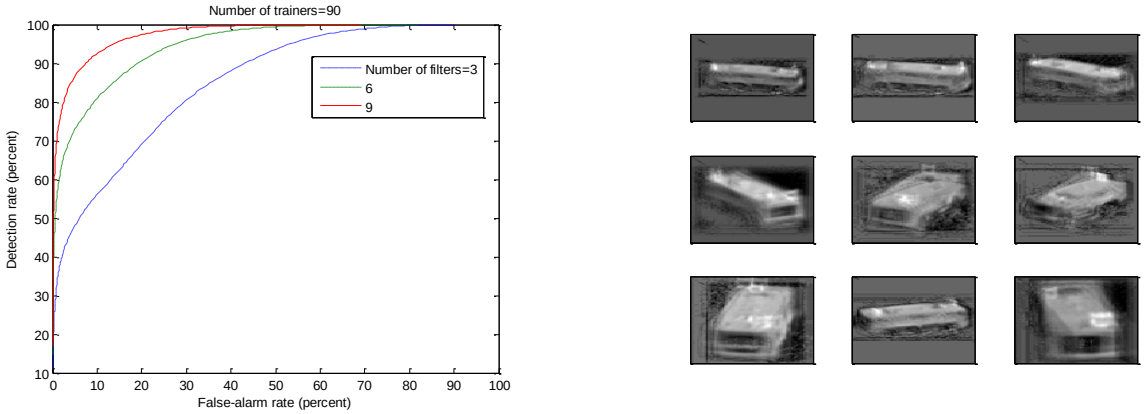

Fig. 4. Left shows the f*ire-engine*-EMFB performance results. The shown on right are the correlation filters for a nine-filter EMFB. Nominal thresholds are 0.6821, 0.6620, 0.6978, 0.6681, 0.6507, 0.6533, 0.6943, 0.6727, 0.6664. All the images on right are 70′ 70- pixel.

In the next example, the performance of the EMFB classifiers pertaining to three vehicle types, namely, *tank*, *truck*, and *fire-engine* are assessed using a different test scenario. For each option of the target class vehicle, the non-target class test images comprise two vehicle types instead of eleven as in the previous cases. This was done in order to test the capability of the classifier under more challenging conditions, where there are higher inter-class degrees of similarity than before. The trainer set consists of 90 randomly selected images of a particular vehicle, and the test set comprises 540 images including 270 non-trained-on target class and 270 non-target class images, equally distributed between the other two vehicles, and randomly selected. Figure 5 shows the ROC plots of the six-filter and the twelve-filter

EMFBs of each vehicle. Plotted in Figure 6 are the intra-class and inter-class distortion histograms, where the horizontal and vertical axes denote, respectively, mutual distortion and the percentage of image pairs whose mutual distortion values are within the specified bands. Here, mutual distortion between two images is defined as the difference between one and peak-correlation. In each plot, only the results for distortion values greater that 0.5 are shown, where dots and circles denote, respectively, intra-class and inter-class distributions, where distortions of each image are computed with respect to images of the same vehicle type as well as those of two other types of vehicles. Plots of Figure 6 show that within class image variability is comparable to cross-class variability.

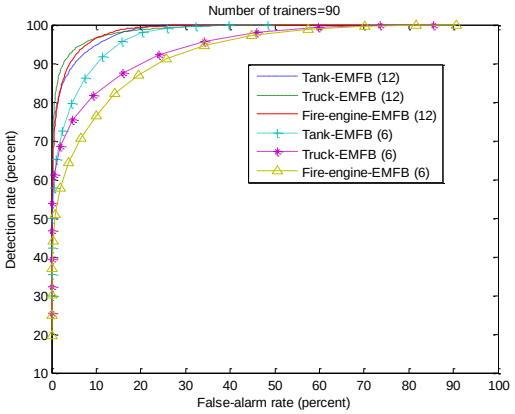

Fig. 5. Effect of number of filters on classifier performance. The non-target test set comprises 540 images equally divided between two objects.

