Multiple Template Matching Using the Expansion Filter

K. Raghunath Rao and Jezekiel Ben-Arie, *Member, IEEE*

Abstract—This paper describes a multiple-template generalization of a newly developed approach for template matching by signal expansion into a set of non-orthogonal template-similar basis functions. The single-template method is proven to be equivalent to "restoration" of undegraded images using the Wiener filter and optimizes a new and more practically defined matching quality criterion that we call Discriminative signal-to-noise ratio (DSNR). Compared to the widely used matched filtering approach (also known as correlation matching) which is based on projection, expansion matching is based on decomposition and is shown to be more robust in conditions of noise, superposition and severe occlusion. In this paper, we extend the DSNR optimization approach to include more than one template. The generalized expansion filter presented here is optimal in terms of DSNR and can be designed to elicit any desired response for each of the templates, while optimizing the DSNR criterion. Our approach considers additive noise as a parameter and leads to a general formulation, of which many previous approaches (such as the Synthetic Discriminant Function) form special cases. In the case of a single template, the formulation reverts to the previously mentioned Wiener restoration filter.

I. INTRODUCTION

IN many applications, template recognition is required within a limited set of template classes. For example, a class of templates could be defined as the set of views of an object thus providing generic recognition of this object. In many cases, it is possible [11] to formulate a single filter or a small number of filters to perform a classification task. Compared to the paradigm of having a separate filter for every template, this approach is more attractive. In principle, it is desired that the filter could be designed to generate a user-specifiable output for each template. As will be demonstrated in the following sections, such a requirement is not only feasible, but yields the optimal discrimination filter and stable expansion results. One of the applications of such a filter could be to discriminate between patterns of different classes. We employ the method to design a filter for generic face recognition as shown in section V.

In previous publications [1]–[4], we have related three issues–non-orthogonal image representation, Expansion Matching (EXM) of templates and implementation of expansion with restoration techniques—and have developed a robust method for single template matching. The fundamental

approach of EXM is to match a given template with a given image by **expanding the image signal in terms of basis functions (BFs) which are all similar to the template**. We have shown [1] that only minor conditions are required for such a set of dense self-similar compact functions to form a complete basis for square-summable discrete functions of compact support. Hence, the image is expanded into a set of BFs that correspond to versions of the template, translated to all the candidate locations in the image. **The expansion coefficients obtained at a particular location signify the presence of the template at that location**.

Expansion Matching (EXM) is fundamentally different from matched filtering (also known as correlation matching) that maximizes the SNR defined by the ratio of the filter's center response to the power of the noisy response. EXM is a **decomposition process** that maximizes a novel matching quality criterion called Discriminative Signal-to-Noise Ratio (DSNR). DSNR emphasizes peak sharpness since it minimizes the sidelobes of the response, and is more relevant to matching than optimizing the traditional SNR since the DSNR considers as "noise" even the filter's off-center response to the template. Hence, EXM yields much sharper peaks in comparison to the widely used correlation matching that maximizes the SNR—which disregards the sidelobe response—and thus generates broad peaks. Experimental results with natural signals show that EXM outperforms correlation matching by more than 20 dB DSNR [1]–[3]. This results in much less spurious responses and a more robust performance in noise and even more important: in conditions of severe occlusion. Experiments show that templates that are occluded up to 80% are still successfully recognized [1]–[3] with diminished spurious responses. Furthermore, expansion being a decomposition process, is ideally suited for template recognition in sonar and radar signals, where superposition of templates with clutter or echo signals frequently occur. Expansion has also been successfully employed for edge detection, feature extraction and generic face recognition [4].

We have also established that signal expansion with dense template-similar BFs is precisely equivalent to restoration of the image with a "blurring function" that matches the template. The image is usually sharp and does not need restoration, but a Minimum Squared Error (MSE) restoration (using the Wiener filter) with the template as the blurring function, yields exactly the coefficients of the EXM. Similar results can also be obtained by using the adaptive lattice described in [3], [5], [6].

In this paper, we extend the DSNR optimization paradigm to include more than one template. The response to each

individual template can be constrained to any desired value, and optimizing the DSNR criterion results in a generalized formulation of a single filter that can respond to multiple templates. For specific template recognition, the desired response from each template could be different for different classes, while for generic or invariant recognition, the same response is desired from all templates within a class. Unlike all other previous works, our approach considers additive noise as a parameter in the filter formulation. This important consideration leads to a generalized filter and we find that special cases of our filter exactly correspond to previous formulations. The case of zero additive noise corresponds to the Minimum Average Correlation Energy (MACE) filter [12], and if the noise power approaches infinity, our filter leads to the Synthetic Discriminant Function (SDF) [7] or the Minimum Variance SDF (MVSDF) approach [8], depending on the color of the noise. We present the formal derivation of our filter in the section III and relevant experimental results in section VI.

A. Comparitive Survey of Correlation-Based Matching Methods

We would like to point out that all linear, shift-invariant filtering methods are correlation-based, since the end result is obtained by correlating the image with a filter. In the case of direct correlation (matched filtering in white noise) the filter takes the form of the template itself. We now present a survey of works that developed different filters for correlation-based matching.

An early approach to the synthesis of a single linear shift-invariant filter that caters to multiclass recognition tasks is due to Casasent [7]. The Synthetic Discriminant Function (SDF) that he proposed, was formulated as a weighted average of the various templates of the classes under consideration. The weights are computed to satisfy user defined constraints on the filter's correlation peak response to each template. Different constraints correspond to different types of template recognition problems, eg. intraclass, interclass, multilevel SDFs etc. However, this scheme does not address the optimization of peak sharpness, and in the case of single template recognition, simply reverts to the direct correlation matching approach. Furthermore, additive noise is not considered in the formulation at all.

A minimum variance (MV) paradigm was applied to the SDF [8] and this optimizes the SNR (as defined in the conventional matched filtering sense) at the correlation peak. For the single template recognition case this approach yields exactly the matched filter result. This method does not constrain the sharpness of the correlation peak in any manner.

Mostafavi and Smith [9] define a metric called Peak-to-Sidelobe Ratio (PSR) (sometimes called Signal-to-Clutter Ratio or SCR) which is the ratio of the expected peak value at the template center, to the standard deviation of the total response "far" from this center. Though one could possibly formulate one criterion from the other, our DSNR definition is basically different, since we do not consider the notion of "far" from the center in defining the unwanted filter response.

According to the DSNR, any response of the filter which is not at the template center is regarded as unwanted clutter . Our work stems from the use of non-orthogonal expansions for matching. This is also equivalent to applying the linear operator that maximizes the DSNR criterion, which in turn is implemented by the Wiener restoration filter. These newly developed relationships are briefly discussed in Appendix B.

A "maximization" of the peak PSR was suggested by Kallman [10] with a minimax formulation based on the assumption that the ideal filter is a linear combination of the templates. Apart from the huge computation complexity of this approach, the basic assumption of the approach restricts it from yielding the filter with maximal possible PSR, and the formulation also does not consider additive noise.

The PSR correlation filter [11] maximizes the PSR, but only in the vicinity of the correlation peak and not all over the filter response. Furthermore, PSR filters do not allow any control on the amplitude of the peak response for the various template classes and require shifted versions of the templates as an input, and thus, a very large number of templates.

Another measure called Peak-to-Correlation Energy (PCE) [14] also characterizes peak sharpness. Works that performed optimization on this criterion yielded the inverse filter which is unstable and very sensitive to noise. There have been no attempts to include the noise as a parameter in the formulation and perform the peak sharpness optimization to yield stable results.

The Minimum Average Correlation Energy (MACE) formulation [12] seeks a single filter that satisfies the user defined constraints (as in the SDF) and also minimizes the average energy of the filter's response to all the templates. In essence, this leads to a sharper correlation peaks than the matched filter approach. The MACE filter is generalized and special cases yield the minimum variance SDF [8] or the matched filter [13]. However, as in almost all the works available in the literature, this filter also reverts to the matched filter formulation in the single template classification problem. Furthermore, additive noise is not considered in the formulation.

Another class of filters called Fractional Power Filters (FPFs) [14] use nonlinearities in the frequency domain of the correlator. This class of filters specialize to the inverse filter, the matched filter and the Phase Only Filter (POF), but offer no new conceptual contribution.

After an extensive survey of work done in this area, we find that most previous approaches neglect the sharpness of the correlation peak. Those works that do consider sharpness of the correlation peak by some measure or other, do not address the problem of additive noise as an integral part of their formulation.

II. A BRIEF OVERVIEW OF EXPANSION MATCHING

This section briefly presents the results of the Expansion Matching (EXM) formulation for matching a single template. A more detailed discussion of these results is presented in Appendix B. These results serve as a background for the design of the expansion filter in section 3. We use discrete 1D signals for the sake of simplicity. Extension to higher dimensions is

straightforward using lexicographically ordered vectors and block circulant matrices.

