Wavelet packet correlation methods in biometrics

Pablo Hennings, Jason Thornton, Jelena Kovačević, and
B. V. K. Vijaya Kumar

We introduce wavelet packet correlation filter classifiers. Correlation filters are traditionally designed in the image domain by minimization of some criterion function of the image training set. Instead, we perform classification in wavelet spaces that have training set representations that provide better solutions to the optimization problem in the filter design. We propose a pruning algorithm to find these wavelet spaces by using a correlation energy cost function, and we describe a match score fusion algorithm for applying the filters trained across the packet tree. The proposed classification algorithm is suitable for any object-recognition task. We present results by implementing a biometric recognition system that uses the NIST 24 fingerprint database, and show that applying correlation filters in the wavelet domain results in considerable improvement of the standard correlation filter algorithm. © 2005 Optical Society of America

OCIS codes: 100.5010, 100.6740, 100.7410, 100.2960, 100.4550.

1. Introduction

Correlation filter classifiers are typically designed and applied in the image-intensity domain. In applications such as biometric recognition, the images of interest (e.g., iris or fingerprint images) are often characterized by localized high-frequency features. The image domain may not be the most convenient characterization of these images and may not represent the features that effectively distinguish one class from another. In contrast, wavelet analysis provides space–frequency localization that may be more effective for classification. The goal of this study is to find the wavelet spaces for an image class that are most conducive to correlation filter recognition.

The theory of wavelet analysis, which in the past two decades has been extensively documented,[1,2] describes how wavelet packet bases are used to represent signals effectively. This is achieved by partitioning the time–frequency plane into regions with different resolutions in frequency and time (or, in the case of images, space) domains. Wavelet packets can be implemented efficiently and have therefore proved to be practical in many different applications in which a good time–frequency or space–frequency resolution trade-off is needed. This is why we believe that wavelet analysis can be advantageous to correlation filter classifiers. The proposed algorithm uses a pruning method to search for the best wavelet packet subspaces and computes a set of subspace correlation filters for each class. During testing, this set of filters is applied to the corresponding subspaces of a test image; each output subspace correlation plane is then used to compute a final match score for this image.

Applying correlation filters in the wavelet domain is a new technique recently used by Daniell *et al.*[3] for fast recognition of compressed imagery available in the subband domain. We address a different problem in this paper. Instead of aiming for efficiency in the recognition of subband coefficients, we start with uncompressed images and address the problem of increasing recognition accuracy by finding optimal wavelet spaces. In particular, we use the NIST 24 fingerprint database to show the improvements that can be achieved by using this method instead of standard correlation filter algorithms. In the case of fingerprint recognition, using the proposed wavelet-domain algorithm is an attractive alternative to previous studies that used standard correlation filter algorithms.[4]

The paper is organized as follows. We first formulate the biometric problem (Section 2) and then present an overview of correlation filter classifiers and of wavelet transforms (Sections 3 and 4, respectively). Next we describe the proposed algorithm (Section 5), and, finally, the new classification scheme is

All the authors are with Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania 15213. P. H. Hennings Yeomans (phh@ece.cmu.edu), J. Thornton, J. Kovačević, and B. V. K. Vijaya Kumar are with the Department of Electrical and Computer Engineering. J. Kovačević is also with the Department of Biomedical Engineering.

Received 9 June 2004; revised manuscript received 19 October 2004; accepted 19 October 2004.
0003-6935/05/050637-10$15.00/0

evaluated and compared with standard correlation filter techniques (Section 6).

2. Biometric Problem Formulation

To start the discussion, we first formulate the problem. We are given an image $\mathbf{x} \in \mathbb{R}^N \times \mathbb{R}^N$, which goes either through an identification or a verification system, after which a decision needs to be made.

A. Identification System

Given p image-class templates $\mathbf{y}_i$, $i = 1, \ldots, p$, that correspond to p individuals stored in a database, we need to find the closest match to our input $\mathbf{x}$ as follows:

$$\hat{\mathbf{y}} = \mathbf{y}_k \qquad \text{if } f(\mathbf{x}, \mathbf{y}_k) = \min_{\mathbf{y}_i} f(\mathbf{x}, \mathbf{y}_i), \tag{1}$$

where $f(\mathbf{x}, \mathbf{y})$ is a suitably chosen cost function that is dependent on the application.

B. Verification System

Given the input $\mathbf{x}$ that claims to belong to the class $\mathbf{y}_k$, we need to verify whether this is true. The answer a is a binary "yes" or "no":

$$a = \begin{cases} \text{yes} & \text{if } f(\mathbf{x}, \mathbf{y}_k) \le T \\ \text{no} & \text{otherwise} \end{cases}, \tag{2}$$

where T is a given threshold.

3. Review of Correlation Filters

Correlation filters have been successfully applied in biometric recognition.[5–7] In particular, correlation filters have been demonstrated to work well for fingerprint recognition.[4]

Filtering an image is equivalent to performing a spatial cross correlation between the image and the flipped impulse response of the filter. If the filter matches the image (i.e., there is strong correlation between the two), the output will contain a correlation peak. Advanced correlation filters are designed to produce correlation peaks when they are applied to any image containing a specific pattern class (such as fingerprints from the same finger).[8]

A correlation filter takes the form of a two-dimensional complex-valued array in frequency and is applied as shown in Fig. 1. The image is converted into the frequency domain with a fast Fourier transform (FFT), multiplied by the filter, and converted back to the image domain by an inverse fast Fourier transform; this produces a correlation plane. This process is equivalent to, but much faster than, computing the cross correlation in the original image domain. This process, which produces a correlation plane from a given image, can be modeled as

$$\mathbf{v}_i = (\mathbf{DFT})^{-1}\, \mathbf{C}_i\, \mathbf{DFT}\, \mathbf{x}, \tag{3}$$

where $\mathbf{x}$ is the vectorized form of the n image pixels, $\mathbf{DFT}$ is an $n \times n$ matrix containing the basis of a

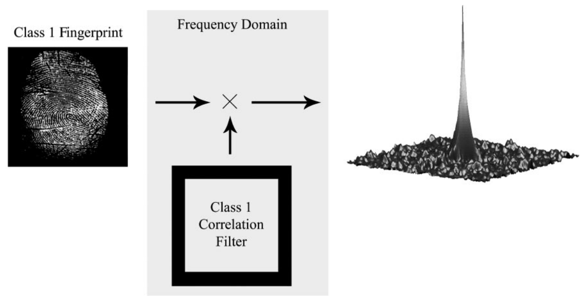


Fig. 1. Application of a correlation filter in the frequency domain.

two-dimensional discrete Fourier transform, and $\mathbf{C}_i$ is a diagonal matrix containing the correlation filter values (designed for class i) in the Fourier domain along the diagonal.