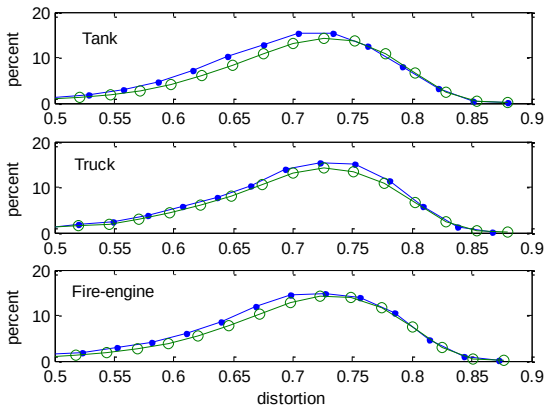

Fig. 6. distributions of image pair numbers with respect to mutual image distortions. Dots and circle correspond to intra-class and interclass distortions.

5. Conclusions

This paper presents an effective machine learning algorithm for computing robust image classifiers. The training algorithm is computationally efficient, and the resultant classifier can be applied to sensor images with low latency using modest data storage and processing resources. The EMFB classifier is trained using exemplar images representing only the target class of interest in various states and view conditions, and it does not rely on any information regarding the non-target class images that it is required to guard against. The EMFB comprises a small number of Fourier filters that can be applied to sensor images in parallel in real-time target recognition and tracking applications. The EMFB training and performance has been studied using realistic test scenarios and challenging problems, and some representative results have been presented. It has been shown that the EMFB provides concurrently high probability of detection and low probability of false alarm. Formal analysis of the computational complexity of the EMFB classifier training process, and the effects of lighting variations on the classifier performance will be reported in a forthcoming paper. Work in the areas of EMFB applications to real-time video processing and target tracking is ongoing. Efforts on the development of EMFBs for detection of partially obscured targets in sensor images is also ongoing. One of the areas of related future research includes development of EMFBs for color and multispectral imagery data.

Acknowledgements

Partial support for this research was provided by U.S. Department of Defense through contract W911NF-13-1-0136.

References

Belongie S., Malik J., Puzicha J. (2002). Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 24, 509-522.

Casasent D., Ravichandran G., Bollapragada S. (1991). Gaussian-minimum average correlation energy filters. Applied Optics, Vol. 30, No. 35, 5176-5181.

Casasent D., Ravichandran G. (1992). Advanced distortion-invariant minimum average correlation energy (MACE) filters. Applied Optics, Vol. 31, No. 8, 1109-1116.

Caulfield H.J., Weinberg, M.H. (1982). Computer recognition of 2-D patterns using generalized matched filters. Applied Optics, Vol. 21, No. 9, 1699-1704.

Nene A., Nayar S.K., Murase H. (1996). Columbia Object Image Library (COIL-100). Columbia University Technical Report No. CUCS-006-96.

Davies E.R. (2012). Computer and Machine Vision: Theory, Algorithms, Practicalities, Academic Press.

Georghiades A.S., Belhumeur, P.N. and Kriegman D.J. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 23, Issue 6, 643-660.

Heidary K., Caulfield H.J. (2005). Application of supergeneralized matched filters to target classification. Applied Optics Vol. 44 No. 1, 47-54.

Heidary K., Caulfield H.J. (2009). Nonlinear Fourier correlation. Proceedings of SPIE Vol. 7340A, 1-11.

Heidary K. (2013). Distortion tolerant correlation filter design. Applied Optics, Vol. 52, No. 12, 2570-2576.

Heidary, K. (2014). Fourier filter augmented with trainer histograms. Applied Optics, Vol. 53, No. 28, 6464-6471.

Johnson R.B., Heidary K. (2011) A unified approach for database analysis and application to ATR performance metrics,' Proceedings of SPIE, Vol. 7696Z, 1-20.

Kumar B.V.K. (1986). Minimum-variance synthetic discriminant functions. JOSA A, Vol. 3, No. 10, 579-1584.

Landeau S., Dagobert T. (2006). Image database generation using image metric constraints: an application within the CALADIOM project. Proc. SPIE, Vol. 6234, 623410-1–623410-12.

Mahalanobis A., Kumar B.V.K., Casasent D. (1987). Minimum average correlation energy filters. Applied Optics, Vol. 26, No. 17, 3633-3640.

Nene A., Nayar S.K., Murase H. (1996). Columbia Object Image Library (COIL-100). Columbia University Technical Report No. CUCS-006-96.

Savvides M., Kumar B.V.K. (2003). Quad Phase Minimum Average Correlation Energy Filters for Reduced Memory Illumination Tolerant Face Authentication. Lecture Notes in Computer Science, Vol. 2688, 1056-1065.

Turin G. (1960). An introduction to matched filters. IRE Trans. Information Theory, Vol. 6, issue 3, 311-329.