Given an M-point signal $s(x)$ that has a corresponding vector form $\mathbf{S}$, it is desired to match a template $\psi(x) \equiv \Psi$ by correlating (not convolving) the signal $\mathbf{S}$ with a filter $\theta(x) \equiv \Theta$ to yield a result $c(x) \equiv \mathbf{C}$. The corresponding matrix-vector notation is

$$[\mathbf{S}]\Theta = \mathbf{C}; [\mathbf{S}] = [\mathbf{S}^{(1)}\mathbf{S}^{(2)}\cdots\mathbf{S}^{(M)}]^T \qquad (1)$$

where $\mathbf{S}^{(i)}$ corresponds to the signal shifted to the i-th position, and thus $[\mathbf{S}]$ forms a circulant correlation matrix. We also know that the signal $\mathbf{S}$ contains only the template at location l, i.e. $s(x) = \psi(x - l)$ or $\mathbf{S} = \Psi^{(l)}$. The Discriminative Signal-to-Noise Ratio (DSNR) is a matching quality measure and is defined as

$$\text{DSNR} = 10 \log \frac{(\Theta^T \mathbf{S}^{(l)})^2}{\sum_{j \neq l}(\Theta^T \mathbf{S}^{(j)})^2} = 10 \log \frac{c^2(l)}{\sum_{j \neq l} c^2(j)} \qquad (2)$$

Unlike the SNR criterion, the DSNR penalizes any response that is not at the center of the template. Optimization of the DSNR criterion will therefore enhance peak sharpness in the response. Such an optimization [1–4] leads to a filter Θ given by

$$\Theta = \frac{1}{k}\left[[\mathbf{S}]^T[\mathbf{S}]\right]^{-1}\mathbf{S}^{(l)} = \frac{1}{k}[\mathbf{R}_{ss}]^{-1}\mathbf{S}^{(l)};$$

$$k = \frac{\Theta^T \mathbf{S}^{(l)}}{\Theta^T[\mathbf{S}]^T[\mathbf{S}]\Theta} \qquad (3)$$

where k is only an amplitude scaling factor (a scalar) and $[\mathbf{R}_{ss}]$ is the autocorrelation matrix of the signal $\mathbf{S}$. Note that the autocorrelation matrices of the signal and the template are identical, i.e. $[\mathbf{R}_{ss}] = [\mathbf{R}_{vv}]$.

An interesting equivalence to (3) is derived [1] if we expand the signal $s(x)$ into basis functions (BFs) $\psi^{(i)}(x)$ that correspond to shifted versions of the template $\psi(x)$, i.e.

$$s(x) = \sum_{i=1}^{M} c_i \psi^{(i)}(x) ; \ \psi^{(i)}(x) = \psi(x - i) \qquad (4)$$

By minimizing the squared error in this non-orthogonal signal representation, the result for the coefficients $\mathbf{C}$ turns out to be

$$\mathbf{C} = [\mathbf{R}_{vv}]^{-1}[\Psi]\mathbf{S} \qquad (5)$$

This result is identical to the result of correlating the signal with the optimal DSNR filter Θ in (3). Thus DSNR optimization is equivalent to a non-orthogonal expansion of the signal with BFs that are shifted versions of the template itself.

Another look at (4) reveals that it can be written slightly differently as

$$s(x) = \sum_{i=1}^{M} c_i \psi(x - i) = c(x) * \psi(x) \qquad (6)$$

where $*$ denotes discrete convolution. Thus, the signal is the convolution of the coefficients with the template. To obtain the coefficients, one has to "restore" the signal $s(x)$ with the

template $\psi(x)$ as the blurring function. We use a minimum squared error (MSE) restoration paradigm [1] and obtain as the result, the Wiener filter

$$\Theta = \frac{1}{k}[[\mathbf{R}_{\psi\psi}] + [\mathbf{R}_{\lambda\lambda}]]^{-1}\Psi^{(l)} \qquad (7)$$

The matrices $[\mathbf{R}_{\psi\psi}]$ and $[\mathbf{R}_{\lambda\lambda}]$ are the autocorrelation matrices of the template and the additive noise $\lambda(x)$ respectively. If there is no noise, this is exactly equivalent to (3).

Thus, we establish that for matching a single template to an image, the optimal DSNR operator, the template-similar non-orthogonal expansion and the MSE restoration of the signal by the template, all lead to exactly the same result.

III. OPTIMAL DSNR MULTIPLE TEMPLATE MATCHING: THE EXPANSION FILTER

In this section, we use the DSNR criterion to design a single filter for the recognition of multiple templates. In the most general form, we desire a filter that maximizes the DSNR criterion as well as elicits a set of desired responses from the given set of templates. In our formulation, we include the concept of additive noise, and correspondingly, seek to satisfy the given conditions using statistical expectations. The problem is formulated in one-dimension using matrix-vector notation. The results can easily be generalized into two dimensions by using lexicographically ordered vectors and block-circulant matrices.

Given a set of templates $\Psi_i: i = 1 \cdots N$, we expect a set of noisy input templates $\mathbf{S}_i = \Psi_i + \Lambda$, where Λ is a random noise vector. We wish to design a filter Θ that yields as its expected peak correlation to each template $\mathbf{S}_i$, a desired response u_i, i.e.,

$$E\{\Theta^T \mathbf{S}_i\} = u_i; i = 1 \cdots N \qquad (8)$$

The correlation coefficient vectors $\mathbf{C}_i$ (filter response) for the i-th template are obtained as

$$\mathbf{C}_i = [\mathbf{S}_i]\Theta; i = 1 \cdots N \qquad (9)$$

where $[\mathbf{S}_i]$ is a circulant correlation matrix obtained by circulating the noisy input template $\mathbf{S}_i$. Note that Θ is correlated with the signal and not convolved. Ideally $\mathbf{C}_i = [0 \cdots 0, u_i, 0 \cdots 0]^T$ where u_i is located at the i-th pattern's center. $[\mathbf{S}_i]$ is also equivalent to

$$[\mathbf{S}_i] = [\Psi_i] + [\Lambda]; i = 1 \cdots N \qquad (10)$$

where $[\Psi_i]$ and $[\Lambda]$ are circulant correlation matrices for the templates Ψ_i and the noise Λ respectively.

In addition to the constraints of (8), we also wish to maximize the DSNR criterion for the correlation response of the filter to each input template. The DSNR to be maximized for the i-th template is defined by the ratio between the power of the response at the template center to the total off-center power. In terms of expected values, this is given by

$$DSNR_i = 10 \log \frac{E\{(\Theta^T \mathbf{S}_i)^2\}}{E\{\mathbf{C}_i^T \mathbf{C}_i\} - E\{(\Theta^T \mathbf{S}_i)^2\}} \to \max \qquad (11)$$

Thus, high DSNR implies a large response at the template-center and relatively small response off the center of the

template. Since the center-response power is included in the constraint of (8) (the power of the noisy response being identical all over the signal), to maximize $DSNR_i$ we need only minimize the first term in the denominator

$$E\{\mathbf{C}_i^T \mathbf{C}_i\} \to \min \tag{12}$$

Since a global minimization of all the DSNR criteria for each input template is not possible [7], we seek to minimize a weighted sum of the individual DSNRs by

$$E\left\{\sum_{i=1}^{N} \alpha_i \mathbf{C}_i^T \mathbf{C}_i\right\} \to \min; \sum_{i=1}^{N} \alpha_i = 1 \tag{13}$$

where α_i are weighting factors, typically chosen to be $1/N$. Combining the requirements of (8) and (13), we need to minimize a single objective function defined as

$$J(\Theta) = E\left\{\sum_{i=1}^{N} \alpha_i \mathbf{C}_i^T \mathbf{C}_i\right\} + \sum_{i=1}^{N} \xi_i (E\{\Theta^T \mathbf{S}_i\} - u_i) \to \min$$

$$= \sum_{i=1}^{N} \left(\alpha_i E\{\mathbf{C}_i^T \mathbf{C}_i\} + \xi_i \Theta^T E\{\mathbf{S}_i\} - \xi_i u_i\right) \to \min \tag{14}$$

where the Lagrange multipliers ξ_i are to be chosen to satisfy the constraints of (8).

We assume the noise Λ to be zero-mean and stationary, and thus the first-order noise terms in the expectations drop out, yielding the identities

$$E\{\mathbf{S}_i\} = E\{\Psi_i + \Lambda\} = \Psi_i \tag{15}$$

and

$$E\{\mathbf{C}_i^T \mathbf{C}_i\} = \Theta^T E\{[\mathbf{S}_i]^T [\mathbf{S}_i]\} \Theta$$
$$= \Theta^T E\{[\Psi_i]^T [\Psi_i] + [\Psi_i]^T [\Lambda]$$
$$+ [\Lambda]^T [\Psi_i] + [\Lambda]^T [\Lambda]\} \Theta$$
$$= \Theta^T [[\mathbf{R}_{\psi\psi i}] + [\mathbf{R}_{\lambda\lambda}]] \Theta \tag{16}$$

where $[\mathbf{R}_{\psi\psi i}] = [\Psi_i]^T [\Psi_i]$ and $[\mathbf{R}_{\lambda\lambda}] = E\{[\Lambda]^T [\Lambda]\}$ are defined as the autocorrelation matrices of the i-th template and the noise respectively.

We can now rewrite the objective function of (14) as

$$J(\Theta) = \sum_{i=1}^{N} \left(\alpha_i \Theta^T [[\mathbf{R}_{\psi\psi i}] + [\mathbf{R}_{\lambda\lambda}]] \Theta\right.$$
$$\left. + \xi_i \Theta^T \Psi_i - \xi_i u_i\right) \to \min \tag{17}$$

Setting the first derivative to zero, we obtain

$$\frac{\partial J(\Theta)}{\partial \Theta} = \sum_{i=1}^{N} (2\alpha_i [[\mathbf{R}_{\psi\psi i}] + [\mathbf{R}_{\lambda\lambda}]] \Theta + \xi_i \Psi_i) = 0 \tag{18}$$

which yields the solution for Θ as

$$\Theta = -\frac{1}{2} \left[\sum_{i=1}^{N} \alpha_i [\mathbf{R}_{\psi\psi i}] + [\mathbf{R}_{\lambda\lambda}]\right]^{-1} \sum_{j=1}^{N} \xi_j \Psi_j$$

$$= [\mathbf{R}]^{-1} \sum_{j=1}^{N} \tilde{\xi}_j \Psi_j \tag{19}$$

where $[\mathbf{R}] = \sum_{i=1}^{N} \alpha_i [\mathbf{R}_{\psi\psi i}] + [\mathbf{R}_{\lambda\lambda}]$, and $\tilde{\xi}_j = -\xi_j/2$.