If the image belongs to the pattern class of the filter, the correlation plane output contains a sharp peak (as in Fig. 1); if not, no such peak exists. We derive a match metric from the correlation plane by measuring the peak-to-correlation energy (PCE), defined as

$$\mathrm{PCE}(\mathbf{v}_i) = \frac{\max(|\mathbf{v}_i|) - \mathrm{mean}(|\mathbf{v}_i|)}{\mathrm{stdev}(|\mathbf{v}_i|)}, \tag{4}$$

which yields our cost function

$$f(\mathbf{v}_i) = \frac{1}{\mathrm{PCE}(\mathbf{v}_i)}. \tag{5}$$

Note that correlation filters are shift invariant (i.e., a shift in the input results in a corresponding shift in the correlation plane); the PCE remains constant because it is computed after the correlation peak is located.

When correlation filters are designed based on a set of multiple training images, they are called composite filters. The design options for composite filters have been thoroughly studied.[8,9] Each design satisfies a different criterion for the correlation planes generated by the training images.

For the purpose of this paper, we consider the minimum average correlation energy (MACE) filter.[10] In the MACE filter design, the origin of each training-image correlation-output plane is constrained to equal some specific value (typically one for training images that belong to the class of interest and zero for training images that do not). The correlation filter usually has many more degrees of freedom than are needed to satisfy this set of constraints. Thus, in addition, the MACE filter design minimizes the average correlation plane energy (ACE), which can be expressed as

$$\mathrm{ACE} = \mathbf{h}^{+}\mathbf{D}\mathbf{h}, \tag{6}$$

where $\mathbf{h}$ contains the correlation filter values in the

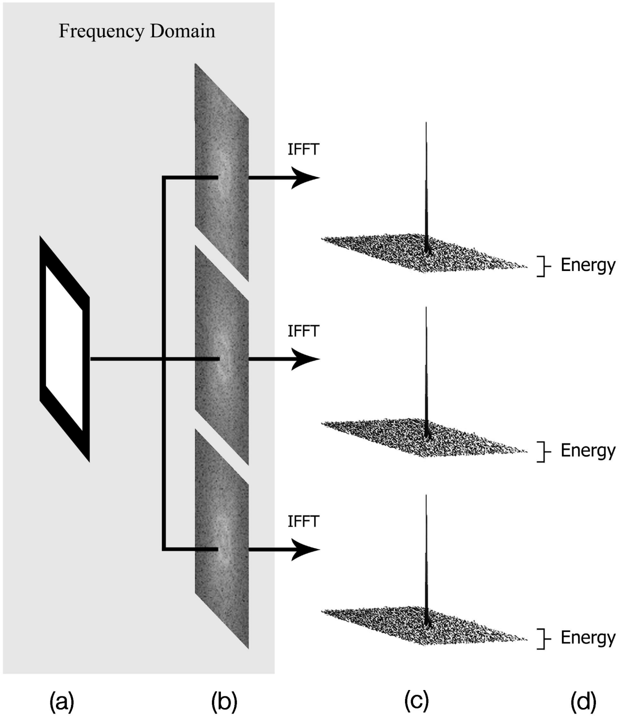


Fig. 2. MACE-constrained minimization problem: (a) correlation filter to be designed; (b) transforms of the training images; (c) correlation planes that result when the filter is applied to each training image; (d) energy of each correlation plane. The filter in (a) is designed to give specified correlation peaks at the origins of (c) while minimizing the combined energies at (d). Simply stated, the filter is designed to give the sharpest-possible correlation peaks for the within-class training set.

frequency domain, ordered into a column vector, and $\mathbf{D}$ is a diagonal matrix with the average power spectrum of the training images along its diagonal. Figure 2 shows the constrained minimization problem that the MACE filter design solves. The measure ACE is minimized because lower correlation plane energy means sharper correlation peaks and better PCE scores for images belonging to the filter's class. The filter that meets these design criteria is given by

$$\mathbf{h} = \mathbf{D}^{-1}\,\mathbf{X}(\mathbf{X}^{+}\,\mathbf{D}^{-1}\,\mathbf{X})^{-1}\,\mathbf{u}, \tag{7}$$

where matrix $\mathbf{X}$ contains the Fourier transforms of the training images, one in each column, and vector $\mathbf{u}$ contains the correlation plane origin constraints.

Robustness to noise can be achieved instead by minimization of the output variance of the correlation peak. If the noise in the training set is additive zero mean and stationary, with power spectral density values arranged in a diagonal matrix $\mathbf{P}$, then the measure minimized is the output noise variance (ONV),

$$\mathrm{ONV} = \mathbf{h}^{+}\mathbf{P}\mathbf{h}. \tag{8}$$

This design is the minimum variance synthetic discriminant function filter proposed by Kumar.[11]

In this paper we use the optimal trade-off synthetic discriminant function (OTSDF) filter design.[12–14] The OTSDF can be considered to be a generalized form of the MACE filter, with a parameter that allows for within-class noise tolerance. Specifically, the OTSDF filter finds a compromise in the minimization of the ACE and the ONV measures by minimizing the energy function

$$E(\mathbf{h}) = \alpha(\mathrm{ONV}) + \beta(\mathrm{ACE}), \tag{9}$$

where $\beta = \sqrt{1-\alpha^2}$. Then the OTSDF filter is given by

$$\mathbf{h} = (\alpha\mathbf{I} + \beta\mathbf{D})^{-1}\,\mathbf{X}(\mathbf{X}^{+}(\alpha\mathbf{I} + \beta\mathbf{D})^{-1}\,\mathbf{X})^{-1}\,\mathbf{u}, \tag{10}$$

where $\mathbf{I}$ is the identity matrix. (We have modeled the noise in the training set as additive and white.) The advantage of this filter is that it allows for the minimization of the energy in the correlation plane while adjusting for noise tolerance.

