Object recognition in subband transform-compressed images by use of correlation filters

Cindy Daniell, Abhijit Mahalanobis, and Rod Goodman

We introduce subband correlation filters (SCFs) as a solution to the problem of object recognition at multiple resolution levels in quantized transformed imagery. The approach synthesizes correlation filters that operate directly on subband coefficients rather than on image data. We explore two techniques to accomplish the reduced-resolution recognition: (1) training the correlation filters to incorporate downsampling tolerance and (2) adaptation of the subband decomposition filters to accommodate the reduced resolutions. For compression ratios of 20:1, SCFs demonstrate recognition performance of at least 90%, 85%, and 75%, respectively, on 2-, 4-, and 8-ft-resolution synthetic aperture radar data. © 2003 Optical Society of America

OCIS codes: 100.5010, 100.7410.

1. Introduction

Images take an enormous amount of memory; hence they are typically stored in a compressed format, yet many applications need quick searches of large image databases. When one works with compressed images, however, the objective is often to do a fast cursory search across the data, even if that entails a reduction in recognition accuracy. Performing a quick search of images in the quantized transform space suggests that we exploit the multiple resolutions available in a subband decomposition of image data.

We address two parallel questions: Can we perform object recognition at lower resolutions with reasonable accuracy? and How does the accuracy degrade with decreasing bit rate? The approach manifested in subband correlation filters (SCFs) is twofold: (1) We perform both the training and the operating phases of object recognition directly on the subband coefficients at multiple resolution levels, and (2) we combine the multiple subband responses with polynomial correlation filters (PCFs), a powerful architecture developed by Mahalanobis and Vijaya Kumar[1] for fusing multiple inputs into a single correlation output.

Toward this end, we pursue the intuitive idea of synthesizing a correlation filter from each of the individual subbands in a subband decomposition and then processing each subband with its corresponding filter to achieve a pattern match for object classification. We call this approach subband coefficient-domain processing. Aside from the issue of fusing the multiple outputs, a larger issue that affects this approach is the variation in the subband patterns owing to the downsampling properties of subband coders.

We combine two techniques to mitigate effectively the downsampling effects encountered when using subband coder architectures to perform reduced-resolution recognition. First, we train the correlation filters over multiple shifts of the training data to instantiate downsampling tolerance directly in the object-recognition domain. Second, we design the subband decomposition filters to reduce the downsampling effects present in the reduced resolutions of the compression domain. Together, these two procedures aid our approach to combining correlation filters with subband coders as a solution to the problem of object recognition at multiple resolution levels in quantized transformed imagery.

Our purpose in this paper is not to advance the field of correlation filtering or subband decomposition but to combine them for the purpose of object recognition at multiple resolutions of transformed data. Current systems typically perform the reconstruction

C. Daniell (daniell@hrl.com) is with HRL Laboratories, LLC, 3011 Malibu Canyon Road, RL69, Malibu, California 90265. A. Mahalanobis is with Lockheed Martin, 5600 Sandlake Road, MS 450, Orlando, Florida 32819. R. Goodman is with Cyrano Sciences, 73 North Vinedo Avenue, Pasadena, California 91107.

(i.e., the inverse of compression) and recognition tasks serially. They fully uncompress the stored images and then apply an object recognizer to the reconstructed data. In general, the reconstructed data are of a lower quality than the original owing to quantization during the compression step. Walls and Mahalanobis[2] used this method on synthetic aperture radar (SAR) imagery; Liu and Mitra[3] followed the technique using fingerprint data; Shin and Kil[4] employed the paradigm with sonar images; and visual images of tanks underwent the same approach by Miller.[5] In addition, multiresolution analysis does not benefit from current systems of working with compressed imagery. The Walls and Mahalanobis paper[2] provides us with baseline results for comparison with our own experiments.

As a brief initiation into subband coders and correlation filters, Section 2 provides a review of these research areas. Notation and terminology that we use throughout this paper are defined in Section 3. Section 4 introduces PCF theory and its function in the fusion of subband information. We present the concept of subband coefficient-domain processing in Section 5. In addition, this section discusses the downsampling issue inherent in coefficient-domain processing and suggests two methods to ameliorate the effect. Next, we examine the two methods proposed in Section 5: a training technique and a quadrature mirror filter (QMF) design technique, in Sections 6 and 7, respectively. Section 8 reports the performance of the SCF system along with the effects of decreasing bit rates. For future modifications, Section 9 proposes possible extensions to the SCF model presented here. Finally, Section 10 concludes the paper with a review of the SCF system: its performance, benefits, and limitations.

2. Brief Review of Correlation Filters and Subband Coders

To date, correlation filters and subband decomposition have been treated as separate disciplines. Thus we provide a little background on the history of both fields.

A. Correlation Filters

Some of the earliest research in object recognition grew out of the matched filter originating in the field of communications.[6,7] It is well known that a matched spatial filter is optimum for the detection of an object in a single image in the presence of additive white noise.[6] In such a system, a cross correlation is performed between the stored filter and the image in question. If the correlation exceeds a predetermined threshold, then the image is declared to contain the object of interest. The use of matched spatial filters grew throughout the 1960's and 1970's, with a good discussion and synopsis given in Andrews,[8] Duda and Hart,[9] and Pratt.[10] Unfortunately, wide variability in an object's characteristics makes the matched spatial filter a brittle method of object recognition. Matched spatial filters, however,

provided the basis for the evolution of the more sophisticated correlation filters.

Classical pattern-recognition techniques began with correlation-filtering models in the 1980's. Correlation filters are specially formulated generalizations of the matched spatial filter, which are able to recognize many different views of an object in the presence of signature variations and in clutter. They were born out of synthetic discriminant functions, which themselves evolved from linear combinations of matched spatial filters.[11] A good introduction to correlation filter theory is given in Vijaya Kumar et al.[12] and in Vijaya Kumar.[13] The two most important steps in the evolution of correlation filters came when Vijaya Kumar[14] introduced minimum-variance SDFs to maximize SDF noise tolerance and when Mahalanobis et al.[15,16] introduced minimum average correlation energy filters capable of producing sharp correlation peaks at the same location as the shifted input.

In general, correlation filters are composites of several training images representative of a particular object, thus mitigating the brittle characteristics of a matched spatial filter. Moreover, correlation filter theory suggests that the minimum average correlation height (MACH) filter is statistically optimum for detecting objects in additive noise, when Gaussian assumptions hold.[17] Correlation filters, however, are still not robust enough to perform recognition over wide variations of an object. Multiple correlation filters are formed for different aspects of an object.

SAR imagery forms a specific domain of object recognition. To date, successful object recognition in SAR imagery primarily involves template matching and correlation filters. Matched spatial filters (as known as template matching) were first applied to SAR imagery in the mid-1980's.[18,19] Currently, both Novak et al.[20–22] and Hostetler[23] perform cross correlation with template filters. These templates are formed from a mean image, over 5° spans of each object within each object class. Thus, for 360° viewing aspects of each object class, the template method requires 72 templates for each class.

Correlation filters, although similar to template matching, form more global descriptions of the objects under consideration. They were first used on SAR imagery in Mahalanobis et al.,[24–26] and recently good results have been shown by use of correlation filters over 45° spans of the object classes. This method requires only eight correlation filters per object class.

Most recently Mahalanobis and Vijaya Kumar have developed PCFs[1] that provide a vehicle for more powerful object recognition by fusion of correlation filter responses over multiple input sources. The PCF architecture permits object recognition with multisensor data and forms the backbone of the subject of this paper, SCFs. In a novel twist on PCFs, SCFs employ correlation over the multiple subbands in a subband decomposition.

B. Image Compression

The objective of image compression is to reduce the number of bits of information necessary to represent a given image by elimination of redundancy in the image or by introduction of distortion into the image in a manner that is acceptable to the viewer or, in our case, to the object recognizer. A good summary of image-compression techniques is given in Sayood.[27] Transform coding, developed by Andrews[8] and Andrews and Pratt[28,29] was the big step in compression of two–dimensional data. A current popular image-compression technique, the JPEG method,[30] combines transform and predictive coding. The main detriment of JPEG coding is that it produces blocking artifacts at low bit rates.

The most recent image-compression technique is subband coding. Detailed information about subband coding can be found in Vaidyanathan,[31] Vetterli and Kovacevic,[32] and also in Strang and Nguyen.[33] The technique, first introduced for speech coding in the 1970's, centers around splitting an image into multiple frequency bands and then coding the resulting coefficients. In that regard, it is similar to transform coding. Subband coding differs, however, in that it is not block based but rather applies the filtering operation to the entire image.