To find $\tilde{\Xi} = [\tilde{\xi}_1 \cdots \tilde{\xi}_N]^T$ we enforce the constraints of (8)

$$u_i = \Theta^T \Psi_i = \left([\mathbf{R}]^{-1} \sum_{j=1}^{N} \tilde{\xi}_j \Psi_j\right)^T \Psi_i = \sum_{j=1}^{N} \tilde{\xi}_j \Psi_j^T [\mathbf{R}]^{-1} \Psi_i;$$
$$i = 1 \cdots N \tag{20}$$

Or in matrix form, the solution to this set of equations is written as

$$\tilde{\Xi} = [\mathbf{A}]^{-1} \mathbf{U} \quad ; \quad [\mathbf{A}]_{ij} = \Psi_i^T [\mathbf{R}]^{-1} \Psi_j \tag{21}$$

where $[\mathbf{A}]$ is a real symmetric $N \times N$ matrix which is non-singular for distinct and non-zero Ψ_i, and $\mathbf{U} = [u_1 \cdots u_N]^T$ defines the constraints.

Thus (19) and (21) yield the desired expansion filter Θ. If we choose $N = 1$, i.e. the single template recognition case, the solution reverts to the Wiener filter formulation, and corresponds exactly to a minimum squared error non-orthogonal expansion of the signal with template-similar basis functions [1]–[4]. This is elaborated in Appendix B.

IV. FREQUENCY DOMAIN FORMULATION OF THE EXPANSION FILTER

We now present a frequency domain interpretation of the formulation of section III and show that it leads to an efficient implementation of the expansion filter design technique. We can diagonalize the circulant matrix $[\mathbf{R}]$ using the unitary DFT matrix $[\mathbf{W}]$ and write (19) as

$$\Theta = [[\mathbf{W}][\mathbf{D}_{RR}][\mathbf{W}]^{-1}]^{-1} \sum_{j=1}^{N} \tilde{\xi}_j$$

$$\Psi_j = [\mathbf{W}][\mathbf{D}_{RR}]^{-1} \sum_{j=1}^{N} \tilde{\xi}_j [\mathbf{W}]^{-1} \Psi_j \tag{22}$$

where $[\mathbf{W}]_{mn} = M^{-0.5} \exp(j2\pi mn/M)$, and $[\mathbf{D}_{RR}]$ is a diagonal matrix with the diagonal elements containing the DFT of the first row of the autocorrelation matrix $[\mathbf{R}]$. These diagonal elements correspond to the weighted sum of the individual power spectrums $[\mathbf{D}_{\psi\psi i}]$ and $[\mathbf{D}_{\lambda\lambda}]$ of each given template and the additive noise respectively

$$[\mathbf{D}_{RR}] = \sum_{i=1}^{N} \alpha_i [\mathbf{D}_{\psi\psi i}] + [\mathbf{D}_{\lambda\lambda}] \tag{23}$$

where $[\mathbf{D}_{\psi\psi i}] = [\mathbf{W}]^{-1} [\mathbf{R}_{\psi\psi i}][\mathbf{W}]$ and $[\mathbf{D}_{\lambda\lambda}] = [\mathbf{W}]^{-1} [\mathbf{R}_{\lambda\lambda}][\mathbf{W}]$. We can now rewrite (22) as

$$([\mathbf{W}]^{-1} \Theta) = \left[\sum_{i=1}^{N} \alpha_i [\mathbf{D}_{\psi\psi i}] + [\mathbf{D}_{\lambda\lambda}]\right]^{-1} \sum_{j=1}^{N} \tilde{\xi}_j ([\mathbf{W}]^{-1} \Psi_j) \tag{24}$$

All the vectors in parentheses in (24) correspond to DFTs, and the inverse of the diagonal matrix is easily obtained by

the reciprocal of its diagonal values. Thus, in the frequency domain, we can write

$$\Theta(\omega) = \frac{\sum_{j=1}^{N} \tilde{\xi}_j \mathbf{X}_j(\omega)}{\sum_{i=1}^{N} \alpha_i S_{\psi\psi i}(\omega) + S_{\lambda\lambda}(\omega)} \qquad (25)$$

where $\mathbf{X}_i = [\mathbf{W}]^{-1}\Psi_i$ is the DFT of the i-th template Ψ_i, $\mathbf{X}_i(\omega)$ refers to the w-th component of the vector $\mathbf{X}_i$, $S_{\psi\psi i}$ $(\omega) = \mathbf{X}_i(\omega)\overline{\mathbf{X}}_i(\omega)$, $S_{\lambda\lambda}(\omega)$ is the power spectrum of the noise, and the bar symbol represents complex conjugation. Eq. (25) offers an efficient computational method for the design of the expansion filter Θ. With $N = 1$, this is easily recognized as the Wiener filter result.

We also need to solve the linear system of equations in (21). Here also, instead of evaluating the elements of matrix $[\mathbf{A}]$ in the spatial domain, it is possible to obtain a simpler frequency domain interpretation. From (21), we know that

$$[\mathbf{A}]_{ij} = \Psi_i^T [\mathbf{R}]^{-1} \Psi_j \qquad (26)$$

If we diagonalize $[\mathbf{R}]$ as before, we obtain

$$[\mathbf{A}]_{ij} = \Psi_i^T [\mathbf{W}] \left[\sum_{i=1}^{N} \alpha_i [\mathbf{D}_{\psi\psi i}] + [\mathbf{D}_{\lambda\lambda}] \right]^{-1} [\mathbf{W}]^{-1} \Psi_j$$

$$= ([\overline{\mathbf{W}}]^{-1}\Psi_i)^T \left[\sum_{i=1}^{N} \alpha_i [\mathbf{D}_{\psi\psi i}] + [\mathbf{D}_{\lambda\lambda}] \right]^{-1} ([\mathbf{W}]^{-1}\Psi_j) \qquad (27)$$

which can be rewritten in the frequency domain as

$$[\mathbf{A}]_{ij} = \sum_{\omega} \left(\frac{\overline{\mathbf{X}}_i(\omega)\mathbf{X}_j(\omega)}{\sum_{k=1}^{N} \alpha_k S_{\psi\psi k}(\omega) + S_{\lambda\lambda}(\omega)} \right) \qquad (28)$$

The frequency domain design approach offers considerable savings in computation, since the DFTs of the templates need be computed only once. The square matrix $[\mathbf{A}]$ of (21) has dimension equal to the number of templates N which is usually much smaller than the signal dimensionality, and thus can be inverted directly.

V. SPECIAL CASES OF THE EXPANSION FILTER FORMULATION

Our extensive literature survey reveals that our approach is the only one that considers additive noise as a parameter in the filter design and thus results in a generalized formulation. If we assume the noise to be white and of very large magnitude compared to the templates, our filter approaches the Synthetic Discriminant Function (SDF) [8]. If the noise is not white, but still very large compared to the templates, our filter tends to the Minimum Variance SDF (MVSDF) paradigm [9]. If we neglect the noise totally (equivalent to zero noise condition), then our filter design results in the Minimum Average Correlation Energy (MACE) method [7]. The proofs of these results are presented in Appendix A.

Thus we see, that as the noise parameter is varied, at the zero noise end our filter yields the MACE filter and at the other extreme, with infinite noise we obtain the MVSDF or SDF results. However, in reality the noise almost always assumes an intermediate value and therefore (25) is more practical.

Another important upshot of our work is the relation between DSNR optimization, non-orthogonal expansion and Wiener MSE restoration in the case of the filter for a single template [1]–[3].

VI. EXPERIMENTAL RESULTS

The method of Expansion Matching (EXM) is equivalent to the above expansion filter formulation when the number of templates $N = 1$. Experiments under conditions of additive noise [1]–[4] show that EXM is superior to matched filtering (correlation approach). In fact, in the limiting case of infinite noise, the expansion filter approaches the matched filter. EXM is also very robust in conditions of severe occlusion. Templates occluded up to 80% have been successfully recognized [1]–[4]. EXM is especially suited to matching superimposed templates that occur in radar or sonar applications. Since expansion is a linear decomposition process it yields perfect and distinct peaks, while correlation fails to resolve between superimposed templates [1]–[4].

Fig. 1(a) shows two anti-symmetric truncated ramp functions used as synthetic templates. The filter requirements are to elicit a response of -1 from the left hand template and $+1$ from the right hand template. The test signal of Fig. 1(b) has the two templates with additive noise to give a SNR of 11 dB. The expansion filter for the case of zero noise power is shown in Fig. 1(c). This corresponds exactly to the MACE filter. Fig. 1(d) is the result of applying the MACE filter to the noisy test signal of Fig. 1(b), and shows noisy spurious peaks (denoted by question marks), substantial off-center response and a DSNR of 15.5 dB. The optimal DSNR expansion filter in Fig. 1(e) yields the matching results in Fig. 1(f) which displays two sharp and well localized peaks (marked by arrows) with minimal off-center response and the maximum possible DSNR = 21.3 dB. Fig. 1(g) shows the SDF or MVSDF filter (expansion filter with noise power approaching infinity). Fig. 1(h) shows the response to the test signal, which has broad peaks and large off-center spurious response and spurious peaks (DSNR = 9.7 dB).

Fig. 2(a) shows two templates extracted from natural images embedded in white noise (SNR = 20 dB). These templates correspond to features in one scanline of an image of a real world scene. The optimal DSNR multiple template approach is used to design a filter that elicits -1 for the first template and $+1$ for the second template. The result in Fig. 2(b) uses zero noise as the design parameter, and corresponds to the MACE approach. The DSNR is only 17.8 dB and noisy response and many spurious peaks (denoted by question marks) are observed. Our approach (Fig. 2(c)) results in two distinct impulse functions (marked by arrows) with minimal off-center response and the maximum possible DSNR = 22.2 dB. Fig. 2(d) shows the SDF or MVSDF approach (infinite noise expansion filter) which yields broad peaks and large off-center spurious response (DSNR = 10.3 dB).