4. Review of Wavelet Transforms

The theory of wavelets and multirate signal processing have become standard signal processing tools. Iterated multirate filter banks perform decompositions of signals into wavelet coefficients. The main characteristic of wavelet transforms is their capability to represent a signal by achieving a convenient partition of the time–frequency plane (or the space–spatial frequency for images). These partitions provide us with features that have joint locality in space and frequency, which can be valuable for biometric image characterization.

Wavelets, filter banks, and wavelet packets. Filter banks are signal processing devices used to implement wavelet transforms. With filter banks we can construct orthonormal bases for the space of finite-energy sequences. In particular, we can project a signal onto spaces of low-pass and bandpass signals.[1]

To achieve this, we start by filtering the signal by using a low-pass filter that captures the coarse approximation and a complementary high-pass filter that captures the details. The impulse responses of these filters and their even translates (versions of the filter that have been shifted by a multiple of 2) are designed to form an orthonormal basis.[1] The filter outputs are downsampled by two to avoid redundancy. The resulting device is called an orthogonal, critically sampled, two-channel analysis filter bank, depicted in Fig. 3(a). Conceptually, this filter bank splits a space V_0 into subspaces V_1 and W_1, where V_1 contains the approximation of the signal and W_1 contains the detail information. The iteration of this filter bank on the low-pass channel, as shown in Fig. 3(b), computes the discrete wavelet transform (DWT) coefficients.[1,2]

A generalized iterated filter bank algorithm that allows the high-pass channel to be decomposed as well is the discrete wavelet packet transform, introduced by Coifman *et al.*[16] With this generalization we

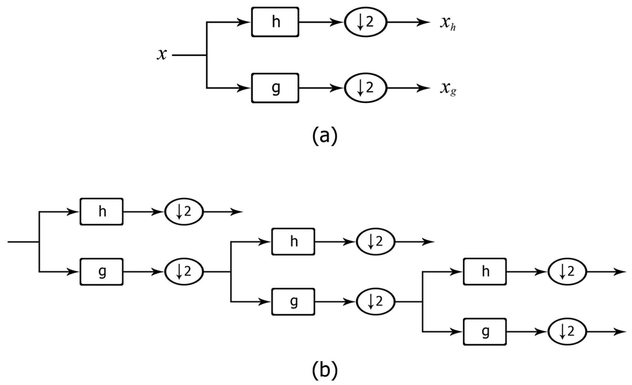


Fig. 3. (a) Two-channel analysis filter bank structure used for wavelet analysis. (b) Multichannel filter bank algorithm that uses two-channel filter bank blocks iteratively on the low-pass channel (DWT with three levels).

can design wavelet bases that achieve an almost-arbitrary time–frequency tiling.

At iteration level i, the discrete wavelet packet transform computes an orthonormal expansion in each space V_i and W_i. Specifically, given an orthogonal low-pass filter g and its complementary high-pass filter h, both of finite support, the single-level expansion of a signal $x^{(i)}$ at level i is given by

$$x^{(i)}[n] = \sum_{k\in\mathbb{Z}} x_g^{(i+1)}[k]g[n-2k] + \sum_{k\in\mathbb{Z}} x_h^{(i+1)}[k]h[n-2k], \tag{11}$$

where

$$x_g^{(i+1)}[n] = \sum_{k\in\mathbb{Z}} x^{(i)}[k]g[k-2n], \tag{12}$$

$$x_h^{(i+1)}[n] = \sum_{k\in\mathbb{Z}} x^{(i)}[k]h[k-2n]. \tag{13}$$

We have used superscripts to indicate the level of decomposition that the signals belong to and subscripts to indicate whether a signal is generated through the low-pass channel or the high-pass channel. This can also be expressed in matrix notation as

$$\mathbf{x}_l = \mathbf{WP}_l\,\mathbf{x}, \tag{14}$$

where $\mathbf{x}$ is the input image vector, $\mathbf{WP}_l$ is the matrix computing the wavelet packet coefficients in subspace l, and $\mathbf{x}_l$ is the output image in subspace l.

We can use these equations iteratively to compute all wavelet packet coefficients for a given wavelet tree. All wavelet packet decompositions in this paper are trees in which each node has either 0 or 2 children (or 0 or 4 in the two-dimensional case); such trees are called *admissible trees*,[2] and their leaf subspaces contain the complete information to reconstruct the original signal. Note that a full decomposition tree is also an admissible tree.

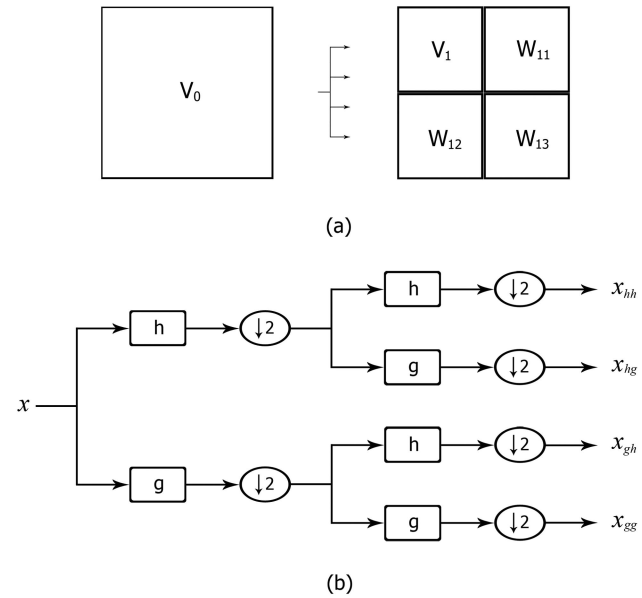


Fig. 4. (a) Single-level splitting of an image space V_0 into four spaces V_1, W_{11}, W_{12}, and W_{13} with quadtree implementation. (b) Quadtree implementation of single-level wavelet packet decomposition.

In the case of images, wavelet packet decompositions are implemented by iterating quadtrees, which divide the two-dimensional frequency plane into square regions of different sizes. The initial splitting of space V_0 is into four spaces, V_1, W_{11}, W_{12}, and W_{13} [see Fig. 4(a)].

In this study we use only two-dimensional wavelet packets whose basis functions are separable products of one-dimensional wavelet packet basis functions[2]; i.e., the quadtree filter bank implementation uses one-dimensional filters applied along the rows and along the columns of a two-channel orthogonal filter bank structure as shown in Fig. 4(b). This generates four subspaces from each parent space.

Our goal is to find the best admissible tree that provides the subspaces for correlation filter–based classification.

