Eigenface-Domain Super-Resolution for Face Recognition

Bahadir K. Gunturk, *Student Member, IEEE*, Aziz U. Batur, *Student Member, IEEE*,
Yucel Altunbasak, *Senior Member, IEEE*, Monson H. Hayes, III, *Fellow, IEEE*, and Russell M. Mersereau, *Fellow, IEEE*

***Abstract*—Face images that are captured by surveillance cameras usually have a very low resolution, which significantly limits the performance of face recognition systems. In the past, super-resolution techniques have been proposed to increase the resolution by combining information from multiple images. These techniques use super-resolution as a preprocessing step to obtain a high-resolution image that is later passed to a face recognition system. Considering that most state-of-the-art face recognition systems use an initial dimensionality reduction method, we propose to transfer the super-resolution reconstruction from pixel domain to a lower dimensional face space. Such an approach has the advantage of a significant decrease in the computational complexity of the super-resolution reconstruction. The reconstruction algorithm no longer tries to obtain a visually improved high-quality image, but instead constructs the information required by the recognition system directly in the low dimensional domain without any unnecessary overhead. In addition, we show that face-space super-resolution is more robust to registration errors and noise than pixel-domain super-resolution because of the addition of model-based constraints.**

***Index Terms*—Dynamic range extension, face recognition, multiframe reconstruction, super-resolution.**

I. Introduction

THE performance of existing face recognition systems decreases significantly if the resolution of the face image falls below a certain level. This is especially critical in surveillance imagery where often only a low-resolution video sequence of the face is available. If these low-resolution images are passed to a face recognition system, the performance is usually unacceptable. Therefore, super-resolution techniques have been proposed for face recognition that attempt to obtain a high-resolution face image by combining the information from multiple low-resolution images [1]–[4]. In general, super-resolution algorithms try to regularize the ill-posedness of the problem using prior knowledge about the solution, such as smoothness or positivity [5]–[8]. Recently, researchers have proposed algorithms that attempt to use model-based constraints in regularization. While [1] demonstrates how super-resolution (without model-based priors) can improve the face recognition rate, [2]–[4] provide super-resolution algorithms that use face-specific constraints for regularization.

All these systems propose super-resolution as a separate preprocessing block in front of a face recognition system. In other words, their main goal is to construct a high-resolution, visually improved face image that can later be passed to a face recognition system for improved performance. This is perfectly valid as long as computational complexity is not an issue. However, in a real-time surveillance scenario where the super-resolution algorithm is expected to work on continuous video streams, computational complexity is usually a very critical issue. In this paper, we propose an efficient super-resolution method for face recognition that transfers the super-resolution problem from the pixel domain to a low dimensional face space. This is based on the observation that nearly all state-of-the-art face recognition systems use some kind of front-end dimensionality reduction, and that a lot of redundant information generated by the preprocessing super-resolution algorithm is not used by the face recognition block. Hence, we perform the super-resolution reconstruction in the low-dimensional framework so that only the necessary information is reconstructed. In addition, we show that face-space super-resolution is more robust to registration errors and noise than the pixel-domain super-resolution because of the addition of model-based constraints.

There are two important sources of noise in this problem. One is the observation noise that results from the imaging system. The other is the representation error, which is a result of the dimensionality reduction. We derive the statistics of these noise processes *for the low-dimensional face space* by using examples from the human face image class. Substitution of this model-based information into the algorithm provides a higher robustness to noise. We test our system on both real and synthetic video sequences.

Currently, by far the most popular dimensionality reduction technique in face recognition is to use subspace projections based on the Karhunen–Loeve Transform (KLT). This type of dimensionality reduction has been central to the development of face recognition algorithms for the last ten years. We propose to use a similar KLT-based dimensionality reduction technique to decrease the computational cost of the super-resolution algorithm by transforming it from a problem in the pixel domain to a problem in the lower-dimensional subspace, which is called the face space.

In Section II, we briefly review the KLT-based dimensionality reduction method for face recognition. Then, in Section III, we formulate the super-resolution problem in the low-dimensional framework. Section IV details the reconstruction algorithm, and Section V provides experimental results addressing

Manuscript received March 21, 2002; revised November 6, 2202. This work was supported in part by the Office of Naval Research (ONR) under Award N00014-01-1-0619 and by the National Science Foundation under Award CCR-0113681. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Philippe Salembier.