Fig. 3(a) shows a test image with three different faces. Three "generic" face filters are designed using the MACE, Expansion and SDF approaches by presenting templates of these faces and demanding unit response to each face. There
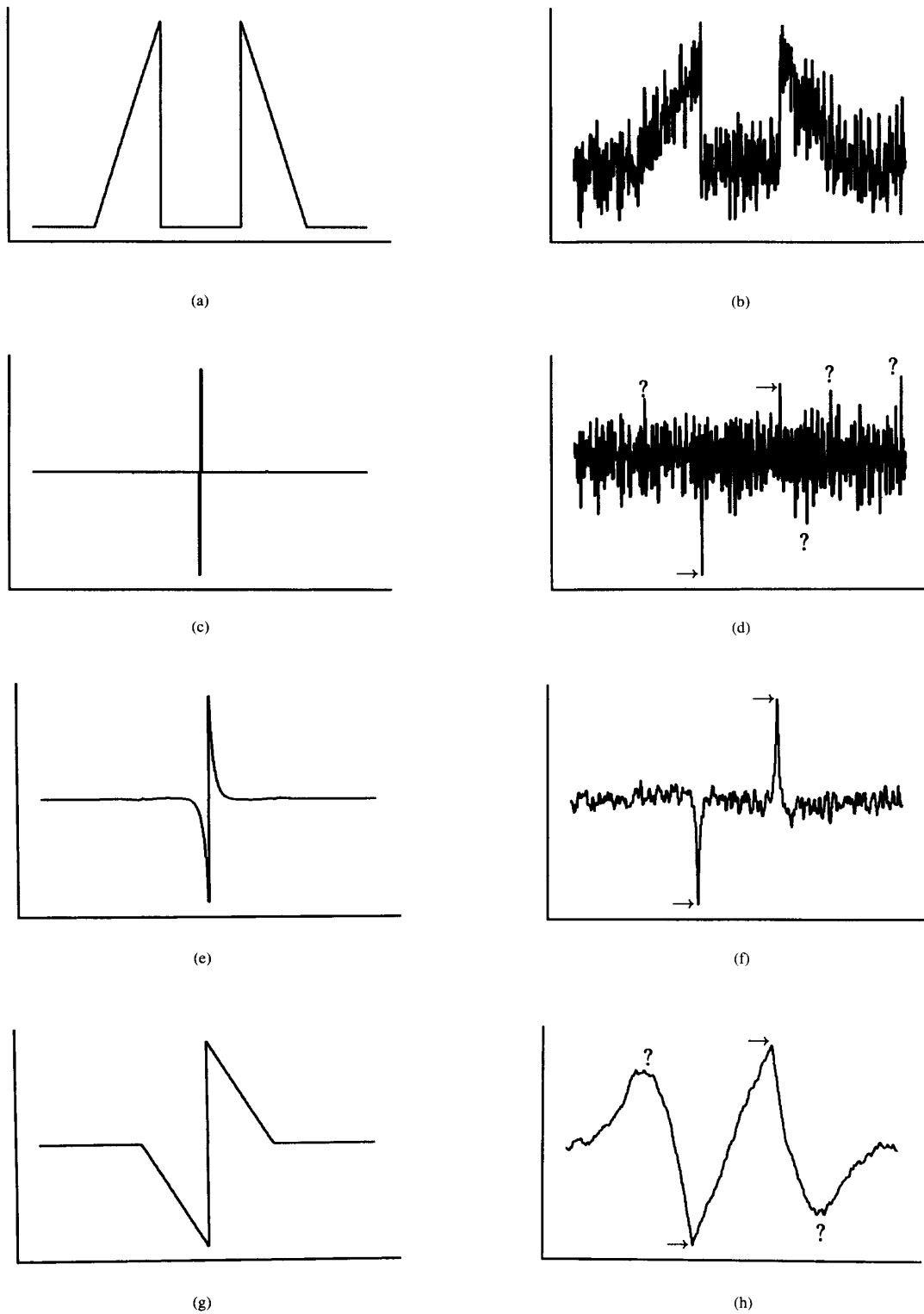
Fig. 1. (a) Two synthetic ramp templates; (b) Noisy Test signal (SNR = 11 dB); (c) MACE filter (zero noise Expansion filter); (d) MACE filter response. Note large and noisy off-center reponse with unwanted spurious peaks denoted by '?' (proper peaks denoted by arrows). DSNR = 15.5 dB. (e) Optimal DSNR Expansion filter (Noise term = 11 dB); (f) Expansion filter response. Note clean and sharp peaks with minimal off-center spurious. DSNR = 21.3 dB; (g) SDF or MVSDF filter (infinite noise Expansion filter); (h) SDF or MVSDF filter response. Note the large off-center response and broad peaks. DSNR = 9.7 dB.
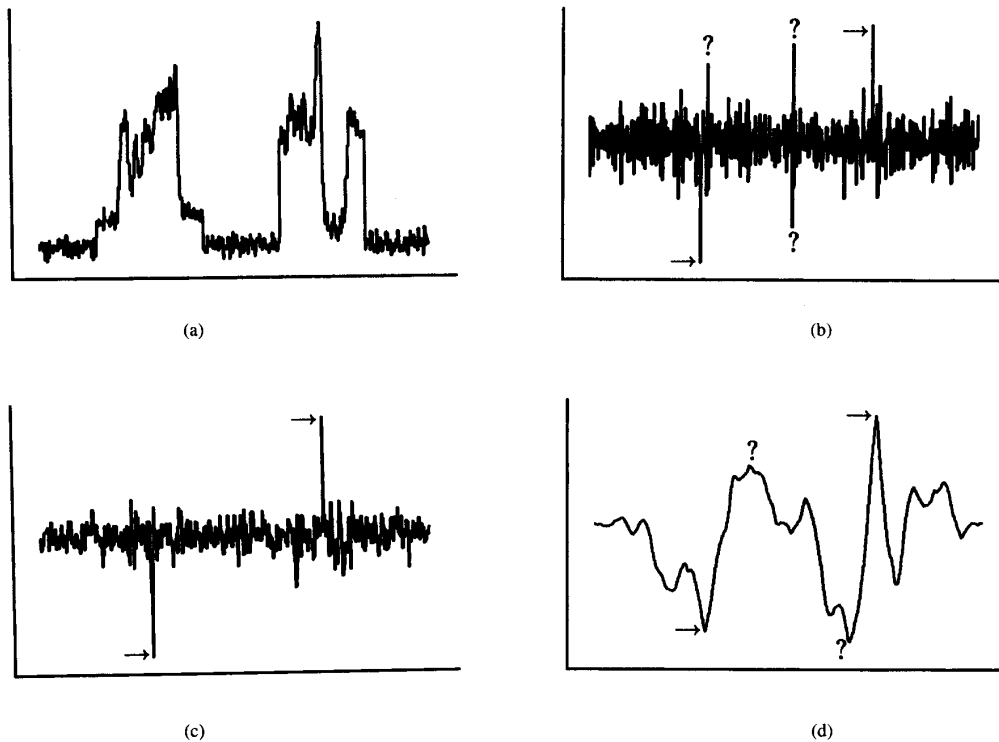
Fig. 2. (a) A signal with two dissimilar templates and additive noise. SNR = 20 dB; (b) Matching results with MACE filter (zero noise Expansion filter), DSNR = 17.8 dB. Note noisy response and large spurious peaks (marked by '?'); (c) Expansion Matching results with noise term = 20 dB. Only desired peaks are present (marked by arrows) and minimal off-center response. DSNR = 22.2 dB; (d) Matching results with the SDF or MVSDF method (infinite noise Expansion filter), DSNR = 10.3 dB. Note the blurred peaks and large spurious response.

is no additive noise in this example and thus the only distortion between the templates and the input image is due to the windowing of the templates. Fig. 3(b) shows the result of using the MACE approach which corresponds to the zero noise expansion filter. One can observe isolated but strong spurious responses (marked by '?'), though the matching peaks are sharp (marked by arrows) since there is no additive noise in the image. However, the expansion filter with an assumed input white noise of SNR = 25 dB yields the superior results in Fig. 3(c). Here, the peaks are sharp and stronger with relatively low background response. The SDF result in Fig. 3(d) (infinite noise expansion filter) shows broad and poorly defined peaks with stantial spurious response (marked by '?').

As a comparison to conventional correlation-based single template matching, Fig. 4(a) shows the result of matched filtering (or direct correlation) of Fig. 3(a) using the template corresponding to the top-left face in Fig. 3(a). Similarly Fig. 4(b) and Fig. 4(c) present matched filtering results using the templates of the center and top-right faces of Fig. 3(a) respectively. It can be seen that these matching results have very weak and broad peaks (marked by arrows) that are hard to detect. Furthermore, there is substantial off-center spurious response and clutter (marked by '?').

Fig. 5(a) shows a test image with many different office items. The pliers and scissors are substantially occluded.

The MACE, Expansion and SDF filters are designed using unoccluded templates of the pliers and scissors and demanding unit response to each. These templates were independently grabbed using a TV camera and thus have inherent noise, scaling, and rotation distortions in them. Fig. 5(b) shows the MACE response which detects only the peak for the scissors (marked by the arrow) and has substantial spurious response (marked by '?'). The Expansion filter with an assumed input white noise of SNR = 18 dB yields the results in Fig. 5(c). One can observe the sharp and well localized peaks at the centers of the pliers and scissors and the minimal off-center response. The SDF result in Fig. 5(d) yields only one highly blurred peak (the scissors) and very large off-center response.