Subband coding is easily extended to two dimensions by analysis of the input signal independently in the horizontal and vertical directions. A typical level of a multiresolution image decomposition will have four subbands of information: one composed of only low-frequency components, the low–low (LL) subband; two containing both high- and low-frequency components, the high–low (HL) and low–high (LH) subbands; and one consisting of only high-frequency components, the high–high (HH) subband. Often, subband literature refers to the high-frequency subbands (HL, LH, and HH) collectively as the upper subbands.

Subband coders became more popular and viable with the advent of perfect reconstruction, made independently by Vaidyanathan[34,35] and Vetterli.[36] In the absence of quantization and ignoring machine round-off error, subband coders can thus provide lossless compression. Subband coders are most often used in lossy compression; however, in that case the compression rate relates directly to the degree of quantization. Compression applications typically use perfect reconstruction filters[34–37] in both lossless and lossy compression. A popular set of perfect reconstruction filters are the QMFs,[38,35,37] which we use in this research. Again, our purpose is not to improve the compression power and quality of QMFs but to optimize their effectiveness with an object-recognition model, the correlation filter.

3. Notation and Definitions

Images in the space domain are denoted in lower-case italics, and upper-case italics represent their counterpart in the frequency domain. Thus a two-dimensional image $x(m, n)$ has Fourier transform $X(k, l)$. Vectors are expressed by lower-case bold characters, and matrices are denoted by upper-case bold characters. Either $x(m, n)$ or $X(k, l)$ can be expressed as a column vector $\mathbf{x}$ by lexicographical scanning. We define lexicographical scanning to proceed columnwise, unless otherwise stated. The quantities $\mathbf{X^T}$, $\mathbf{X^*}$, $\mathbf{X^t}$, and $\mathbf{X^{-1}}$ signify, respectively, the transpose, conjugate, transpose conjugate, and inverse of $\mathbf{X}$. Correlation and convolution are designated by the symbols $\otimes$ and $*$, respectively. The quantity $\hat{x}(n)$ refers to a reconstructed version of the original signal $x(n)$.

To bridge the two disparate fields brought together in this paper, we provide some nuances of the terminology used throughout the paper. Note that the definitions used in this paper do not necessarily hold universally throughout the literature.

- The term correlation filter is synonymous with object-recognition filter and MACH filter.
- The subband coder employed in this paper uses a QMF filter, which we use interchangeably with the term subband decomposition filter.
- When we reconstruct the image with a subband coder, we use synthesis filters; therefore the terms reconstructed signal and synthesized signal represent the same thing. Yet, when we create the MACH filter, we synthesize it from its training input.
- A subband coder forms the heart of a compression system. When no ambiguity results, the term coding is at times used to mean compressing; sometimes, however, we also encode symbols with their representative code words.
- Compression does not take place until the coefficients in the subband coder are quantized to total fewer than the number of bits in the original image (typically, 8 bits per pixel). At the typical digitization of 8 bits per coefficient, the subband coder is simply another representation of the original image.
- A multiresolution decomposition, pyramid decomposition, subband decomposition, and hierarchical decomposition are all the same entity, for the purposes of this paper.
- Analysis filters form the forward compression transform, whereas the inverse compression transform uses synthesis filters (also known as interpolation filters).
- The subband literature denotes the first level of decomposition as the highest level because it contains the highest-resolution data. The last level of decomposition is known as the lowest level.
- The high-frequency subbands in a subband decomposition are known collectively as the upper subbands.

4. Subband Fusion with Polynomial Correlation Filters

Although many of the fusion techniques residing in the information-processing literature may be applicable, we selected PCFs because they are a direct extension of standard correlation filters. A detailed description of PCFs is given in Mahalanobis and Vijaya Kumar,[1] in which the use of PCFs with mul-
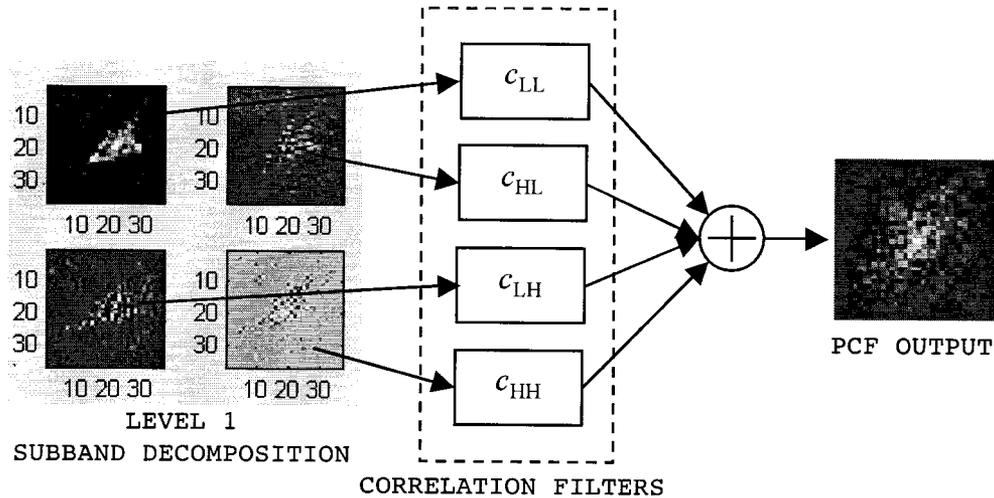
Fig. 1. SCF architecture. The PCF provides a way to simultaneously correlate all subbands. Each subband takes its own input channel and thus its own PCF. We use an analogous structure for each level of the subband decomposition.

tiresolution data was first suggested. PCFs embody a multiple-input–single-output system. Given multiple inputs, a PCF system calls for a like number of correlation filters, each corresponding to one input. Yet, the system produces a single correlation surface output.

In our application the number of inputs is equal to the number of subbands in one level of a pyramid decomposition. The architecture diagrammed in Fig. 1 illustrates our use of PCFs. In future research, however, a system may integrate any subset of a subband decomposition, as discussed in Section 9.

Mathematically, the output of a general PCF is expressed as

$$y(m, n) = \sum_{p=1}^{P} c_p(m, n) \otimes x_p(m, n), \qquad (1)$$

where P is the number of input channels, $x_p(m, n)$ is the data present at the input of each channel, and $c_p(m, n)$ is the filter function for each channel. In our case, each subband forms one channel. These filters are formulated to jointly maximize the peak value of $y(m, n)$ while minimizing the effects of distortions and noise. Equation (1) gives the appearance that PCF theory augments standard correlation by only a simple summation. The real power behind PCF theory, however, lies in the simultaneous optimization of the filters for multiple channels during training.

Typically, correlation processing is done in the frequency domain for the purpose of computational efficiency. In that case, the correlation output is given as

$$y(m, n) = \Im^{-1}\left\{\sum_{p=1}^{P} C_p{}^*(k, l) X_p(k, l)\right\}, \qquad (2)$$

where $C_p(k, l) = \Im\{c_p(m, n)\}$ and $X_p(k, l) = \Im\{x_p(m, n)\}$. $\Im\{\cdot\}$ and $\Im^{-1}\{\cdot\}$ represent the forward and in-

verse discrete Fourier-transform operations, respectively. Within the PCF architecture, we employ the formulation from the current state of the art, the MACH filter.[17]

For a set of N training images, the average training image for the pth channel is defined by

$$M_p(k, l) = \frac{1}{N} \sum_{i=1}^{N} X_p{}^i(k, l), \qquad (3)$$

where $X_p{}^i(k, l)$ represents the ith training image of the pth input channel. We define the cross-power spectrum of the training set for the pth and qth input channels as

$$D_{pq}(k, l) = \frac{1}{N} \sum_{i=1}^{N} X_p{}^i(k, l) X_q{}^{i*}(k, l) \qquad (4)$$

and a spectral variance term as

$$S_{pq}(k, l) = D_{pq}(k, l) - M_p(k, l) M_q{}^*(k, l). \qquad (5)$$

Using these quantities, we define the cross-channel spectral statistics as

$$B_{pq}(k, l) = \alpha S_{pq}(k, l) + \beta D_{pq}(k, l) + \gamma \mu_{pq}, \qquad (6)$$

where the term μ_{pq} is simply the mean value of the spectral variance term, S_{pq}. The constants α, β, and γ are used for optimally trading between distortion tolerance, noise tolerance, and correlation energy minimization. The interested reader is referred to Carlson[39] for further details about the use of these parameters in the design of MACH filters.