5. Wavelet-Domain Correlation Filters

Correlation filters are traditionally applied to image intensities, but they may be applied to other image features instead, provided the feature space maintains a consistent spatial relationship. For this reason, filters may be designed in individual wavelet subspaces of an image. Instead of designing a single correlation filter for a pattern class, we propose to design a correlation filter for each leaf in the best wavelet packet admissible tree that we can find for that class. To do this, we define a measure of performance for correlation filter classifiers as a criterion for searching for the best set of subspaces; this will help us to discriminate between subspaces that are more useful for recognition than the original image domain.

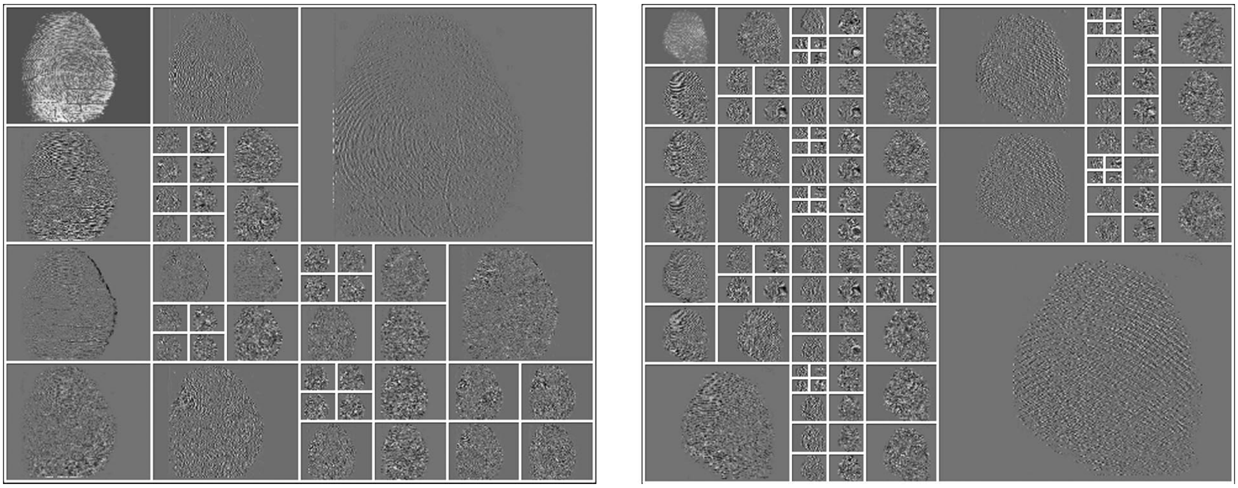


Fig. 5. Samples of image decomposition into subspaces with the best wavelet packet for two different fingerprint classes.

A. The Fitness Metric

This metric provides a score for how well we expect a correlation filter to perform in a given subspace. Because we achieve sharp correlation peaks by minimizing the average correlation energy, we propose to use the minimum average correlation plane energy that a filter can achieve on a subspace as a performance metric. This is derived by substituting Eq. (7) into Eq. (6) and simplifying as follows:

$$\begin{aligned}\tilde{E} &= (\mathbf{D}^{-1}\,\mathbf{X}(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X})^{-1}\,\mathbf{u})^{+}\,\mathbf{D}(\mathbf{D}^{-1}\,\mathbf{X}(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X})^{-1}\,\mathbf{u}) \\ &= \mathbf{u}^{T}\,(\mathbf{X}^{+}\mathbf{D}^{-1}\mathbf{X})^{-1}\,\mathbf{u}. \qquad (15)\end{aligned}$$

In this case each column of matrix $\mathbf{X}$ holds the spectrum of one of the training images after projecting the image onto the wavelet subspace. Because lower energy implies a better performance, we define the fitness metric as

$$F = 1/\tilde{E}. \qquad (16)$$

Intuitively, this is a good metric because it measures how sharp we can make the correlation peaks in this subspace. The higher the metric, the better the correlation filter is expected to perform.

Note that the fitness metric is directly related to the size of the space it is applied to. This means that if we use a four-channel wavelet filter bank to decompose space V_0 into four spaces, V_1, W_{11}, W_{12}, and W_{13}, and we wish to evaluate their relative fitness for correlation filter classification, then we can evaluate the inequality

$$F(V_0) > F(V_1) + F(W_{11}) + F(W_{12}) + F(W_{13}), \qquad (17)$$

where operator $F(\cdot)$ computes the fitness metric of the space. Here V_1, W_{11}, W_{12}, and W_{13} are the spaces of signals $x_{gg}^{(i+1)}$, $x_{gh}^{(i+1)}$, $x_{hg}^{(i+1)}$, and $x_{hh}^{(i+1)}$, respectively. The sum of the subspace fitness metrics is used in this comparison because the summed areas of the subspaces equal the area of the original space V_0.

If the left-hand side of inequality (17) is greater, we consider V_0 a better space for applying a correlation filter; but if the right-hand side of inequality (17) is greater, we consider the individual subspaces V_1, W_{11}, W_{12}, and W_{13} to be better suited for correlation filter classification.

B. Wavelet Packet Tree Pruning

In the training stage we need to find the best admissible wavelet packet tree to represent each pattern class. Then we compute the set of subspace correlation filters for each class by using the corresponding tree.

The process of obtaining the tree for a given class is as follows. We start with the full wavelet packet decomposition of all the images in the class and choose this tree as the best admissible wavelet packet tree available so far. Because we are working with two-dimensional images, any four leaves from the same parent node of the current admissible tree will be pruned unless the sum of the fitness metrics of the leaves is greater than the fitness metric from the parent node, as in inequality (17). This generates a new admissible tree, and the process continues iteratively across the tree until no leaves can be pruned. The final admissible tree is used as the wavelet packet tree for the class. Figure 5 shows samples of fingerprint images decomposed by use of such trees.

After pruning the tree, we build a correlation filter for each leaf subspace of the tree. This involves decomposing all the training images of the class by use of the tree just found. Then, for each subspace, a correlation filter is computed by use of the corresponding subspace images. For example, assume that we found the best wavelet packet tree for a class is a single quadtree, which decomposes the image space V_0 into subspaces V_1, W_{11}, W_{12}, and W_{13}; i.e., we would train four subspace correlation filters for this class. So we start with decomposing the M images in the class training set with this quadtree so that we have M subspace training images for each

Table 1. Training Algorithm for Each Class

FOR EACH image in the class

Compute the full wavelet packet decomposition of the image.