VII. DISCUSSION

In this paper, we have presented a method for simultaneous matching of multiple templates that optimizes a novel similarity measure called Discriminative Signal-to-Noise Ratio (DSNR). The DSNR is more practical than the SNR since it considers all off-center response of the filter as unwanted noise, and thus optimal DSNR filters respond to their templates with sharp peaks and minimal off-center spurious response. The special case of matching a single template reverts to a non-orthogonal expansion of the image with template-similar
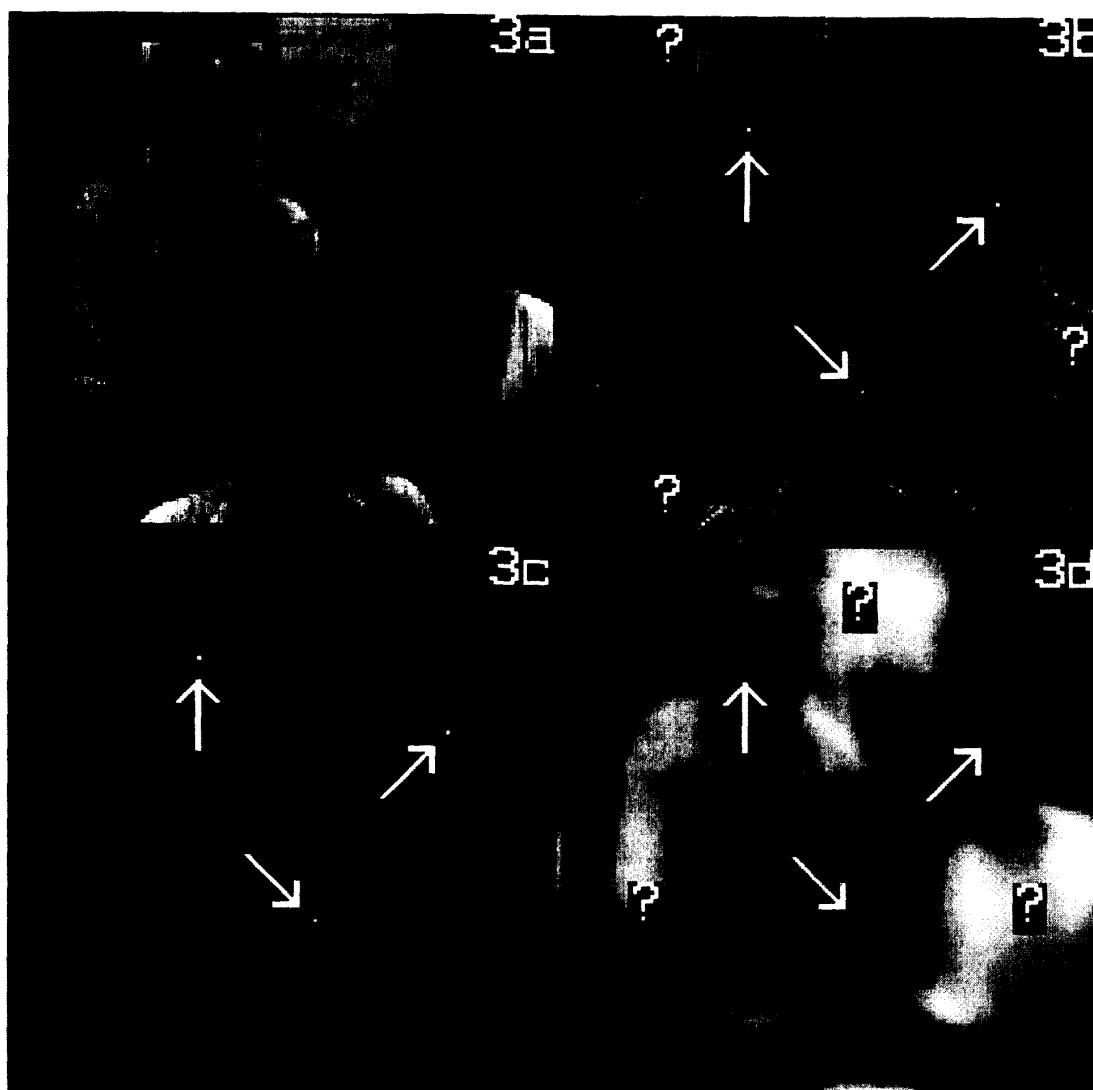
Fig. 3. (a) Test image with three dissimilar faces; (b) MACE response (zero noise Expansion filter) for matching generic face with unit response. Note spurious off-center responses; (c) Corresponding Expansion filter response (assumed SNR = 25 dB). Note the sharp peaks and less background response than MACE result; (d) SDF response (infinite noise Expansion filter) has broad and poorly defined peaks with substantial off-center spurious.

basis functions, which has been proven to be equivalent to minimum squared error restoration using the Wiener filter [1].

Additive noise is part of our formulation and this results in a generalized filter. Special cases include the MACE, MVSDF and SDF methods. It can be seen from practical experiments in Section V that the Expansion filter yields better results than these methods, since the amount of expected input noise can be varied to accomodate the amount of distortion of the template in the image. The MACE result is very sensitive to template distortions such as noise, occlusion etc. Choosing an appropriate noise power in the Expansion formulation solves this problem. In practice, even rough estimates of the noise power demonstrate marked improvement over MACE results. At the other extreme, the SDF response is not discriminative enough, with broad, poorly defined peaks and large responses to unrelated features in the image. The Expansion filter always yields sharp peaks and suppresses response to irrelevant features.

The expansion approach is robust to noise as demonstrated by Figs. 1 and 2. Since the Expansion approach is tuned to the amount of additive noise, the matching results are optimally sharp and have minimal spurious response. However, the effectiveness of the Expansion approach is not limited to additive noise conditions alone. This is particularly visible in the results of Fig. 4 in matching occluded templates. Furthermore, a simple white noise assumption in the filter design yields sharp and robust matching results. The matching results are not sensitive
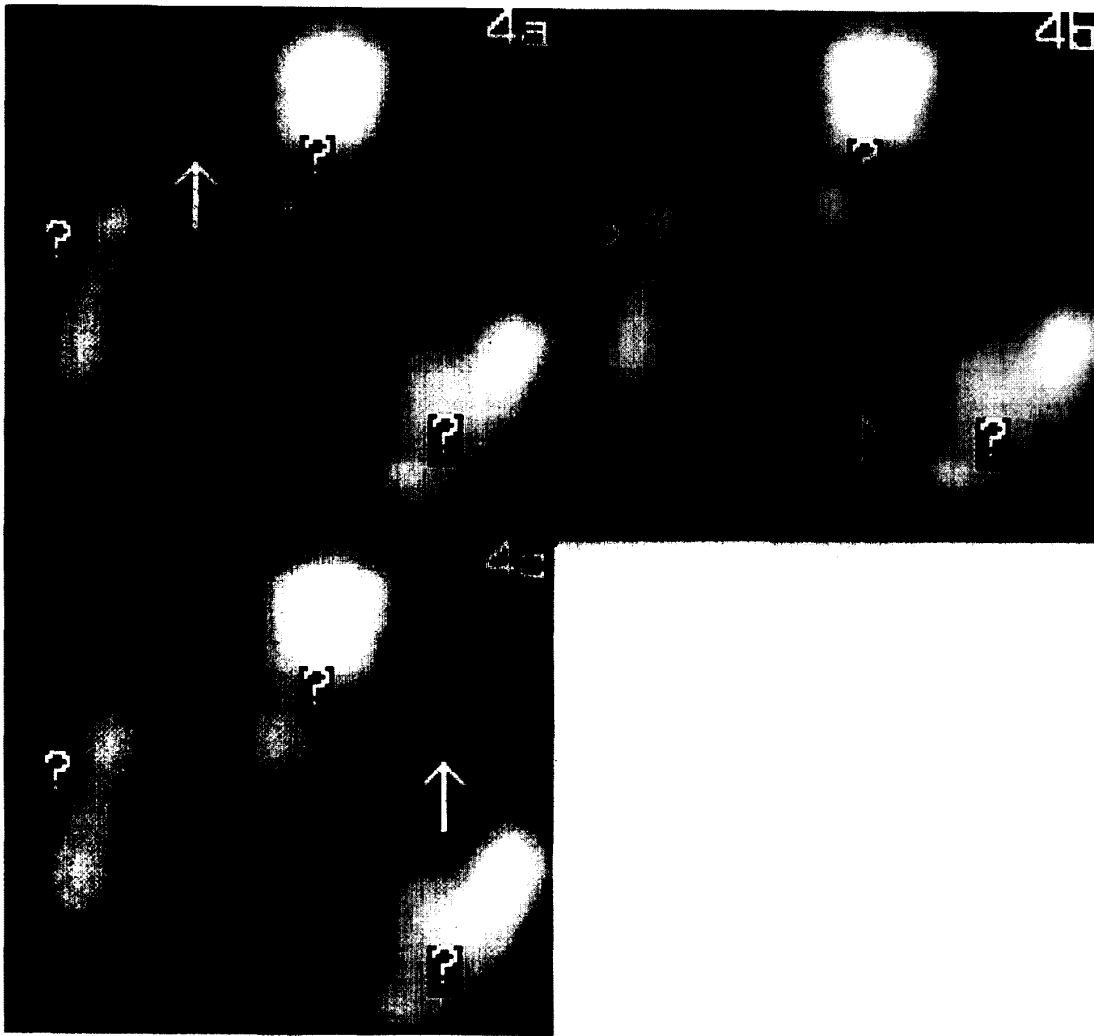
Fig. 4. (a) Test image of Fig. 3(a) correlated with template of top-left face in the image; (b) Test image of Fig. 3(a) correlated with template of center face in the image; (c) Test image of Fig. 3(a) correlated with template of top-right face in the image. Note the weak peaks (marked by the arrows) and substantial spurious responses (marked by '?').

to the design noise parameter and we find that even coarse estimates produce sharp peaks and low spurious response.