The authors are with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250 USA (e-mail: bahadir@ece.gatech.edu; batur@ece.gatech.edu; yucel@ece.gatech.edu; mhh3@ece.gatech.edu; rmm@ece.gatech.edu).

Digital Object Identifier 10.1109/TIP.2003.811513

1057-7149/03$17.00 © 2003 IEEE

several issues, such as sensitivity to noise and motion estimation errors. Conclusions are given in Section VI.

II. Dimensionality Reduction for Face Recognition

KLT-based dimensionality reduction for face images was first proposed by Sirovich and Kirby [9]. They showed that face images could be represented efficiently by projecting them onto a low-dimensional linear subspace that is computed using the KLT. Later, Turk and Pentland demonstrated that this subspace representation could be used to implement a very efficient and successful face recognition system [10]. Since then, eigenface-based dimensionality reduction has been used widely in face recognition.

Mathematically, the eigenface method tries to represent a face image as a linear combination of orthonormal vectors, called eigenfaces. These eigenfaces are obtained by finding the eigenvectors of the covariance matrix of the training face image set. Let $\mathbf{I}_1, \mathbf{I}_2, \ldots, \mathbf{I}_K$ be a set of K face images, each ordered lexicographically. The eigenvectors of the matrix

$$\mathbf{C} = \sum_{i=1}^{K} \mathbf{I}_i \mathbf{I}_i^T \tag{1}$$

that correspond to the largest L eigenvalues span a linear subspace that can reconstruct the face images with minimum reconstruction error in the least squares sense. This L-dimensional subspace is called the face space. Assuming $\mathbf{x}$ is a lexicographically ordered face image and $\mathbf{\Phi}$ is the matrix that contains the eigenfaces as its columns, we can write

$$\mathbf{x} = \mathbf{\Phi a} + \mathbf{e_x} \tag{2}$$

where $\mathbf{a}$ is the feature vector that represents the face, and $\mathbf{e_x}$ is the subspace representation error for the face image. As a larger training data set is used and the dimensionality of the face space is increased, the representation error $\mathbf{e_x}$ gets smaller. Letting

$$\mathbf{a} \triangleq [a_1 \quad a_2 \quad \cdots \quad a_L]^T \tag{3}$$

be the feature vector, and

$$\mathbf{\Phi} \triangleq [\phi_1 \quad \phi_2 \quad \cdots \quad \phi_L] \tag{4}$$

be the matrix where $\phi_1, \ldots, \phi_L$ are the eigenface vectors, a_i is computed as follows:

$$a_i = \phi_i^T \mathbf{x}, \qquad \text{for } i = 1, \ldots, L. \tag{5}$$

III. Super-Resolution in the Face Subspace

In this section, we formulate the super-resolution problem in the low-dimensional face subspace. In such a formulation, the observations are inaccurate feature vectors of a subject, and the reconstruction algorithm estimates the true feature vector. We start with the observation model for pixel-domain super-resolution, and then derive the observation model for face-space super-resolution using the eigenface representation. In pixel-domain super-resolution, the observations are low-resolution images that are related to a high-resolution image by a linear mapping. By ordering images lexicographically, such a relation can be written in matrix-vector notation as follows:

$$\mathbf{y}^{(i)} = \mathbf{H}^{(i)}\mathbf{x} + \mathbf{n}^{(i)}, \qquad \text{for } i = 1, \ldots, M \tag{6}$$

where $\mathbf{x}$ is the unknown high-resolution image, $\mathbf{y}^{(i)}$ is the ith low-resolution image observation, $\mathbf{H}^{(i)}$ is a linear operator that incorporates the motion, blurring, and downsampling processes, $\mathbf{n}^{(i)}$ is the noise vector, and M is the number of observations. Assuming that s is the downsampling factor $(0 < s < 1)$, and that the high-resolution image is of dimension $N \times N$; $\mathbf{y}^{(i)}$, $\mathbf{H}^{(i)}$, $\mathbf{x}$, and $\mathbf{n}^{(i)}$ have dimensions $s^2N^2 \times 1$, $s^2N^2 \times N^2$, $N^2 \times 1$, and $s^2N^2 \times 1$, respectively. The matrix $\mathbf{H}^{(i)}$ can be written as