END

FOR EACH level i, from the $(I - 1)$th level up to 0

FOR EACH parent node

(i) Compute the fitness metrics F_k for every space V_1, W_{11}, W_{12}, W_{13}

(ii) Compute the fitness metric G for parent space V_0

(iii) Prune the leaves of spaces V_1, W_{11}, W_{12}, W_{13}, **IF** $G \geq \sum_k F_k$

(iv) **IF** the children are not pruned, define the fitness metric of the parent as $G = \sum_k F_k$

END

END

Table 2. Algorithm for Applying Wavelet Packet Correlation Filters

Compute the wavelet packet decomposition of the image by using the tree of the class of the filter.

Initialize the match metric M to zero.

FOR EACH subspace

Compute the PCE value and add it to the match metric.

END

subspace V_1, W_{11}, W_{12}, and W_{13}. Then we train a correlation filter for each subspace by using their corresponding subspace images. A filter is not trained for the parent image space V_0.

The training stage thus generates a tree of correlation filters for each class. As with any wavelet decomposition into admissible trees, the size of the leaf subspace filters add up to the size of the images in the class. Table 1 outlines the pseudocode for training correlation filters for each class.

C. Match Metric Computation

In the evaluation stage we require a rule to calculate a match metric when wavelet-domain filters are applied to a new image. After decomposing an image by using the filter's packet tree, we compute the subspace correlation planes. We then compute PCE values for all subspaces and define our match metric for the test image as the sum of all the PCE values.

This computation can be expressed as follows. For a new image $\mathbf{x}$ and wavelet packet correlation filters of class i contained in matrices $\mathbf{C}_{i,\,l}$ for every subspace l, the subspace correlation planes are given by

$$\mathbf{v}_{i,\,l} = (\mathbf{DFT})^{-1}\, \mathbf{C}_{i,\,l}\, \mathbf{DFT}\, \mathbf{WP}_{i,\,l}\, \mathbf{x}, \tag{18}$$

and the cost function becomes

$$f(\mathbf{v}_i) = \frac{1}{\sum_l \mathrm{PCE}(\mathbf{v}_{i,\,l})}. \tag{19}$$

Table 2 outlines the pseudocode for applying wavelet-domain correlation filters in the testing stage.

D. Maintaining Shift Invariance

Although correlation filters are shift invariant in the image-intensity domain, they are not shift invariant in the wavelet domain. This is because the wavelet decomposition is not a shift-invariant operation. At every level in a wavelet packet tree decomposition, the subspaces are downsampled by a factor of 2 in each dimension. As a result, a subspace at level i of a wavelet packet tree is periodically shift invariant with the factor 2^i in each dimension.

If a fingerprint image is shifted an arbitrary distance before it is decomposed with a wavelet packet tree, the result will not be perfect shifts of the wavelet subspaces. Instead, a shift in the original image domain will only periodically correspond to a shift in the wavelet subspaces. Then we can expect the shift invariance of correlation filters to erode somewhat when they are applied in the wavelet domain (i.e., recognition performance peaks and decreases with varying levels of image translation).

To solve this problem, we design one extra correlation filter in the original image domain. This filter is meant only for alignment preprocessing and does not contribute to the match metric computation. Before we decompose a test image, we apply this filter and examine the correlation plane, which if the image is authentic, will contain a correlation peak. In this domain the correlation peaks may not be sharp enough to yield a good discrimination, but they do indicate shifts of the original image. We simply find the location of the correlation peak and correct for any translation. Then we may continue with the decomposition and the use of subspace filters, which allow for superior discrimination. This technique maintains shift invariance without adding to the computational complexity of the algorithm.

To demonstrate the shift invariance of our proposed algorithm, we selected a sample test image and shifted it horizontally across a range of 20 pixels (including interpolated subpixel shifts). We applied wavelet-subspace correlation filters with and without using prefiltering for alignment, and the resulting match scores are plotted in Fig. 6. Without alignment, the PCE match scores are sensitive to the amount of translation. The periodicity of the match scores ranges over $2^3 = 8$ pixels, corresponding to the largest downsampling factor in our wavelet decompositions. However, with the aid of prefiltering, shift invariance is clearly restored.

E. Computational Complexity

We consider the computational complexity of applying correlation filters in both the image and wavelet

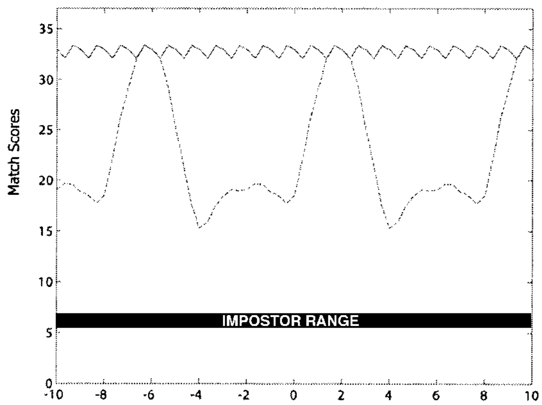


Fig. 6. PCE values for a range of possible horizontal shifts for a sample authentic image. A solid curve is used for the wavelet packet correlation filters with prefiltering, and a dashed curve for those without prefiltering. Shift invariance improves separation between the scores of authentic and impostor images.

domains; these are the relevant comparisons because the training stage is performed off-line. For simplicity, we assume that our images are of size $N \times N$, where N is a power of 2.

The most expensive operations in applying an image domain correlation filter are the FFTs, which require an order of $\mathcal{O}(N^2 \log N)$ multiplications. Then the entire filtering process is also $\mathcal{O}(N^2 \log N)$.

To apply a wavelet-domain correlation filter set to an image, we must first compute a wavelet packet decomposition of the image and then apply the set of filters on its subspaces. If we assume that the two-dimensional wavelet filters used for the decomposition are separable, composed of one-dimensional filters of length K, then the worst case of a full decomposition (with $\log_2 N$ levels) requires a computational complexity[2] of $\mathcal{O}(KN^2 \log N)$. Note that if we prefilter the image for alignment before performing the decomposition [an $\mathcal{O}(N^2 \log N)$ operation], the order of complexity does not increase.