It is important to note here that the assumption on the circulant nature of the data is only for mathematical convenience. In practice the signal is assumed to be appropriately zero padded to make the convolution (or correlation) linear. The frequency domain results of Section III are still valid even without assuming circularity or finite extent.

The computational complexity of the matching scheme is simply one of a linear filtering process, just as in matched filtering or direct correlation. However, the Expansion filter has to be designed (only once), given the templates to be matched and the corresponding desired responses. We use Fast Fourier Transforms and perform our filter design in the frequency domain as outlined in Section III. Thus, for designing a 1D M-point filter for the given N templates (assuming sufficient

storage and neglecting all additions), it is required to compute N FFTs (or $NM \log_2 M$ real multiplications) and compute N corresponding power spectrums (or NM multiplications). Computing the weighted power spectrum, i.e. the denominator in (25), requires an additional NM multiplications. The symmetric constraint matrix requires computation of $(N^2 - N)/2$ cross power spectrums (or $M(N^2 - N)/2$ multiplications), and $M(N^2 + N)/2$ divisions (assumed as equal complexity to a multiplication). Note that complex multiplications are assumed as two real multiplications and since all the templates are real, only $M/2$ points are effectively involved in the frequency domain computation. Assuming that the matrix inversion can be neglected (since the number of templates N is typically small), the realization of the desired filter in the frequency domain requires another MN multiplications and M divisions. Thus the overall computational cost of designing
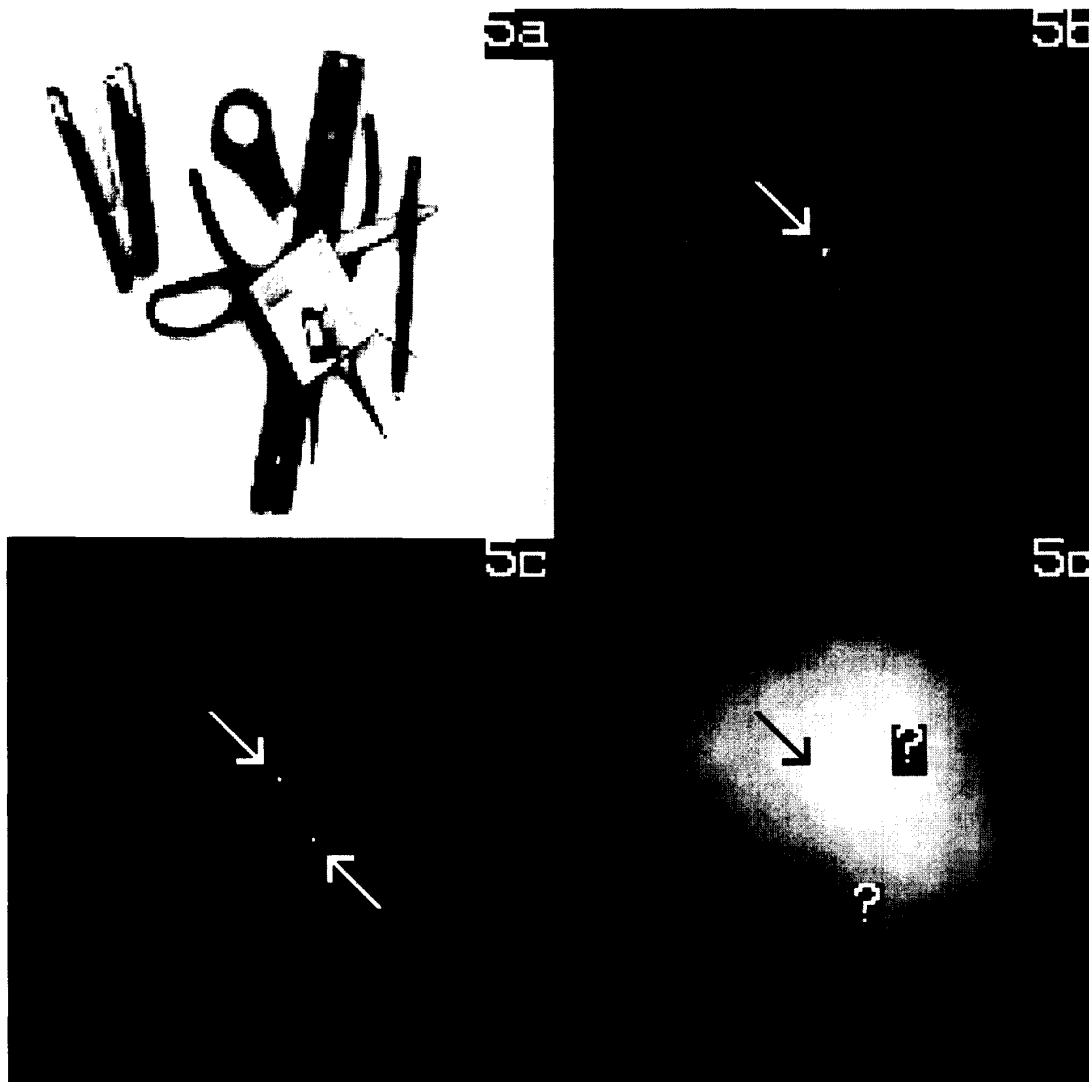
Fig. 5. (a) Test image with occluded scissors and pliers; (b) MACE response (zero noise Expansion filter) for matching scissors and pliers with unit response. Only the scissors are detected; (c) Corresponding Expansion filter response (assumed SNR = 18 dB). Both the scissors and pliers are detected and there is minimal off-center response; (d) SDF response (infinite noise Expansion filter) has only one broad peak for the scissors and substantial spurious response.

the filter is approximately $M(N^2 + 2N + N\log_2 M + 1)$ multiplications. The computational complexity required for the MACE, MVSDF and SDF designs is of the same order as for designing the Expansion filter.

Another consideration in designing such a multiple template recognition filter is its "capacity", or the number of templates that can be matched using a single filter. Obviously, as N becomes indiscriminately large, the filter will not be able to recognize all the N patterns successfully. We have observed that the degradation of the matching peak with N is also dependent on the nature of the patterns itself and the desired responses specified. Eq. (20) guarantees the value of this desired response at the template-center but as N increases the sharpness of the peak degrades. When template discrimi-

nation is required (as in Figs. 1 and 2), the filter performance degenerates with increasing number of templates. For large N, the template center may be overshadowed by larger spurious peaks or even cease to be a local extremum. However, if generic recognition is required and the templates themselves have some common features (as in Fig. 3), then large N does not typically degrade performance, but makes the expansion filter more generic.

APPENDIX A. THE MACE, MVSDF AND SDF APPROACHES AS LIMITING CASES OF THE EXPANSION FILTER

In this appendix we analyze the formulation of the Minimum Average Correlation Energy (MACE) filter [12], the Minimum Variance Synthetic Discriminant Function (MVSDF) [8] and

SDF [8] approaches and show these to be special cases of our formulation. The notation used is consistent with our previous formulation. However, wherever possible, we conform to the notation used by Mahalanobis et al. in their paper [12].

We use the vectors $\mathbf{X}_i$ to represent the DFTs of the templates Ψ_i, i.e., $\mathbf{X}_i = [\mathbf{W}]^{-1}\Psi_i$. As before $\mathbf{X}_i(\omega)$ represents the ωth frequency component of this DFT. The matrix $[\mathbf{X}]$ is defined as

$$[\mathbf{X}] = [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_N]; [\mathbf{X}]_{ij} = \mathbf{X}_j(i) \qquad (A1)$$

As before, the power spectrum of each of the templates is given by a diagonal matrix $[\mathbf{D}_{\psi\psi i}]$, i.e. $[\mathbf{D}_{\psi\psi i}]_{kk} = |\mathbf{X}_i(k)|^2 = S_{\psi\psi i}(k)$.

The most general solution for the MACE filter [12] is a vector $\mathbf{H}$ (which is the DFT of the filter's impulse response) given by

$$\mathbf{H} = [\mathbf{B}][\mathbf{X}][[\overline{\mathbf{X}}]^T[\mathbf{B}][\mathbf{X}]]^{-1}\mathbf{U} \qquad (A2)$$

where $\mathbf{U}$ defines the user constraints as before, and $[\mathbf{B}]$ is a given matrix of appropriate dimensions.

A. The MACE Filter Solution

To obtain the MACE filter [12], in (A2) the matrix $[\mathbf{B}]$ is chosen to be $[\mathbf{D}]^{-1}$, where $[\mathbf{D}]$ is real and **diagonal**, given by

$$[\mathbf{D}] = \sum_{i=1}^{N} \alpha_i[\mathbf{D}_{\psi\psi i}] \qquad (A3)$$

i.e. the weighted sum of the power spectrum matrices of the templates. Evaluating each part of (A2) separately, we can write

$$[[\overline{\mathbf{X}}]^T[\mathbf{D}]^{-1}[\mathbf{X}]]_{ij} = \sum_k \overline{\mathbf{X}}_i(k)[\mathbf{D}]_{kk}^{-1}\mathbf{X}_j(k)$$
$$= \sum_k \left(\frac{\overline{\mathbf{X}}_i(k)\mathbf{X}_j(k)}{\sum_{n=1}^{N} \alpha_n S_{\psi\psi n}(k)} \right) \qquad (A4)$$

Note that (A4) corresponds exactly to the definition of matrix $[\mathbf{A}]$ in (28) when the noise power is zero. Thus, we can write out

$$\tilde{\Xi} = [[\overline{\mathbf{X}}]^T[\mathbf{D}]^{-1}[\mathbf{X}]]^{-1}\mathbf{U} \qquad (A5)$$

to yield exactly the same set of $\tilde{\xi}_i$ as from (21) with zero noise. Noting that

$$[\mathbf{X}][[\overline{\mathbf{X}}]^T[\mathbf{D}]^{-1}[\mathbf{X}]]^{-1}\mathbf{U} = [\mathbf{X}]\tilde{\Xi} = \sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i \qquad (A6)$$

we can now write out the MACE filter solution from (A2) explicitly as

$$\mathbf{H} = [\mathbf{D}]^{-1}\sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i \qquad (A7)$$

or

$$\mathbf{H}(k) = \frac{\sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i(k)}{\sum_{i=1}^{N} \alpha_i S_{\psi\psi i}(k)} \qquad (A8)$$

which is exactly the expansion filter Θ of (25) with zero noise power term.