Central to the strength of PCF theory is the fact that each filter is a conglomerate influenced by the spectral content of all the input channels. Specifically, each filter is influenced by the power spectral density of its own inputs as well as the cross-spectral density between its input channel and all the others. The term $B_{pq}(k, l)$ of Eq. (6) contains the collective spectral terms used in the derivation of the MACH

filter (see Mahalanobis and Vijaya Kumar[17]), with modifications made to reflect multiple input channels. Consequently, the PCF solution analytically optimizes the same performance criterion as the MACH filter but with respect to the cross-channel statistics.

Using matrix–vector notation, we denote the cross-spectral statistics in Eq. (6) by a block-diagonal matrix $\mathbf{B}_{pq}$. Further, let $\mathbf{m}_p$ be the mean image of the pth channel also expressed as a vector. The two-channel PCF solution is then given by

$$
\begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{pmatrix} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}^{-1} \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{pmatrix}, \tag{7}
$$

where $\mathbf{c}_p$ is the optimum correlation filter vector for the pth input channel. Thus the PCF amalgamates information from all input channels to obtain the best set of correlation filters that simultaneously process all inputs and combine them into a single result. Essential to the filter synthesis is that the system estimates the $\mathbf{m}_p$ and $\mathbf{B}_{pq}$ terms over a large number of inputs representative of the objects under study (i.e., a training set). Certainly, Eq. (7) is easily extendable to N input channels. Detailed information regarding the solution to the inverse of the generalized form of the matrix in Eq. (7) is provided in Mahalanobis and Vijaya Kumar.[1]

5. Downsampling Effects in Coefficient-Domain Processing

Subband coders downsample by 2 between levels of decomposition. This operation produces different subband responses to single-pixel-shifted copies of the original input image. Herein lies the crux of the issue. For example, assume an object at some reference point in the original image has produced a one-level subband decomposition. (We call this the reference decomposition.) Shifting the original image by one pixel in any direction will result in a similar yet distinctly different pattern in a second subband decomposition. This is because, after the low-pass and high-pass filters are applied to the original image, only every other pixel remains in the image representation at level one. The issue is more of a problem for the upper subbands because of the high-frequency nature of the data; however, it is still present in the subband containing only low-frequency components (also known as the LL subband). We illustrate the issue for the LL subband in Fig. 2.

Now from the reference point, a two-pixel shift of the original image in any direction will produce the exact same decomposition pattern (ignoring boundary effects) at level one as the reference decomposition. However, the output will be translated by one pixel. This type of displacement is not an issue for correlation filters. Correlation filters are well known for their shift-invariant properties[12]; thus they are well suited for recognizing translations of the pattern they are trying to match. They are not,

however, downsampling invariant, nor are they very downsampling tolerant.

Furthermore, the downsampling issues are present and multiply on every level of the decomposition. Although the first level has only four possible patterns, the second level will produce four variations of each of them. Recursive downsampling leads to a tetradic tree (i.e., a tree branching fourfold) of pattern variations. Moreover, when processed with identical correlation filters, similar yet distinctly different patterns produce dissimilar correlation surfaces.

To mitigate the downsampling effects of a subband coder, we propose two complementary solutions. The first requires synthesizing the correlation filters over multiple shifts of the training data. The second modifies the subband decomposition's QMF properties to account for correlation between neighboring pixels in the training data. We designed the two solutions to work in concert with each other; the first was developed with the high-pass information in mind, whereas the second was constructed primarily for the low-pass information. SCFs may use these two solutions separately or together to substantiate downsampling tolerance.

6. Incorporating Downsampling Tolerance through Training

As elucidated in Section 5, single-pixel variations in the placement of the original image produce differing subband images (excluding edge effects). The algorithm described here frames the following question: Is it possible for a correlation filter to accommodate these variations by training over multiple shifts of each training image?

In reference to a one-level decomposition, the training algorithm we use incorporates each input image four times: once in its original state, once after the image undergoes a single-pixel shift up, again after we translate the image by one pixel to the left, and yet again after we shift the image up and to the left one pixel in each direction. We utilize a circular shift technique so that the first column of $x(m, n)$ becomes the last column of the shifted image. Figure 3 depicts such a procedure.

We then employ this procedure at each level. That is, at level two, the LL subband from level one acts as the input image; it is shifted the requisite number of times, and the four subband images on level two make up the PCF channel inputs. All experiments in this paper employ the training tree structure of Fig. 3.

The above procedure is but a subset of the set of shifts necessary to cover all the downsampling variations. To cover all possibilities, level two would have to incorporate four shifts for each of the four shifts in level one, and so on at each level, spawning a proliferating tetradic tree of training images. Such a massive training set is a large and unreasonable burden.
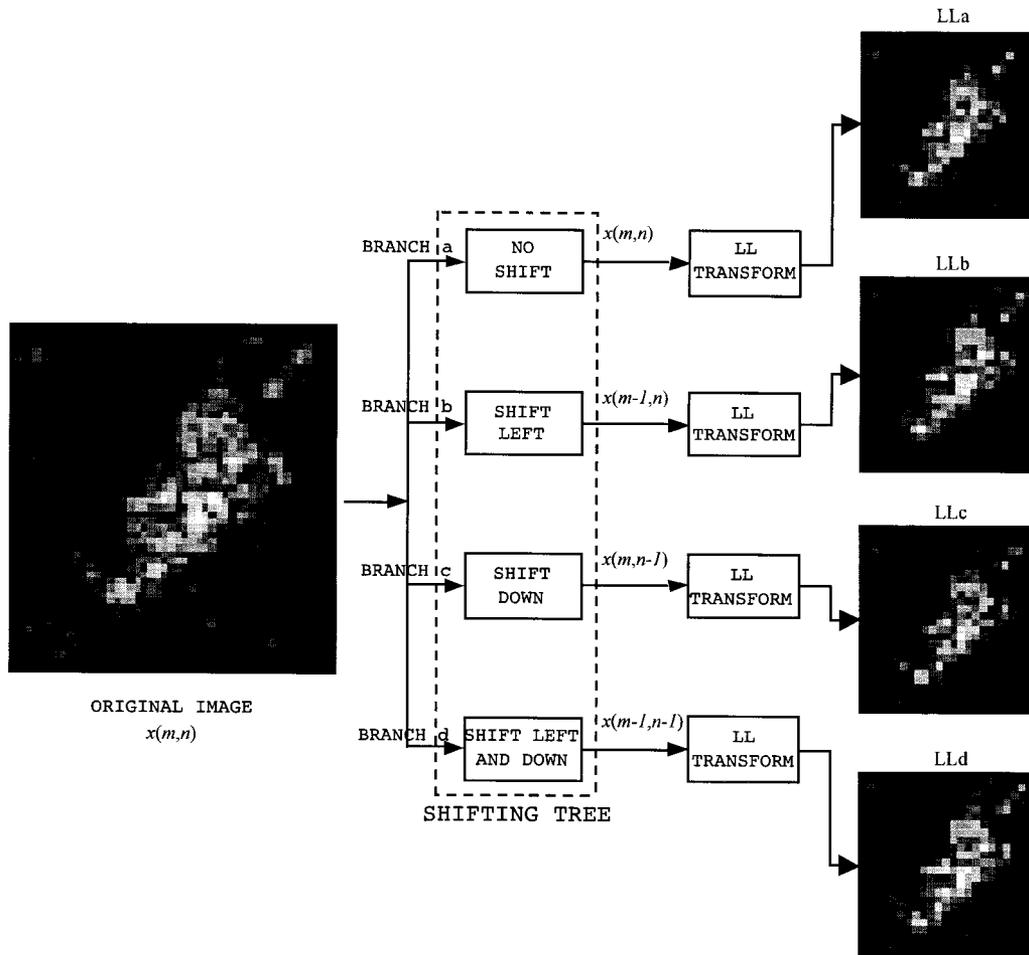
Fig. 2. Downsampling effect on the LL subband patterns at level one. The subband downsampling process is not shift invariant. We illustrate the issue for the LL subband only, although the upper-band coefficients are most affected. Note that the four LL images shown are not simply shifted versions of one another. Rather, they result from shifts of the original image. On close inspection, one can see that the LL subband of each branch exhibits a different pattern. Furthermore, the problem explodes at the lower decomposition levels, as each shifted version on level n spawns four shift variations on level $n - 1$.

7. Incorporating Downsampling Tolerance through Quadrature Mirror Filter Design

To avoid an exploding tree structure for training, we modify the QMF bank of the subband coder as a second and complementary method of instilling downsampling tolerance in the system. If the LL subband response is sufficiently robust, the tree used for shift training becomes linear, as discussed in Section 6. This idea gives rise to the following question: Can we tune the QMF used for image compression to compensate for the downsampling within the subband coder?