The wavelet packet divides the image space into a set of smaller spaces, as illustrated in Fig. 5. Note that the cost of computing a FFT of the entire $N \times N$ space will always be greater than or equal to the combined cost of computing the FFT of each subspace. As stated above, the complexity of filtering is determined by the complexity of the FFTs. This means that applying a set of subspace filters always incurs less computational cost than applying a filter to the full space; so, in the worst case, the computation is $\mathcal{O}(N^2 \log N)$.

Combining the packet decomposition step with the subspace filtering step gives an overall complexity of $\mathcal{O}(KN^2 \log N)$. This is similar to the complexity of the image-domain correlation filters, with the added factor of the filter length K, typically a small constant. In summary, although the proposed technique does incur a greater computational cost to generate the wavelet subspaces that give better accuracy, it does not increase the order of complexity.

Fig. 7. Samples of fingerprint images from two different classes: a typical class (top row), and a class with particularly difficult distortions (bottom row).

6. Experiment Specifications and Results

A. Data Set

For testing, we used images from a subset of the NIST 24 fingerprint database.[17] Our data set consists of 20 classes from 2 different fingers, each class containing 100 images. When the images were captured, the subjects were instructed to roll their thumbs continually; the resulting variation makes these data more challenging but also more realistic for the recognition task. Because the image classes are influenced by subject behavior (i.e., subjects roll their fingers to varying extents), some classes are harder to classify than others.

Figure 7 shows sample images from two different classes of the data set. The upper images belong to a class that is representative of most of the data set. There is some variation in position and some deformation of the ridge patterns, but for the most part, the general fingerprint pattern remains. The lower images in Fig. 7 are from a relatively difficult class, in which the images show more extreme variations in position and blurring artifacts that distort the ridge patterns. In some images almost the entire print is missing.

For the classes in our data set that fit the second description, recognition with a single standard correlation filter is problematic. It has been shown that increasing the number of training images and using multiple filters in the image domain can improve performance.[4] However, to make a fair comparison, in our experiments we restrict our training data to a fixed number of images and expect to see improvement when the proposed algorithm is used.

B. Wavelet Filter Selection

To implement the wavelet-domain correlation filter algorithm in our tests, we had to select the type of wavelet filter to use in the decompositions. Ideally, the training algorithm would select for each image class the best wavelet filter from a full range of known wavelet filters. However, it is not practical to conduct such an exhaustive search. For simplicity, we worked only with wavelets from the Daubechies family.[15] After trying Daubechies filters of different lengths, the filter that performed best for all finger-

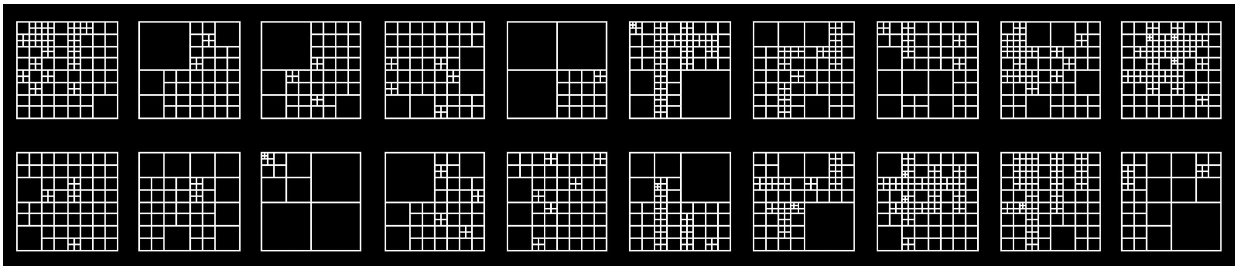


Fig. 8. Best decompositions for wavelet-domain correlation filters in each class, with classes 1–10 in the top row, and classes 11–20 in the bottom row.

print classes was the shortest, length-two filter: the Haar wavelet.

Short wavelets may extract more useful local space–frequency features from fingerprint images because the ridges that characterize these images have a narrow width. Employing other short-length wavelet filters may improve the algorithm's performance. Also, if the restriction to separable wavelet filters is relaxed (at increased computational cost), it may be advantageous to use wavelets that are specifically designed for curved structures, such as curvelets or ridgelets.[18,19] In our tests described in Subsection 6.C, we used the Haar wavelet for all fingerprint classes.

C. Evaluation

We sequestered eight images from each fingerprint class as training data. Using these images, we trained the two types of classifiers: standard correlation filters and wavelet-domain correlation filters, as described in Subsection 6.B. In both cases we designed OTSDF filters with the trade-off parameter set to 10^{-6} (which has been shown to be a reasonable value for fingerprint images[4]) and normalized the energy of the training images. Figure 8 shows the wavelet packet decompositions for each class.

The decompositions shown are dissimilar in structure and depth. Although for class 5 the tree further decomposes only the first-level high-pass channel W_{13}, some other trees do not iterate in this space but do iterate in V_1, W_{11}, and W_{12}, as is the case in classes 6, 13, 17, and 20. It is interesting to note that the decomposition found by the pruning algorithm for class 13 is actually the discrete wavelet transform. This dissimilarity is, of course, what will help us to distinguish among the classes effectively.