B. The Minimum Variance SDF Approach

We define the noise power spectrum $S_{\lambda\lambda}(\omega)$ to be an amplitude scaled version of a normalized noise power spectrum $S_{nn}(\omega)$

$$S_{\lambda\lambda}(\omega) = c_n S_{nn}(\omega); \quad 0 \le S_{nn}(\omega) \le 1;$$
$$c_n = \max_{\omega}(S_{\lambda\lambda}(\omega)) \qquad (A9)$$

where c_n is the amplitude scaling constant. Note that if $c_n \to \infty$, the noise power goes to infinity as well.

To obtain the MVSDF solution [12], in (A2) we choose the matrix $[\mathbf{B}]$ to be $[\mathbf{K}]^{-1}$ where $\mathbf{K}$ is the normalized noise power spectrum matrix, i.e.

$$[\mathbf{K}] = \frac{1}{c_n}[\mathbf{D}_{\lambda\lambda}] \quad ; \quad [\mathbf{K}]_{kk} = S_{nn}(k) \qquad (A10)$$

Proceeding similarly as in (A1), we can write

$$[[\overline{\mathbf{X}}]^T[\mathbf{K}]^{-1}[\mathbf{X}]]_{ij} = \sum_k \overline{\mathbf{X}}_i(k)[\mathbf{K}]_{kk}^{-1}\mathbf{X}_j(k)$$
$$= \sum_k \left(\frac{\overline{\mathbf{X}}_i(k)\mathbf{X}_j(k)}{S_{nn}(k)} \right) \qquad (A11)$$

Using the constant c_n, we can normalize the matrix $[\mathbf{A}]$ of (28) as

$$c_n[\mathbf{A}]_{ij} = \sum_\omega \left(\frac{\overline{\mathbf{X}}_i(\omega)\mathbf{X}_j(\omega)}{c_n^{-1}\sum_{k=1}^{N} \alpha_k S_{\psi\psi k}(\omega) + c_n^{-1}S_{\lambda\lambda}(\omega)} \right) \qquad (A12)$$

When the constant c_n approaches infinity, the noise power also tends to infinity and thus in the limit, we obtain a normalized matrix $[\mathbf{A}_n]$ given by

$$[\mathbf{A}_n]_{ij} = \lim_{c_n \to \infty} c_n[\mathbf{A}]_{ij} = \sum_\omega \left(\frac{\overline{\mathbf{X}}_i(\omega)\mathbf{X}_j(\omega)}{S_{nn}(\omega)} \right) \qquad (A13)$$

which is exactly the same matrix as used in (A11), i.e.

$$[\mathbf{A}_n] = \lim_{c_n \to \infty} c_n[\mathbf{A}] = [[\overline{\mathbf{X}}]^T[\mathbf{K}]^{-1}[\mathbf{X}]] \qquad (A14)$$

Therefore, in the limit, (21) becomes

$$\tilde{\Xi} = [\mathbf{A}_n]^{-1}\mathbf{U} \qquad (A15)$$

which is the same set of $\tilde{\xi}_i$ as obtained from (A11)

$$[\mathbf{X}][[\overline{\mathbf{X}}]^T[\mathbf{K}]^{-1}[\mathbf{X}]]^{-1}\mathbf{U} = [\mathbf{X}]\tilde{\Xi} = \sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i \qquad (A16)$$

Since, we have proven that the weighting constants $\tilde{\xi}_i$ are identical for both the MVSDF and the expansion filter with infinite noise, it only remains to be shown that the final result is also identical. The MVSDF solution can be computed from Eqs. (A2) and (A16) as

$$\mathbf{H} = [\mathbf{K}]^{-1}\sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i \qquad (A17)$$

or

$$\mathbf{H}(k) = \frac{\sum_{i=1}^{N} \tilde{\xi}_i\mathbf{X}_i(k)}{S_{nn}(k)} \qquad (A18)$$

Normalizing the expansion filter of (25) yields

$$c_n \Theta(\omega) = \frac{\sum_{j=1}^{N} \tilde{\xi}_j \mathbf{X}_j(\omega)}{c_n^{-1} \sum_{i=1}^{N} \alpha_i S_{\psi\psi i}(\omega) + c_n^{-1} S_{\lambda\lambda}(\omega)} \quad \text{(A19)}$$

As before, if $c_n \to \infty$, the noise power is increased to infinity and we obtain the normalized expansion filter as

$$\Theta_n(\omega) = \lim_{c_n \to \infty} c_n \Theta(\omega) = \frac{\sum_{j=1}^{N} \tilde{\xi}_j \mathbf{X}_j(\omega)}{S_{nn}(\omega)} \quad \text{(A20)}$$

which is exactly the MVSDF in (A18).

Thus as the noise power tends to infinity, the expansion filter filter approaches the MVSDF result, with only the difference of an amplitude scaling constant.

C. The SDF Approach

In (A2), if the matrix $[\mathbf{B}]$ is chosen to be the identity matrix $[\mathbf{I}]$, then the general solution reduces to the SDF solution. This is equivalent to the MVSDF result with a white noise spectrum. A similar analysis as performed above yields that both, the SDF and the expansion filter with infinite white noise term, result in the following filter

$$\Theta_n(\omega) = \mathbf{H}(\omega) = c_{SDF} \sum_{j=1}^{N} \tilde{\xi}_j \mathbf{X}_j(\omega) \quad \text{(A21)}$$

where c_{SDF} is an amplitude scaling constant.

APPENDIX B. EXPANSION MATCHING: THE RELATION BETWEEN NON-ORTHOGONAL SIGNAL EXPANSION, DISCRIMINATIVE SNR OPTIMIZATION AND RESTORATION TECHNIQUES

A. Non-Orthogonal Signal Expansion

The method of matching by expansion [1]–[4] is based on expanding the signal with respect to basis functions (BFs) that are all shifted versions of the template. Such an expansion is feasible if these BFs are linearly independent and complete [15], [16]. This method cannot be implemented by orthogonal ba· s functions, since practically, the shifted templates are mutually non-orthogonal.

The general derivation of expansion is as follows: Suppose one wishes to estimate a d-dimensional discrete signal $s(x, y, \ldots)$ by a set of basis functions $\{\psi^{(i)}(x, y, \ldots)\}$ with the sum

$$\hat{s}(x, y, \ldots) = \sum_{i=1}^{m} c_i \psi^{(i)}(x, y, \ldots); x, y, \cdots = 1 \ldots M \quad \text{(B1)}$$

where $\{c_i\}$ are the coefficients of the expansion. The d-dimensional signal $s(x, y, \ldots)$ is translated into a vector $\mathbf{S}$, where the dimensionality of the vector $\mathbf{S}$ is M^d, and likewise the basis functions $\{\psi^{(i)}(x, y, \ldots)\}$ are translated to the basis vector set $\{\Psi^{(i)}\}$, and the approximation $\hat{s}$ is translated to $\hat{\mathbf{S}}$. The squared error D to be minimized is given by the inner product

$$D = \langle (\mathbf{S} - \hat{\mathbf{S}}), (\mathbf{S} - \hat{\mathbf{S}}) \rangle$$
$$= \left\langle \left(\mathbf{S} - \sum_{i=1}^{m} c_i \Psi^{(i)} \right), \left(\mathbf{S} - \sum_{i=1}^{m} c_i \Psi^{(i)} \right) \right\rangle \quad \text{(B2)}$$

According to the orthogonality prin iple [15], [16], the MSE approximation error $\mathbf{S} - \hat{\mathbf{S}}$ is orthogonal to the BFs. This could be formulated by the following set of m equations

$$\langle (\mathbf{S} - \hat{\mathbf{S}}), \Psi^{(i)} \rangle = 0; \quad i = 1 \cdots m; \quad \text{(B3)}$$

This leads to a set of equations for deterministic signals which is similar to equations formulated for random variables [15]

$$[\mathbf{R}_{\psi\psi}]\mathbf{C} = [\Psi]\mathbf{S} \quad \text{(B4)}$$

where the matrix $[\mathbf{R}_{\psi\psi}]$ is $[\mathbf{R}_{\psi\psi}]_{ij} = \langle \Psi^{(i)}, \Psi^{(j)} \rangle$, $\mathbf{C} = [c_1, c_2 \cdots c_{M^2}]^t$, $[\Psi] = [\Psi^{(1)} \Psi^{(2)} \ldots \Psi^{(M^2)}]^t$ and $\mathbf{S}$ is the $M^2 \times 1$ vector formulation of the image $s(x, y)$ If the set of BFs $\{\Psi^{(i)}\}$ is linearly independent, then the matrix $[\mathbf{R}_{\psi\psi}]$ is positive definite and the equations (B4) yield a unique solution for the coefficients c_i. Direct solution of this equation is computationally prohibitive. However, we have found an efficient implementation of (B4) using restoration techniques.