A. Our Baseline Qualified Mirror Filter Design

A detailed description of our approach to QMF design is given in earlier publications.[40,41] It evolved from the research of Walls and Mahalanobis,[2] which introduced a matrix representation of subband decomposition based on the Hadamard transform. Central to the design is the development of a matrix representation, $\mathbf{T}$, which combines the filtering and down-

sampling of the subband analysis filter, $h_0(n)$, over multiple levels in a matrix structure. Its rows are shifted versions of $h_0(n)$, appropriately padded with zeros to implement the required filtering and decimation. The matrix construct, $\mathbf{T}$, collapses the complete subband hierarchy into an aggregate structure that provides a direct channel from input to output and vice versa. A complete mathematical definition of $\mathbf{T}$ is given in an earlier publication.[40]

More specifically, our approach to designing a perfect reconstruction QMF bank is to optimize the impulse response of $h_0(n)$ until the error terms in Eq. (8) are minimized. A sum of three error terms derived from our design constraints[40–41] form a function, f, which we optimize numerically, via the MATLAB routine fminu. This approach,

$$f = \sum_{i \neq j} \left[(\mathbf{T}^\mathbf{T}\mathbf{T})_{ij} \right]^2 + (\mathbf{h}_0{}^\mathbf{T}\mathbf{h}_0 - 1)^2 + \left[\sum_{n=1}^{N_f} h_0(n) - \sqrt{2} \right]^2,$$
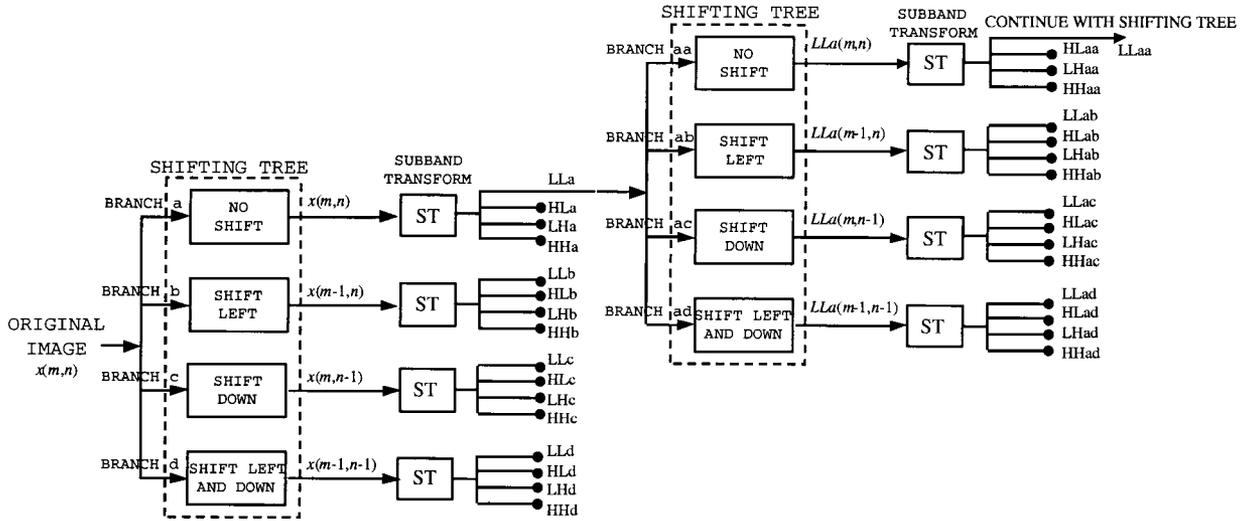(8)

Fig. 3. Partial tetradic training tree for training over image shifts. Only the LL subband of the original placement spawns four shifting branches. Thus there are $4N$ training images per subband of each level, given N images in the original training set. This is the training method we employ in this paper.

is a generalized form of other time-domain techniques reported by Jain and Crochiere,[42] and Hemami and Gray,[43] and Hemami.[44] As in most QMF design strategies, we design only $h_0(n)$ and derive the remaining three filters from this single filter. The terms $h_0(n)$ and $h_1(n)$ are the low-pass and high-pass analysis filters, respectively:

$$h_1(n) = (-1)^n h_0(n), \qquad g_0(n) = h_0(n),$$
$$g_1(n) = -h_1(n). \qquad (9)$$

Likewise, $g_0(n)$ and $g_1(n)$ are low-pass and high-pass synthesis filters, respectively.

By its nature, the QMF compacts the energy in the upper subbands, forming a sparse representation. Coupled with downsampling between levels, this sparseness impairs object recognition, as detailed in Section 5. In contrast, if one could stabilize the QMF response between adjacent coefficients, then the system could better tolerate the downsampling effects. With this objective in mind, we add another term, the downsampling tolerance criterion, to the standard QMF design.

B. Equivocating the Quadrature Mirror Filter Response to Input Shifts

Physically, the shift-sensitivity term refers to the difference of the QMF response to an image and its single-pixel diagonally shifted copy. The two responses should be as similar as possible to compensate for the downsampling of the subband coder in both directions (horizontal and vertical).

The autocorrelation matrix of the difference image, d, is given in Eq. (11):

$$d(m, n) = x(m, n) - x(m - 1, n - 1) \forall (m, n), \quad (10)$$

$$\mathbf{R}_d = E\{\mathbf{dd}^T\}. \qquad (11)$$

We define the shift-sensitivity measure in the LL subband as a filtering operation. Let $h_0(m, n)$ and $h_1(m, n)$ (or $\mathbf{h}_0$ and $\mathbf{h}_1$ in vector notation) represent the QMF filters used for the low-pass and high-pass analysis operations, respectively. Because the LL subband requires a low-pass filter in both directions, we define the shift-sensitivity measure with that in mind. In effect, the shift-sensitivity term, Φ, is a scalar measure of the difference between two low-pass QMF responses: (1) that elicited from an original image and (2) that resulting from a single-pixel diagonally shifted version of the image:

$$\Phi = \mathbf{h}_0^T \mathbf{R}_d \mathbf{h}_0. \qquad (12)$$

In addition, we are interested in measuring a unidirectional shift sensitivity in the upper subbands. Two of the upper subbands (LH and HL) require low-pass filtering in one direction and high-pass filtering in the other. Hence the equation for Φ given below incorporates the effective response for the HL and LH subbands along with the LL response:

$$\Phi = \mathbf{h}_0^T \mathbf{R}_d (\mathbf{h}_0 + \mathbf{h}_1). \qquad (13)$$

Although we attempt to constrain the shifted response of the QMF along only the low-frequency paths of the subband decomposition, the resulting QMF still loses some of its energy compaction. Therefore we balance the shift-sensitivity term with an energy-compaction term. Together, they form the downsampling tolerance constraint.

C. Energy Compaction

Variance of a subband distribution is one measure of its energy compaction or compression capability. Smaller variance corresponds to more compression. We provide a similar measure for our energy-compaction term. We define the energy-compaction measure as a derivative of only the HH upper sub-

band by performing a high-pass filtering operation in both directions on the autocorrelation matrix, $\mathbf{R}_x$, of a training image, $\mathbf{x}$:

$$\Theta = \mathbf{h}_1{}^{\mathbf{T}}\mathbf{R}_x\mathbf{h}_1, \qquad (14)$$

where

$$\mathbf{R}_x = E\{\mathbf{x}\mathbf{x}^{\mathbf{T}}\}.$$

Thus the energy-compaction term, Θ, is a scalar measure of the variance in the HH subband resulting from the training image. Its purpose is to counterbalance the shift-sensitivity term and ensure that the QMF bank maintains powerful compression performance.

D. Shift-Tolerant Quadrature Mirror Filter Solution

Recall the baseline QMF design equation (8). To achieve a shift-tolerant solution, we add a fourth term to this equation, which we call the downsampling tolerance criterion, ϵ_{ds}:

$$\epsilon_{ds} = \lambda\Phi + (1-\lambda)\Theta. \qquad (15)$$

Together, the error-shift term, Φ, and the energy-compaction term Θ form the downsampling tolerance constraint, where λ is a parameter meant to balance the two opposing functions.