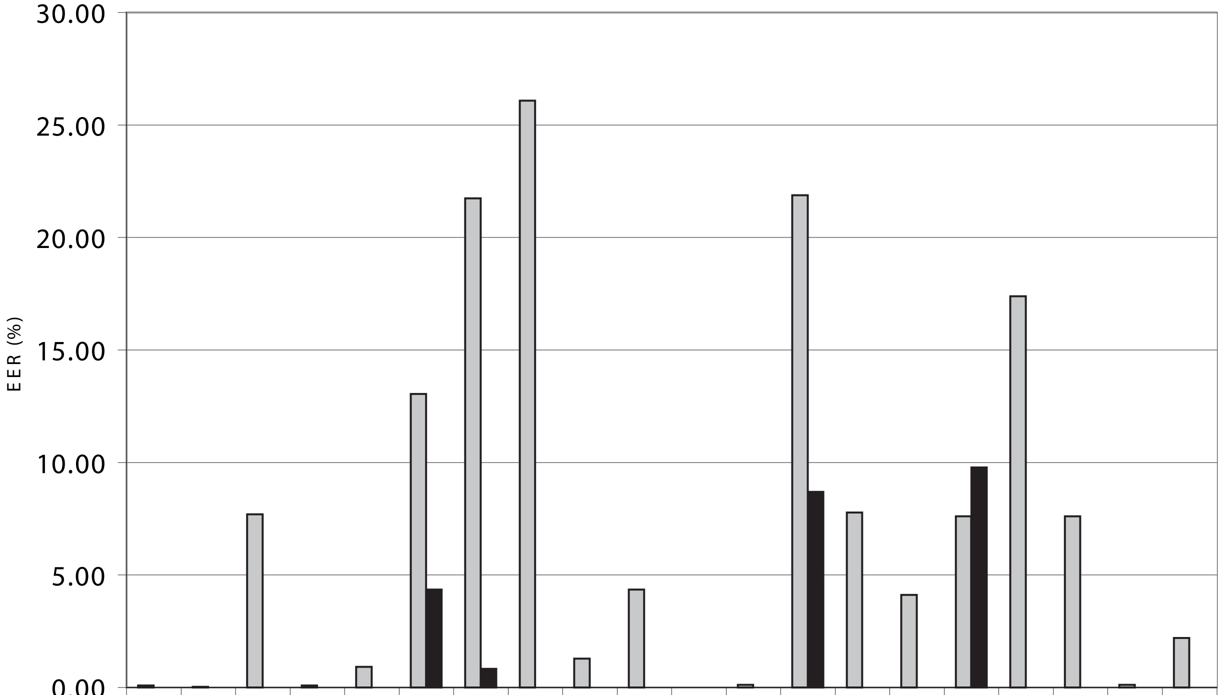
To compare the performance of the two sets, we conducted both verification and identification tests. In the verification tests every filter is applied to every test image to generate match scores. These scores are thresholded to produce "yes" or "no" match decisions. Depending on the threshold, the filter for each class accepts some images belonging to other classes (false acceptances) and rejects some images belonging to its own class (false rejections). When the threshold is set so that these error rates are the same, the resulting error rate is called the equal error rate (EER) of that filter. Figure 9 shows the EER of the filter that corresponds to each fingerprint class.

The results show a substantial overall improvement when wavelet-domain correlation filters are used. The average EER across all classes decreases from 7.21% for standard correlation filters to 1.18% for wavelet-domain correlation filters. Without the proposed algorithm, only 1 out of 20 filters provides complete score separation between authentic and impostor images (EER of 0); with the proposed algorithm, 16 out of 20 classes provide complete score separation. There is one class in which there is an increase in the error in the wavelet domain (from 7.6% to 9.8%). This result is plausible if the training set is not an adequate representation of the entire class. In other words, if the test images deviate dramatically from the sort of images included in the training set, learning a wavelet decomposition may offer no advantage. However, the advantage clearly exists for almost all of the test classes.

The variability in image quality from class to class is reflected in the results, with some classes proving much more difficult to recognize than others. The largest improvement occurs for class 8, which is the most difficult for the standard correlation filter. We note that it is possible for standard correlation filters to achieve much lower error rates (at or near zero), even for difficult classes, if the number of training images is increased and multiple correlation filters are designed instead of single filters. However, the purpose of our tests was to examine the relative improvement in accuracy with the proposed algorithm. Therefore we made the recognition problem a sufficiently challenging one by restricting the training set to eight authentic images per class.

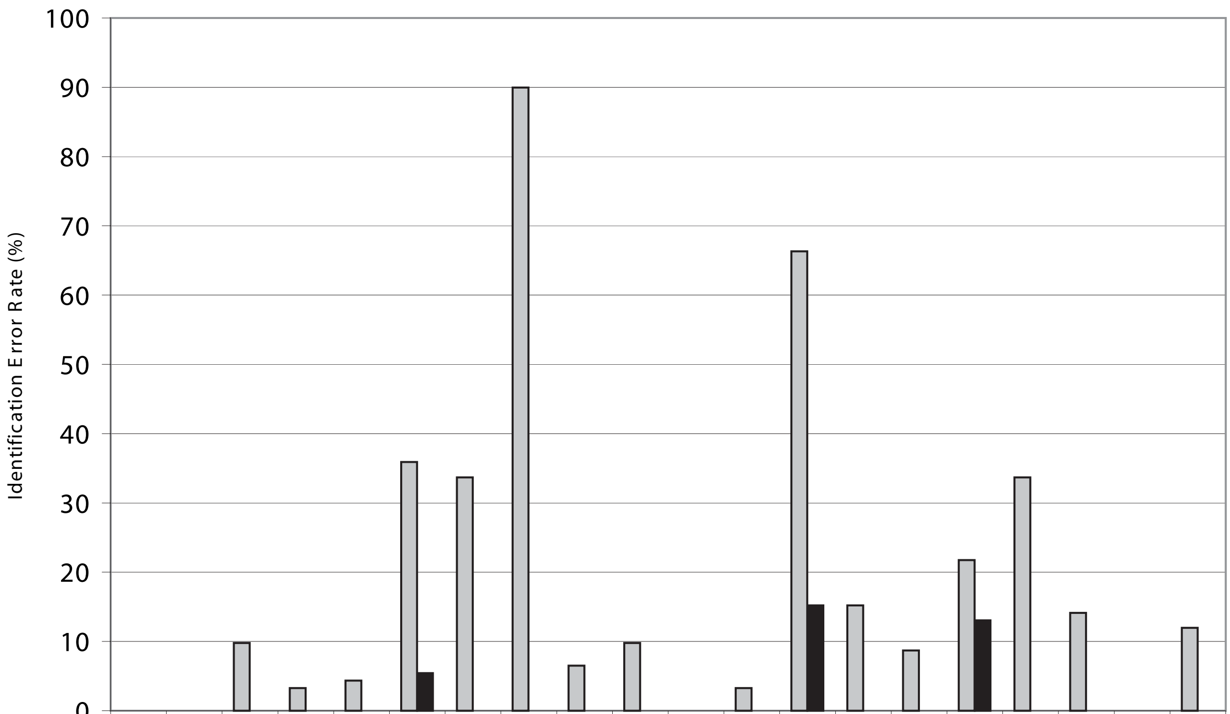
In identification tests, every test image must be identified by estimating to which class it belongs. This estimation is achieved by finding the highest match score from all filters and selecting that class. If the estimated class is incorrect, it is considered to be an identification error. Figure 10 shows the identification error rate for each class.