B. Discriminative SNR Optimization

The problem of template matching can be formulated as follows: Given a $M \times M$ discrete image $s(x, y)$, it is desired to find a sub-image which is similar to a two-dimensional template $\psi(x, y)$ and therefore matches it. Let the image $s(x, y)$ contain the template at a certain location (x_0, y_0) plus additional noise $\lambda(x, y)$. Let the filter $h(x, y)$ be correlated with the signal to yield a response $z(x, y)$

$$z(x, y) = s(x, y) \bigcirc h(x, y)$$
$$= \psi(x, y) \bigcirc h(x, y) + \lambda(x, y) \bigcirc h(x, y)$$
$$= z_\psi(x, y) + z_\lambda(x, y) \quad \text{(B5)}$$

where $\bigcirc$ denotes discrete correlation. $z_\psi(x, y)$ and $z_\lambda(x, y)$ denote the response to the template and the noise respectively. The signal-to-noise ratio (SNR) of the matched filtering (or direct correlation) is defined [15], [16] as the ratio between the maximal response at the location (x_0, y_0) and the variance of the noisy response

$$\text{SNR} = 10 \log \frac{[z_\psi(x_0, y_0)]^2}{\frac{1}{M^2} \sum_i \sum_j [z_\lambda(i, j)]^2} \quad \text{(B6)}$$

The matched filtering (or direct correlation) formulation has a **significant drawback. It overlooks the response** $z_\psi(x, y)$ **at locations other than** (x_0, y_0). These responses are also **unwanted and should also be considered as "noise".** Usually matched filtering generates broad peaks and z_ψ is quite substantial in the neighborhood of (x_0, y_0) and should not be neglected!!

According to the matched filtering approach, a filter that provides a very broad peak is equivalent to a filter that provides a sharp peak with the same amplitude! Thus, a better SNR definition which provides a genuine measure of matching quality, **should be an SNR which defines as "noise", any response that is not located at the sub-image's center.** According to this definition, the new **Discriminative DSNR**

that should be used is

$$\text{DSNR} = 10\log \frac{[z(x_0, y_0)]^2}{\displaystyle\sum_{(i,j)\neq(x_0,y_0)}\sum [z(i,j)]^2} \qquad (B7)$$

Now, suppose one wants to find a filter Θ (in lexicographically ordered vector form), that when correlated with the signal, maximizes the DSNR. The corresponding matrix-vector notation is

$$[\mathbf{S}]\Theta = \mathbf{C} \qquad (B8)$$

where $[\mathbf{S}]$ is a block circulant correlation matrix. Defining the DSNR in accordance with (B7) as the square ratio of the l-th component of $\mathbf{C}$ (location (x_0, y_0)) vs. the remainder

$$\text{DSNR} = 10\log \frac{c^2(l)}{\sum_{j\neq l} c^2(j)} \qquad (B9)$$

Maximizing the DSNR with respect to Θ, we find [1]–[4] that

$$\Theta = \frac{1}{k}[[\mathbf{S}]^T[\mathbf{S}]]^{-1}\mathbf{S}_l = \frac{1}{k}[\mathbf{R}_{ss}]^{-1}\mathbf{S}^{(l)} \qquad (B10)$$

where $k = \Theta^T \mathbf{S}^{(l)}/\Theta^T[\mathbf{S}]^T[\mathbf{S}]\Theta$ is only a scaling factor (a scalar) and $[\mathbf{R}_{ss}]$ is the autocorrelation matrix of the signal which is equivalent to $[\mathbf{R}_{\psi\psi}]$ for the templates. Eq. (B10) is another form of the Yule-Walker equation where the BFs are circulant and $\mathbf{S}_{(l)}$ corresponds to $\Psi^{(l)}$.

C. Restoration Techniques

We have shown [1]–[4] that restoration is a special case of the more generalized expansion. In particular, the expansion matching filter that maximizes the DSNR yields exactly the same result as the Wiener filter. This result is quite logical, since both approaches are based upon the minimization of squared error. Using 1D signals and matrix-vector notation, we can introduce an additive noise term $\Lambda^{(l)}$ to the signal, i.e.,

$$\mathbf{S}_{(l)} = \Psi^{(l)} + \Lambda_{(l)} \qquad (B11)$$

where $\Psi^{(l)}$ is the BF corresponding to the template shifted to location l. The circulant correlation matrix $[\mathbf{S}]$ can now be split into two circulant correlation matrices $[\Psi]$ and $[\Lambda]$ as:

$$[\mathbf{S}] = [\Psi] + [\Lambda] \qquad (B12)$$

Optimization of the DSNR now leads to a filter Θ given by:

$$\Theta = \frac{1}{k}[E\{([\Psi] + [\Lambda])^T([\Psi] + [\Lambda])\}]^{-1}E\{(\Psi^{(l)} + \Lambda^{(l)})\} \qquad (B13)$$

If we assume that $\Lambda^{(l)}$ is of zero mean and stationary, this simplifies to:

$$\Theta = \frac{1}{k}[[\mathbf{R}_{\psi\psi}] + [\mathbf{R}_{\lambda\lambda}]]^{-1}\Psi^{(l)} \qquad (B14)$$

where the matrices $[\mathbf{R}_{\psi\psi}]$ and $[\mathbf{R}_{\lambda\lambda}]$ are defined as the autocorrelation matrices of the template and the noise respectively. The result of (B14) is simply the Wiener filter revisited from another perspective and establishes that **expansion matching and restoration are the same operation** in the case of a dense set of self-similar BFs. Eq. (B14) is also identical to (19) for the case of a single template, i.e., $N = 1$.

REFERENCES

[1] J. Ben-Arie and K. R. Rao, "A Novel approach for template matching by non-orthogonal image expansion," *IEEE Trans. on Circuits &Systems for Video Technology*, vol. 3, no. 1, pp. 71–84, Feb. 1993.

[2] J. Ben-Arie and K. R. Rao, "Image expansion by non-orthogonal wavelets for optimal template matching," in *Proc. of 11th IAPR Intl. Conf. on Pattern Recognition*, Hague, Netherlands, pp. 650–654 Sep. 1992.

[3] J. Ben-Arie and K. R. Rao, "Lattice architectures for signal expansion by Gaussian set wavelets with application to recognition," in *Proceedings, IEEE Intl. Symp. on Circuits and Systems*, San Diego, pp. 955–959, May 1992.

[4] K. R. Rao and J. Ben-Arie, "Restoration with equivalence to non-orthogonal image expansion for feature extraction and edge detection," in *Proc. of 1992 SPIE Conf. on Visual Commn. and Image Proc.*, Boston, pp. 187–197, Nov. 1992.

[5] J. Ben-Arie, "Multi-dimensional linear lattice for Fourier and Gabor transforms, multiple-scale Gaussian filtering and edge detection," in the Book *Neural Networks for Human and Machine Perception*, (H. Wechsler, Ed.), Academic Press, Ch. II3, pp. 231–252, 1991.

[6] J. Ben-Arie and K. R. Rao, "Signal representation by generalized non-orthogonal Gaussian wavelet groups using lattice networks," in *Proceedings, IEEE Intl. Joint Conf. on Neural Networks*, Singapore, pp. 958–973, Nov. 1991.

[7] D. Casasent, "Unified synthetic discriminant function computational formulation," *Applied Optics*, vol. 23, no. 10, May 1984, pp. 1620–1627.

[8] B. V. K. Vijaya Kumar, "Minimum variance synthetic discriminant functions," *J. Optical Soc. of America*, A 3, 1986, p. 1579.

[9] H. Mostafavi and F. W. Smith, Image Correlation with Geometric Distortion Part I: Acquisition Performance, *IEEE Trans. on A. E. S.*, vol. AES-14, no. 3, pp. 487–493, May 1978.

[10] R. R. Kallman, "Construction of low noise optical correlation filters," *Applied Optics*, vol. 25, no. 25, p. 1579, 1986.

[11] D. Casasent and W. T. Chang, "Correlation synthetic discriminant functions," *Applied Optics*, vol. 25, p. 2343, 1986.

[12] A. Mahalanobis, B. V. K. Vijaya Kumar and David Casasent, "Minimum average correlation energy filters," *Applied Optics*, vol. 26, no. 17, pp. 3633–3640, Sep. 1987.

[13] B. V. K. Vijaya Kumar and A. Mahalanobis, "Alternative interpretation for minimum variance synthetic discriminant functions," *Applied Optics*, vol. 25, p. 2484, 1986.

[14] B. V. K. Vijaya Kumar, C. D. Hendrix and D. W. Carlson, "Tradeoffs in the design of correlation filters," *Optical Pattern Recognition*, Eds. J. L. Horner, B. Javidi, vol. CR 40, SPIE Optical Engineering Press, 1992.

[15] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1989.

[16] A. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall 1989.

Jezekiel Ben-Arie (M'91) received the B.Sc. and M.Sc. degrees in Electrical Engineering from the Technion, Israel Institute of Technology. He received the D.Sc. also from the Technion in 1987. During the period 1971–1989 he was the Chief Engineer and Manager of the Image Processing and Flight Control Labs in the Technion. In this capacity, he conducted many research and development projects in the areas of Computer Vision, real-time flight simulation and man-machine control systems. In 1989, he joined the faculty of the Electrical and Computer Engineering department at the Illinois Institute of Technology, Chicago. Currently, he also heads the Computer Vision and Neural Networks laboratory at I.I.T. Prof. Ben-Arie's current research interests are in neural networks for signal processing, non-orthogonal expansions for image representation, image understanding, visual and sound recognition, auditory localization and probabilistic models for vision. In these areas, Prof. Ben-Arie has had more than 60 technical publications.

K. Raghunath Rao received the B.E. degree in Instrumentation Engineering from Bangalore University, India, in 1990. He was a researcher in the Computer Vision and Artificial Intelligence Laboratory in the Indian Institute of Science, Bangalore, India. He received the M.S. degree in Electrical and Computer Engineering from the Illinois Institute of Technology, Chicago, in 1991. He is currently pursuing his Ph.D. under Prof. J. Ben-Arie in the Electrical and Computer Engineering department at I.I.T. He has more than 35 technical publications in the areas of signal representation, signal compression, image processing and image understanding. His research interests include machine vision, image understanding, image/signal processing and neural networks.