Thus the full-fledged optimization equation for a shift-tolerant QMF design follows as

$$f = \sum_{i \neq j} [(\mathbf{T}\mathbf{T}^{\mathbf{T}})_{ij}]^2 + (\mathbf{h}_0{}^{\mathbf{T}}\mathbf{h}_0 - 1)^2 + \left[\sum_{n=1}^{N_f} h_0(n) - \sqrt{2}\right]^2$$
$$+ \ 0.2[\lambda\mathbf{h}_0{}^{\mathbf{T}}\mathbf{R}_d(\mathbf{h}_0 + \mathbf{h}_1) + (1-\lambda)\mathbf{h}_1{}^{\mathbf{T}}\mathbf{R}_x\mathbf{h}_1], \qquad (16)$$

where $h_1(n)$ is given by Eqs. (9) and autocorrelation matrices $\mathbf{R}_d$ and $\mathbf{R}_x$ are defined in Eqs. (11) and (14). We scale the downsampling tolerance constraint by 0.2 so as not to overshadow the important QMF properties.

As in the design of the standard QMF, we use numerical optimization to find the coefficients of $h_0(n)$ that minimize the function f. Note that the $\mathbf{R}$ matrices are computed *a priori* and remain unchanged during the iterative optimization; λ is also a constant parameter. The $\mathbf{T}$ matrix and high-pass filter $h_1(n)$, however, derive from the low-pass filter $h_0(n)$ and thus are computed at every iteration. The image we use to compute the $\mathbf{R}$ matrices is one of the preprocessed training images, which is exemplary of the image statistics of the entire data set.

An earlier publication[40] examines the effects of λ in detail. We select a value of $\lambda = 0.7$ because it provides significantly more shift tolerance than the baseline QMF without a large drop in either compression power or reconstruction performance. Table 1 reports the performance metrics achieved by the baseline QMF and our optimized shift-tolerant QMF. For the remainder of this paper, all results reported for the optimized QMF use $\lambda = 0.7$ in Eq. (15). Figure 4 diagrams the coefficients of $h_0(n)$ for the shift-tolerant optimized QMF with the baseline QMF

Table 1. Performance Metrics of the Baseline and Shift-Tolerant QMF[a]

Measure	Baseline QMF	Shift-Tolerant QMF
PSNR	105 dB	67 dB
MSE	2×10^{-6}	1×10^{-2}
Shift-sensitivity metric	317	239
Compression metric	4.9	2.4

[a]PSNR and MSE are reported from a three-level decomposition of one of the preprocessed BMP training image. The filter has better performance when the shift-tolerant metric is smaller and when the compression metric is larger.

coefficients shown for comparison. The shift-tolerant QMF deviates dramatically from the baseline filter.

We show in Fig. 5 that the low-pass frequency response of the shift-tolerant QMF is similar to a bandstop filter. That is, it tries to attenuate the frequencies that most affect the shifted response. These frequencies are probably related to the size of the objects under investigation. Very small objects elicit an attenuation at higher frequencies. In addition to this frequency selectiveness, the shift-tolerant QMF unexpectedly magnifies some of the higher frequencies. For comparison, we display the low-pass frequency response of our baseline QMF in the same figure.

8. Performance

Finally, we may report on the performance of the SCFs. The first set of experiments examines the SCFs' ability to overcome the downsampling effects of the subband coder. We assess the two proposed solutions (incorporating downsampling tolerance through training and designing a downsampling tolerant QMF) both separately and together. We perform this assessment at only level one of the subband decomposition over unquantized image data. This
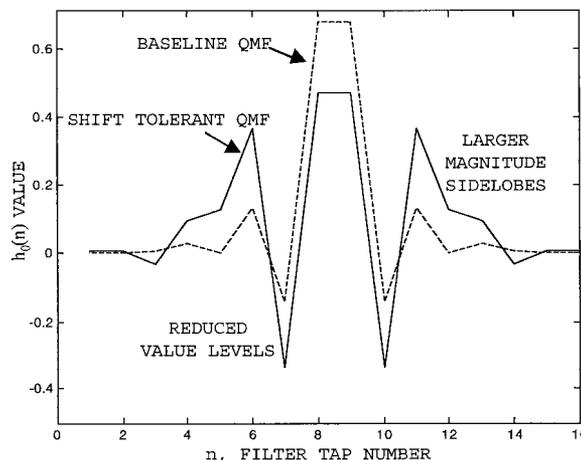

Fig. 4. Shift-tolerant QMF time-domain response. The shift-tolerant QMF values are quite different from our baseline QMF. The three values at its center are reduced, whereas the remaining sidelobes increase in magnitude.
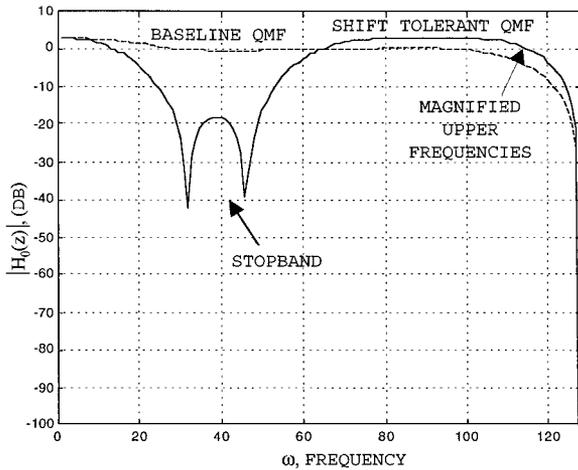
Fig. 5. Shift-tolerant QMF frequency response. The shift-tolerant QMF (solid curve) forms a stop band, as it tries to attenuate the frequencies that most affect the shifted response. The baseline QMF frequency is shown as a dashed curve for comparison.

set of data allows us to select one method for SCF algorithm definition.

Using the algorithm selected in the first set of experiments, we then evaluate the multiresolution performance of SCFs in a second set of experiments. In this case, we investigate system results at each level of the subband decomposition. Additionally, we characterize the performance over increased quantization (i.e., decreasing bit rate).

1. Experiment Conventions. Both sets of experiments observe the following conventions.

• A MACH recognition filter was built during training for each of three object classes: a (BMP) Bradley fighting vehicle, a BTR (Bronyetransporter) armored personnel carrier, and a T72 tank.
• Object classification is determined by one's thresholding the peak-to-sidelobe ratio (PSR) of the correlation surface.[45]
• Objects that produce subthreshold PSR values are classified as rejections.
• Training is done with the original (i.e., unreconstructed and unquantized) image data.
• Images are assumed to be preprocessed[46] prior to any compression.

2. Image Data. We selected the public Moving and Stationary Target Recognition (SAR) 1-ft-resolution image database (1 ft = 0.3048 m).

• The SAR imagery has 2-, 4-, and 8-ft resolutions at a one-, two-, and three-level subband decompositions, respectively.
• All the images were cropped to size 64×64 pixels, which contained the entire vehicle.
• The training set held 30 BMP, 29 BTR, and 28 T72 images at $0-45°$ views and $17°$ elevation.

• The test set held 24 BMP, 28 BTR, and 27 T72 images at $0-45°$ views and $15°$ elevation.

3. Parameter Settings. We use the following parameter settings for all experiments in this paper.

• PCF design: $\alpha = 0.2$, $\beta = 0.8$, and $\gamma = 0.05$ [Eq. (6)].
• Baseline QMF design: filter length $N_f = 16$ [Eq. (8)].[40]
• Shift-tolerant optimized QMF design: $N_f = 16$ and $\lambda = 0.7$ [Eq. (16)].
• PSR thresholds: 6, 5, and 4 at levels one, two, and three, respectively, in the decomposition.

A. Downsampling Tolerance Experiments

We have proposed two possible solutions to the downsampling effects inherent in the coefficient-domain processing of SCFs. First, the filters may incorporate downsampling tolerance through training. Second, the system may benefit from a shift-tolerant QMF design. We examine their effects with the following question: Do the proposed solutions, either separately or together, provide enough downsampling tolerance for reasonable system performance? To that end, we test four variations of the SCF algorithm. The models tested are

1. SCFs that use the baseline QMF and traditional training techniques.
2. SCFs that incorporate the new training method of training over shifts.
3. A system that decomposes the data with a shift-tolerant QMF.
4. A system that employs both the new training methodology and a shift-tolerant QMF.

As this is a proof-of-principle experiment, we test with unquantized image data and analyze results at only the first level of the subband decomposition.

Now, to evaluate a system's tolerance to downsampling, we must effectively process a set of test data twice. That is, we must test with the images in one position, then shift the data diagonally by one pixel and reprocess the data, and finally compare the results. Consequently, we report two sets of results for each of the four SCF models tested. Furthermore, to remove another operating variable, we test on the training set of images. This is because the training images have an established reference point for the image translation. That is, we know the system should perform optimally to the training set in its original position. We do not know what the reference position is with the test images.

1. System Performance
A summary of the system performance metrics is shown in tabular data. Table 2 reports the recognition measures, and Table 3 provides the reconstruction metrics. Note that downsampling tolerance is most exemplified by stability across the two data sets.