As expected from the verification results, a general improvement in all classes is achieved by the wavelet packet correlation filters, and the identification of class 8 proved to be the most difficult for the standard correlation filter algorithm. Again, using more train-



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
SCF	0.09	0.03	7.69	0.09	0.92	13.04	21.74	26.09	1.29	4.35	0	0.11	21.88	7.78	4.12	7.61	17.39	7.61	0.11	2.20
WDCF	0	0	0	0	0	4.35	0.83	0	0	0	0	0	8.70	0	0	9.78	0	0	0	0

Fig. 9. Verification ERR by class: SCF, standard correlation filters; WDCF, wavelet-domain correlation filters. SCF average EER = 7.21%; WDCF average EER = 1.18%.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
SCF	0	0	9.78	3.26	4.35	35.90	33.70	89.96	6.52	9.78	0	3.26	66.30	15.22	8.70	21.74	33.70	14.13	0	11.96
WDCF	0	0	0	0	0	5.43	0	0	0	0	0	0	15.22	0	0	13.04	0	0	0	0

Fig. 10. Identification error rate (IER) by class: SCF average IER = 18.41%; WDCF average IER = 1.68%.

ing images or correlation filters for this class would improve that performance. For the classes that exhibit especially poor identification results (e.g., classes 8 and 13), the score generated by applying the correct correlation filter is often the second- or third-highest score, instead of the maximum, which causes the poor error rates. This is not the case for the wavelet packet correlation filters, and zero error rates are obtained for all but three classes.

7. Conclusion

The wavelet packet correlation filter method, unlike the standard correlation filter method, makes use of image features that have joint locality in space and frequency. Although the wavelet-domain method requires more computational effort in the form of wavelet packet decompositions, it offers the flexibility to find more suitable spaces for correlation filter classification. As a result, the proposed algorithm achieves significantly better performance than standard filters on this data set. This result suggests that these fingerprint images have some features that remain more consistent in the underlying wavelet subspaces than in the spatial domain.

One of the components of the proposed algorithm is the choice of the wavelet-domain spaces for building and applying the filters. Apart from considering the minimum average correlation energy, a fitness metric may also be designed for the wavelet packet tree pruning algorithm to account for other qualities of correlation outputs, such as the maximum average correlation peak.[20] The fitness metric can be tailored to a particular problem. This is left for future study.

This study was supported in part by CyLab at Carnegie Mellon University, by the National Council of Science and Technology (CONACYT) of Mexico, and the Pennsylvania State Tobacco Settlement, Kamlet–Smith Bioinformatics Grant.

References

1. M. Vetterli and J. Kovačević, *Wavelets and Subband Coding* (Prentice-Hall, Englewood Cliffs, N.J., 1995).
2. S. Mallat, *A Wavelet Tour of Signal Processing* (Academic, San Diego, Calif., 1999).
3. C. Daniell, A. Mahalanobis, and R. Goodman, "Object recognition in subband transform-compressed images by use of correlation filters," Appl. Opt. **42,** 6474–6487 (2003).
4. K. Venkataramani, "Reduced complexity correlation filters for fingerprint verification," M.S. thesis (Carnegie Mellon University, Pittsburgh, Pa., 2002).
5. B. V. K. Vijaya Kumar, M. Savvides, K. Venkataramani, and C. Xie, "Spatial frequency domain image processing for biometric recognition," in *Proceedings of IEEE Conference on Image Processing* (Institute of Electrical and Electronics Engineers, New York, 2002), Vol. **1,** pp. 53–56.
6. M. Savvides, B. V. K. Vijaya Kumar, and P. Khosla, "Face verification using correlation filters," in *Proceedings of IEEE Automatic Identification Advanced Technologies* (Institute of Electrical and Electronics Engineers, New York, 2002), pp. 56–61.
7. B. V. K. Vijaya Kumar and J. Thornton, "Distortion-tolerant iris recognition using advanced correlation filters," in *Proceedings of IEEE Multi-Modal User Authentication Workshop* (Institute of Electrical and Electronics Engineers, New York, 2003), pp. 173–180.
8. B. V. K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," Appl. Opt. **31,** 4773–4801 (1992).
9. B. V. K. Vijaya Kumar and A. Mahalanobis, "Recent advances in composite correlation filter designs," Asian J. Phys. **8,** 407–420 (1999).
10. A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," Appl. Opt. **26,** 3633–3640 (1987).
11. B. V. K. Vijaya Kumar, "Minimum-variance synthetic discriminant functions," J. Opt. Soc. Am. A **3,** 1579–1584 (1986).
12. P. Réfrégier, "Filter design for optical pattern recognition: multicriteria optimization approach," Opt. Lett. **15,** 854–856 (1990).
13. P. Réfrégier, "Optimal trade-off filters for noise robustness, sharpness of the correlation peak, and Horner efficiency," Opt. Lett. **16,** 829–831 (1991).
14. B. V. K. Vijaya Kumar, D. Carlson, and A. Mahalanobis, "Optimal trade-off synthetic discriminant function (OTSDF) filters for arbitrary devices," Opt. Lett. **19,** 1556–1558 (1994).
15. I. Daubechies, *Ten Lectures on Wavelets* (Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1992).
16. R. Coifman, Y. Meyer, and V. Wickerhauser, "Wavelet analysis and signal processing," in *Wavelets and Their Applications*, M. B. Ruskai, G. Beylkin, R. Coifman, I. Daubechies, S. Mallat, Y. Meyer, and L. Raphael, eds. (Jones and Bartlett, Boston, Mass., 1992), pp. 153–178.
17. C. Watson, *NIST Special Database 24—Live-Scan Digital Video Fingerprint Database* (National Institutes of Health, Gaithersburg, Md., 1998), available at http://www.nist.gov/srd/nistsd24.htm.
18. E. Candès and D. Donoho, "Curvelets—a surprisingly effective non-adaptive representation for objects with edges," in *Curve and Surface Fitting*, A. Cohen, C. Rabut, and L. L. Schumaker, eds. (Vanderbilt U. Press, Nashville, Tenn., 1999), pp. 123–143.
19. E. Candès and D. Donoho, "Ridgelets and their derivatives: representation of images with edges," in *Curve and Surface Fitting*, A. Cohen, C. Rabut, and L. L. Schumaker, eds. (Vanderbilt U. Press, Nashville, Tenn., 1999), available at http://www.acm.caltech.edu/~emmanuel/papers/StMalo.pdf.
20. A. Mahalanobis, B. V. K. Vijaya Kumar, S. Sims, and J. Epperson, "Unconstrained correlation filters," Appl. Opt. **33,** 3751–3759 (1994).