Overall, the system does fairly well on the data set,

Table 2. Recognition Metrics for the Four SCF Systems[a]

SCF Model Tested	Reference Test Data			Shifted Test Data		
	Pc (%)	Pe (%)	Pr (%)	Pc (%)	Pe (%)	Pr (%)
Baseline QMF–traditional training	97.7	2.3	0	95.4	4.6	0
Baseline QMF–shift training	98.9	0	1.1	97.7	1.15	1.15
Shift-tolerant QMF–traditional training	96.6	3.4	0	94.3	3.4	2.3
Shift-tolerant QMF–shift training	98.9	0	1.1	98.9	0	1.1

[a]We show the probability of correct classification, error, and rejection for both the original training data and the same set after a diagonal shift of one pixel has been applied to all the images. Equivalent results between the two data sets indicate good downsampling tolerance.

which we expected because we are testing on the training set. It is unclear why the results on the unshifted data (i.e., the exact versions we trained on) were not 100% correct when the baseline QMF was used. Although the decomposition contains enough information to reconstruct the original image, it is not the same information as in the original image. Perhaps the information at a one-level decomposition is not linearly separable. In comparison, a quick test of the training data in the uncompressed domain renders 100% correct classification. The baseline QMF does, however, give reasonable reconstruction performance with a peak signal-to-noise ratio (PSNR) of 105.

2. Discussion

The shift-tolerant QMF–shift-training system appears to be the most stable and thus the most tolerant to the downsampling effects of the subband coder. It also provides the best recognition results. The baseline QMF–shift-training system is also fairly stable, providing near the optimum recognition results, but in addition achieves a 57% improvement in the PSNR reconstruction metric.

From this small experiment, it appears that incorporating shifted versions of the input images in the training set has the most impact on downsampling tolerance. Incremental improvements are made by optimization of the QMF for shift tolerance but at a larger cost in image-reconstruction performance. We recommend that the shift training be used alone or in conjunction with the shift-tolerant QMF but note that the shift-tolerant QMF is not effective when used by itself. The baseline QMF employed in this paper exhibits reasonable downsampling tolerance.

B. Experiments Detailing Multiresolution Analysis and Quantization Effects

In this set of experiments we assess the performance of SCFs at all resolutions in the subband decomposition. In this case we are no longer concerned with downsampling tolerance but with performance over the test set of images and the implications of declining bit rate. This set of experiments was designed with the following two questions in mind.

- Which of the SCF systems outlined at the beginning of Section 8 proves capable at lower levels of resolution?
- How will compression degrade recognition performance at each level of resolution?

Corresponding with the architecture of SCFs outlined in Section 4, we decompose the image only to the level required for each experiment below. Thus the quantization tables will be different for the different levels. That is, quantization of a level-two decomposition has to quantize the data effectively at both levels one and two. A level-one decomposition, however, quantizes the information only at level one.

1. System Selection

We considered both the shift-tolerant QMF–shift-training system and the baseline QMF–shift-training system for the multiresolution experiments. We performed an initial review with unquantized test data, with results reported in Table 4.

Observe that the shift-tolerant QMF does not duplicate the success of the baseline QMF system on levels two and three. This is probably due to the

Table 3. Reconstruction Metrics for the Four SCF Systems[a]

SCF Model Tested	PSNR (dB)	MSE
Baseline QMF–traditional training	105	2×10^{-6}
Baseline QMF–shift training	105	2×10^{-6}
Shift-tolerant QMF–traditional training	67	1×10^{-2}
Shift-tolerant QMF–shift training	67	1×10^{-2}

[a]Reconstruction measures are due only to the QMF of the system. Thus systems with the same QMF achieve the same performance. Both PSNR and MSE are given from the reconstruction of one of the BMP test images.

Table 4. Multiresolution Performance of Baseline and Shift-Tolerant QMF[a]

Measure	Baseline QMF–Shift Training	Shift-Tolerant QMF–Shift Training
Level one Pc	96%	93%
Level two Pc	85%	64%
Level three Pc	75%	54%
PSNR	105 dB	67 dB
MSE	2×10^{-6}	1×10^{-2}

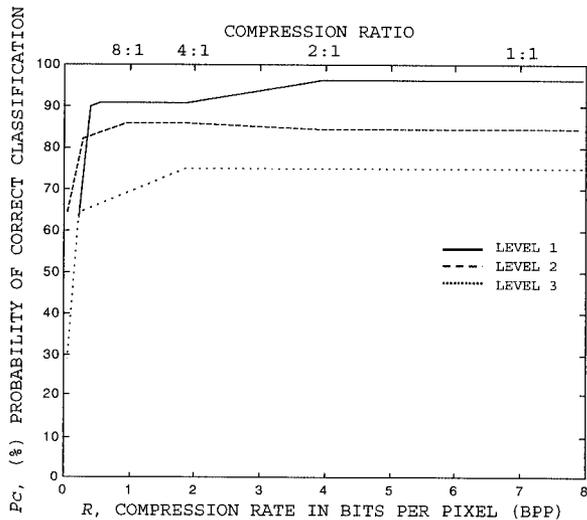[a]One of the BMP test images provides the PSNR and MSE values given above.

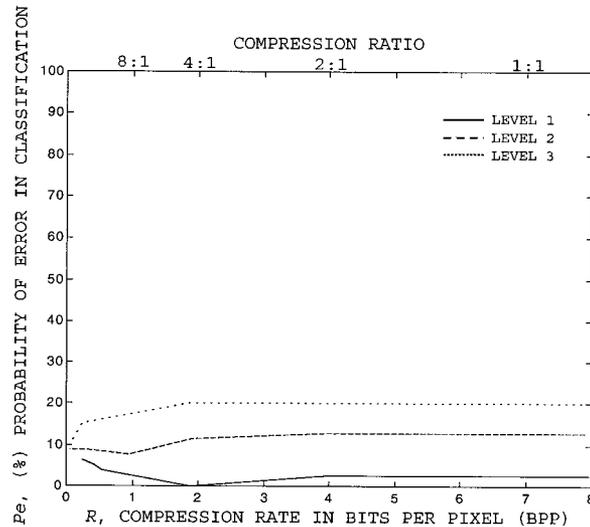Fig. 6. Probability of correct classification versus bit rate.

Fig. 7. Probability of error in classification versus bit rate.

lower PSNR and increased mean squared error (MSE) of the shift-tolerant QMF. PSNR is important not only for visualizing reconstructed data but also for performing optimal representations at the lower levels within a subband decomposition. Therefore we selected the baseline QMF–shift-training system for the remaining experiments. A smaller choice of λ [Eq. (16)] may provide a better-performing shift-tolerant QMF.

Also, note that, for both systems, results decline as we go further into the decomposition. This is an expected consequence of multiresolution. As resolution is decreased, patterns become more ambiguous and thus harder to discriminate.

2. Results

The following data characterize the performance of SCFs over increased quantization (i.e., decreasing bit rate). To compress the data, we employ a simple compression scheme: We use a uniform step size to scalar quantize the subband coefficients and then assume the data will be transmitted by an entropy coder with variable length codes and a finite-length codebook. Figures 6, 7, and 8 illustrate the object-recognition measures Pc, Pe, and Pr, respectively, for a SCF system at each of the three levels of decomposition as a function of bit rate. We also display the corresponding reconstruction metrics PSNR and MSE as we vary the bit rate in Figs. 9 and 10, respectively.

Observe that all three of the recognition-performance curves (Figs. 6–8) possess characteristics similar to that of a rate-distortion curve. That is, they maintain good performance down to very low bit rates and then drop off quickly. The Pe curves display some deviations, probably owing to the small sample size. We see a slight drop off in performance (Pc curves) at approximately 2 bits per pixel (bpp) for the first level and around 0.3 bpp on the second and third levels. The significant drop off occurs at 0.25,

0.0625, and 0.05 bpp on the first, second, and third levels, respectively.

Recall that the purpose of performing object recognition at lower resolutions is simply to execute a fast, cursory search for the object. Thus high recognition accuracy, while desirable, is not critical. A more detailed search at full resolution can be done later on the areas that the low-resolution search deems interesting. At lower levels of resolution, it is more important to have a low rejection rate. Even an error in classification can flag the application to search more thoroughly at a higher resolution of the data. The SCFs exhibit this trait down to compression ratios of at least 4:1 as shown in Fig. 8.

3. Discussion

Several observations are in order. First, we note that these results are slightly worse than the Walls and Mahalanobis results[2] that report good performance results up to compression ratios of 200:1. That
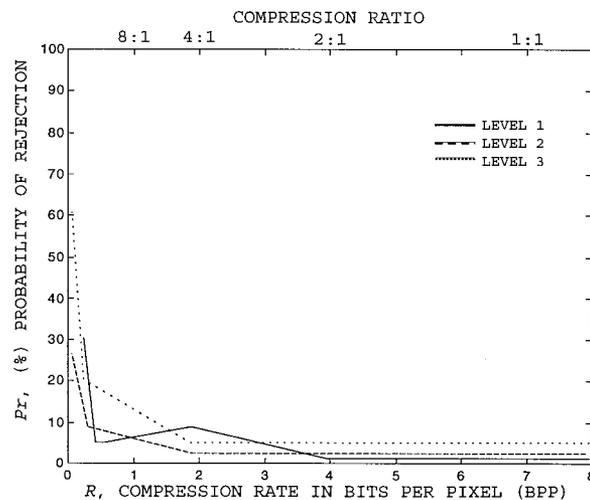

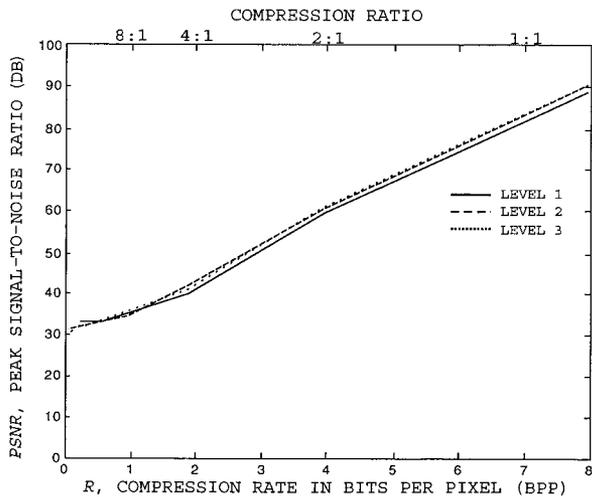Fig. 8. Probability of rejection versus bit rate.

Fig. 9.   PSN ratio versus bit rate.

Table 5.   Subbands Dropped at Low Bit Rates

Approximate Rate (bpp)	Subbands Dropped	
1	On level one:	HL, LH, HH
	On level two:	HH
	On level three:	HH
0.5	On level one:	HL, LH, HH
	On level two:	HH
	On level three:	HH
0.25	On level one:	HL, LH, HH
	On level two:	HL, LH, HH
	On level three:	LH, HH
0.125	On level one:	HL, LH, HH
	On level two:	HL, LH, HH
	On level three:	HL, LH, HH

previous study, however, is much different than that presented in this paper. The authors fully reconstructed the image from a quantized subband decomposition and performed recognition on the full-resolution (albeit reconstructed) image.[2] Moreover, they employed a sophisticated coding technique, the embedded zerotree wavelet coder,[46] whereas our results use the simplistic coding method outlined above. The embedded zerotree wavelet is well known to produce some of the best compression rates currently available for image data.

Second, we observe that SCFs at levels two and three maintain their initial performance at lower bit rates than level-one SCFs. This is primarily due to the fact that the higher levels require more bits to encode each subband, causing subband dropping to occur earlier. At the bit rate of 0.5 bpp, the quantizer does not have enough bits to code all the subbands effectively on level one and encodes the HL, LH, and HH subbands with zero bits (effectively dropping these subbands). The LL subband is then
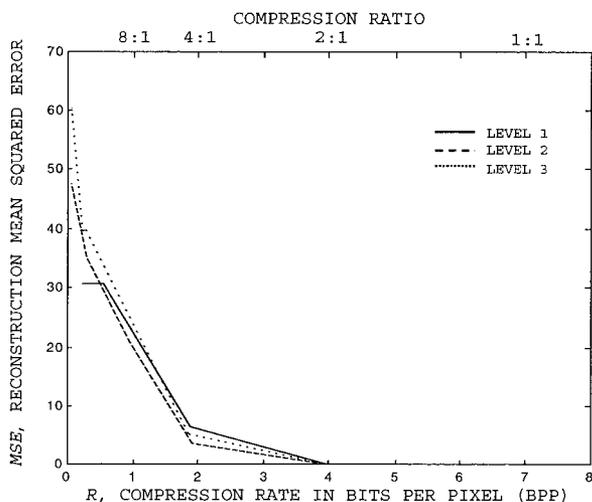
encoded at 2 bpp, and we maintain performance until 0.25 bpp, when the quantization of the LL subband becomes too great. Table 5 displays the bit rates that require an empty subband for each level of decomposition. We see from Table 5 that levels two and three retain more subbands at lower bit rates.

This is expected, as their smaller size requires fewer bits for full encoding. Finally, we note that at low bit rates the output correlation surfaces of the SCFs contain no aliasing components. Even when the bit allocation must zero out certain subbands, the PCF architecture still renders a complete correlation surface.

4.   Conclusions

The following conclusions address the two questions presented earlier.

- A lower PSNR hurts the performance of the shift-tolerant QMF–shift-training system at lower levels of resolution. We selected the baseline QMF–shift-training system for our multiresolution analysis.
- Performance declines on the lower levels of decomposition.
- Reasonable performance is achieved at compression ratios of 32:1, 128:1, and 128:1 at decomposition levels one, two, and three, respectively.
- The SCF system does not produce aliasing at the correlation output, even at low bit rates.

9.   Future Extensions

In this section we provide a brief synopsis of possible extensions to the SCF system discussed so far. First, our marriage of PCFs with a subband decomposition currently weights all the input channels equally. It is simple to scale each input channel (i.e., subband output) differently, which may be fruitful for some applications. Given that most of the downsampling effects arise in the upper subbands, attenuating their output response may actually improve the general performance. In addition, a system may want to apply a weight vector to the spectral terms of Eq. (1) to amplify (or attenuate) specific frequencies.

Fig. 10.   Reconstructed image MSE versus bit rate.

Table 6. Multiresolution Performance Summary of SCF System[a]

Subband Level	Probability of Correct Classification
Level one	96.2%, $R \leq 2{:}1$
(2-ft resolution)	89.9%, $R \leq 20{:}1$
Level two	84.8%, $R \leq 8{:}1$
(4-ft resolution)	82.3%, $R \leq 26{:}1$
Level three	74.7%, $R \leq 8{:}1$
(8-ft resolution)	64.6%, $R \leq 32{:}1$

[a]The table reports that the SCF achieves 96.2% correct classification at compression ratios, R, equal to or less than 2:1 on 2-ft-resolution data; likewise, the system results in 89.9% accuracy at compression ratios equal to or less than 20:1 on the same data. The remaining data are read in an analogous manner.

One extreme case of this scenario would be to drop the HH subband completely. This subband has the least tolerance to downsampling. Not using its correlation response may improve performance.

A second suggestion we have is to try the minimal training method outlined in Section 6. In addition, we recommend integrating all the subbands from a multilevel decomposition (that is, incorporate subbands from more than just one level). This modification, however, will require postprocessing because the correlation responses from different levels vary in size.

10. Summary

The SCF system overcomes the downsampling effects of a subband coder and offers multiresolution object recognition. Novel contributions include

• A new pattern-recognition algorithm that uses correlation filters synthesized in the subband coefficient domain and allows for object recognition at the multiple resolutions of a subband coder by operating directly on the subband coefficients.
• A new training methodology to accommodate the downsampling of the subband coefficients.
• A novel method of jointly optimizing the compression capability of a QMF subband filter and the recognition performance of a correlation filter to compensate for downsampling in the subband coefficients.

Of the two proposed methods for downsampling tolerance, training over multiple shifts of the data is the more effective. Optimizing the QMF for downsampling tolerance offers an additional slight improvement in the downsampling tolerance, but the resulting decreased fidelity of the QMF transform impairs recognition accuracy. In summary, we achieve the correct classification performance denoted in Table 6. A more sophisticated encoder would improve the SCF performance at low bit rates. Another benefit of the SCF system is that its output does not exhibit aliasing, even when low bit rates require the encoder to drop whole subbands of information.

In conclusion, we have answered the questions issued at the beginning of this investigation.

• The SCF performs successful recognition for data up to one eighth the resolution of the original data, for compression ratios of at least 20:1. It can still be effective at higher compression ratios, depending on the level of resolution.
• For every reduction in input resolution of one half, the SCF's recognition accuracy decreases by approximately 10%.
• SCF performance effects a tuning curve similar to a rate-distortion curve, allowing a user to select an operating point, balancing requirements of recognition and compression.

References

1. A. Mahalanobis and B. V. K. Vijaya Kumar, "Polynomial filters for higher order correlation and multi-input information fusion," in *Optoelectronic Information Processing* (SPIE, Bellingham, Wash., 1997), pp. 221–231.
2. B. Walls and A. Mahalanobis, "Performance of the MACH filter and DCCF algorithms in the presence of data compression," in *Automatic Target Recognition IX*, F. A. Sadjadi, ed., Proc. SPIE **3718,** 376–387 (1999).
3. T-C. Liu and S. Mitra, "Fingerprint recognition of wavelet-based compressed images by neuro-fuzzy clustering," in *Applications of Fuzzy Logic Technology III*, B. Bosacchi and J. C. Bezdek, eds., Proc. SPIE **2761,** 76–86 (1996).
4. F. B. Shin and D. H. Kil, "Integrated approach to bandwidth reduction and mine detection in shallow water with reduced-dimension image compression and automatic target recognition algorithms," in *Detection and Remediation Technologies for Mines and Minelike Targets II*, A. C. Dubey and R. L. Barnard, eds., Proc. SPIE **3079,** 203–212 (1997).
5. P. C. Miller, "Reduced-resolution synthetic-discriminant-function design by multiresolution wavelet analysis," Appl. Opt. **34,** 865–878 (1995).
6. D. O. North, "An analysis of the factors which determine signal/noise discrimination in pulsed carrier systems," Proc. IEEE **51,** 1016–1027 (1963).
7. G. L. Turin, "An introduction to matched filters," IRE Trans. Inf. Theor. **6,** 311–329 (1960).
8. H. Andrews, *Computer Techniques in Image Processing*, (Academic, New York, 1970), pp. 55–71.
9. R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, (Wiley-Interscience, New York, 1973), pp. 276–284.
10. W. K. Pratt, *Digital Image Processing*, (Wiley, New York, 1978), Parts 5 and 6.
11. C. F. Hester and D. Casasent, "Multivariant technique for multiclass pattern recognition," Appl. Opt. **19,** 1758–1761 (1980).
12. B. V. K. Vijaya Kumar, A. Mahalanobis, and R. Juday, *Correlation Pattern Recognition* (Cambridge U. Press, Cambridge) (to be published).
13. B. V. K. Vijaya Kumar, "Tutorial survey of composite filter designs for optical correlators," Appl. Opt. **31,** 4773–4801 (1992).
14. B. V. K. Vijaya Kumar, "Minimum-variance synthetic discriminant functions," J. Opt. Soc. Am. A **3,** 1579–1584 (1986).
15. A. Mahalanobis, "New correlation filters for symbolic and rule based pattern recognition," Ph.D. thesis (Carnegie Mellon University, Pittsburgh, Pa., 1987).
16. A. Mahalanobis, B. V. K. Vijaya Kumar, and D. Casasent, "Minimum average correlation energy filters," Appl. Opt. **26,** 3633–3640 (1987).

17. A. Mahalanobis and B. Vijaya Kumar, "Optimality of the maximum average correlation height filter for detection of targets in noise," Opt. Eng. **36,** 2642–2648 (1997).

18. K. G. Leib and J. C. Mendelsohn, "Application of the matched filter image correlator to IR, SAR, and visual target data," in Proc. SPIE **519,** 96–105 (1985).

19. W. M. Boerner and A. B. Kostinski, "On the concept of the polarimetric matched filter in high resolution radar imaging," in *Proceedings of the IEEE 1988 International Symposium Digest: Antennas and Propagation* (Institute of Electrical and Electronics Engineers, New York, 1988), vol. 2, pp. 533–536.

20. L. M. Novak, G. J. Owirka, W. S. Brower, and A. L. Weaver, "The automatic target recognition system in SAIP," Lincoln Lab. J. **10,** 187–202 (1997).

21. L. M. Novak, G. J. Owirka, and W. S. Brower, "Performance of a 20-target MSE classifier," in *Radar Sensor Technology III*, R. Trebits and J. L. Kurtz, eds., Proc. SPIE **3395,** 138–151 (1998).

22. L. M. Novak, G. J. Owirka, and A. Weaver, "Automatic target recognition using enhanced resolution SAR data," IEEE Trans. Aerosp. Electron. Syst. **35,** 157–175 (1999).

23. L. Hostetler, "Template based ATR," presented at the Fifty-Second Automatic Target Recognizer Working Group, Space and Naval Warfare Systems Center San Diego, San Diego, Calif., 8 June 1999.

24. A. Mahalanobis, A. Forman, N. Day, M. Bower, and R. Cherry, "Multiclass SAR ATR using shift-invariant correlation filters," Pattern Recogn. **27,** 619–626 (1994).

25. A. Mahalanobis, L. Ortiz, and B. V. K. Vijaya Kumar, "Performance of the MACH filter and DCCF algorithms on the 10-class public release MSTAR data set," in *Algorithms for Synthetic Aperture Radar Imagery VI*, E. G. Zelnio, ed., Proc. SPIE **3721,** 285–291 (1999).

26. A. Mahalanobis, L. Ortiz, B. V. K. Vijaya Kumar, and A. Ezekial, "Correlation ATR performance using Xpatch (synthetic) training data," in *Algorithms for Synthetic Aperture Radar Imagery VII*, E. G. Zelnio, ed., Proc. SPIE **4053,** 340–344 (2000).

27. K. Sayood, *Introduction to Data Compression* (Morgan Kaufmann, Los Altos, Calif., 1996).

28. H. C. Andrews and W. K. Pratt, "Fourier transform coding of images," in *Proceedings of Hawaii International Conference on Systems and Science* (U. Hawaii Press, Honolulu, 1968), pp. 677–679.

29. H. C. Andrews and W. K. Pratt, "Transform image coding," in *Proceedings of the Symposium on Computer Processing in Communications* (Wiley–Interscience, Chichester, UK, 1970), pp. 63–84.

30. G. K. Wallace, "The JPEG still picture compression standard," Commun. ACM **34,** 31–44 (1991).

31. P. P. Vaidyanathan, *Multirate Systems and Filter Banks* (Prentice-Hall, Englewood Cliffs, N.J., 1993).

32. M. A. Vetterli and J. Kovacevic, *Wavelets and Subband Coding* (Prentice-Hall, Englewood Cliffs, N.J., 1995).

33. G. Strang and T. Nguyen, *Wavelets and Filter Banks* (Wellesley-Cambridge, Wellesley, Mass., 1996).

34. P. P. Vaidyanathan, "Theory and design of M-channel maximally decimated quadrature mirror filters with arbitrary M, having the perfect reconstruction property," IEEE Trans. Acoust. Speech and Signal Process. **ASSP-35,** 476–492 (1987).

35. P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect reconstruction techniques," IEEE ASSP Mag. **4,** June 1987, pp. 4–20.

36. M. A. Vetterli, "A theory of multirate filter banks," IEEE Trans. Acoust. Speech Signal Process. **ASSP-35,** 356–372 (1987).

37. P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," Proc. IEEE **78,** 56–93 (1990).

38. A. Croisier, D. Esteban, and C. Galand, "Perfect channel splitting by use of interpolation/decimation techniques," in *Proceedings of the First International Conference on Information Sciences and Systems* (Hemisphere, Washington, D.C., 1977), pp. 443–446.

39. D. W. Carlson, "Optimal tradeoff composite correlation filters," Ph.D. thesis (Carnegie Mellon University, Pittsburgh, Pa., 1996).

40. C. Daniell, "Object recognition in compressed imagery," Ph.D. thesis (California Institute of Technology, Pasadena, California, 2000).

41. C. E. Daniell, A. Mahalanobis, and R. Goodman, "Joint optimization of correlation and subband transform filters for the dual requirements of recognition and compression," submitted to IEEE Trans. Image Process.

42. V. K. Jain and R. E. Crochiere, "Quadrature mirror filter design in the time domain," IEEE Trans. Acoust. Speech Signal Process. **ASSP-32,** 353–360 (1984).

43. S. S. Hemami and R. M. Gray, "Subband filters optimized for lost coefficient reconstruction," IEEE Trans. Signal Proc. **45,** 763–767 (1997).

44. S. S. Hemami, "Reconstruction of compressed images and video for lossy packet networks," Ph.D. thesis (Stanford University, Stanford, Calif., 1995).

45. A. Mahalanobis, B. V. K. Vijaya Kumar, and S. R. F. Sims, "Distance-classifier correlation filters for multiclass target recognition," App. Opt. **35,** 3127–3133 (1996).

46. J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Trans. Signal Process. **41,** 3445–3462 (1993).