$$\mathbf{H}^{(i)} = \mathbf{D}^{(i)}\mathbf{B}^{(i)}\mathbf{W}^{(i)} \tag{7}$$

where $\mathbf{D}^{(i)}$, $\mathbf{B}^{(i)}$, and $\mathbf{W}^{(i)}$ are the downsampling, blurring, and motion warping matrices, respectively. Details of such modeling can be found in [5], [6], and [11], and we will not elaborate on it in this paper. [Note that it is also possible to include an upsampling matrix in $\mathbf{H}^{(i)}$ that will make the sizes of $\mathbf{y}^{(i)}$ and $\mathbf{x}$ equal.]

The images $\mathbf{x}$ and $\mathbf{y}^{(i)}$ have components that lie in and are orthogonal to the face space. Only the components that lie in the face space are necessary in recognition. We will now derive the observation model for the reconstruction of the components that lie in the face space. The formulation and reconstruction algorithm will not neglect the spatial-domain observation noise and the subspace representation error, which is initially orthogonal to the face space but which has an effect during the imaging process. We start by writing the face-space representation

$$\mathbf{x} = \mathbf{\Phi a} + \mathbf{e_x} \tag{8}$$

$$\mathbf{y}^{(i)} = \mathbf{\Psi}\hat{\mathbf{a}}^{(i)} + \mathbf{e}_\mathbf{y}^{(i)}, \qquad \text{for } i = 1, \ldots, M \tag{9}$$

where $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are $N^2 \times L$ and $s^2N^2 \times L$ matrices that contain the eigenfaces in their columns, $\hat{\mathbf{a}}^{(i)}$ is the $L \times 1$ dimensional feature vector that is associated with the ith observation, and $\mathbf{e_x}$ and $\mathbf{e}_\mathbf{y}^{(i)}$ are the $N^2 \times 1$ and $s^2N^2 \times 1$ representation error vectors. Note that we have two different eigenvector bases, $\mathbf{\Phi}$ and $\mathbf{\Psi}$, corresponding to high and low resolution face images, respectively. [If we had included an upsampling matrix in $\mathbf{H}^{(i)}$, then we could use the same basis matrix.]

We substitute (8) and (9) into (6) to obtain

$$\mathbf{\Psi}\hat{\mathbf{a}}^{(i)} + \mathbf{e}_\mathbf{y}^{(i)} = \mathbf{H}^{(i)}\mathbf{\Phi a} + \mathbf{H}^{(i)}\mathbf{e_x} + \mathbf{n}^{(i)}. \tag{10}$$

Now, we will project (10) into the lower-dimensional face space using the fact that the representations errors $\mathbf{e}_\mathbf{y}^{(i)}$ are orthogonal to the face space $\mathbf{\Psi}$. Using

$$\mathbf{\Psi}^T\mathbf{e}_\mathbf{y}^{(i)} = \mathbf{0}, \qquad \text{for } i = 1, \ldots, M \tag{11}$$

and

$$\mathbf{\Psi}^T\mathbf{\Psi} = \mathbf{I} \tag{12}$$

and multiplying both sides of (10) by $\mathbf{\Psi}^T$ on the left, we obtain

$$\hat{\mathbf{a}}^{(i)} = \mathbf{\Psi}^T\mathbf{H}^{(i)}\mathbf{\Phi a} + \mathbf{\Psi}^T\mathbf{H}^{(i)}\mathbf{e_x} + \mathbf{\Psi}^T\mathbf{n}^{(i)}. \tag{13}$$

This is the observation equation that is analogous to (6). It gives the relation between the unknown "true" feature vector $\mathbf{a}$ and the observed "inaccurate" feature vectors $\hat{\mathbf{a}}^{(i)}$. In the traditional way of applying super-resolution, the unknown high-resolution image $\mathbf{x}$ in (6) is reconstructed from the low-resolution observations $\mathbf{y}^{(i)}$. Then, the reconstructed $\mathbf{x}$ is fed into a face recognition system (see Fig. 1). For eigenface-based face recognition systems, a better way is to directly reconstruct the low-dimensional feature vector. Using the relation provided in (13), accurate feature vectors of a face image can be obtained from the in-

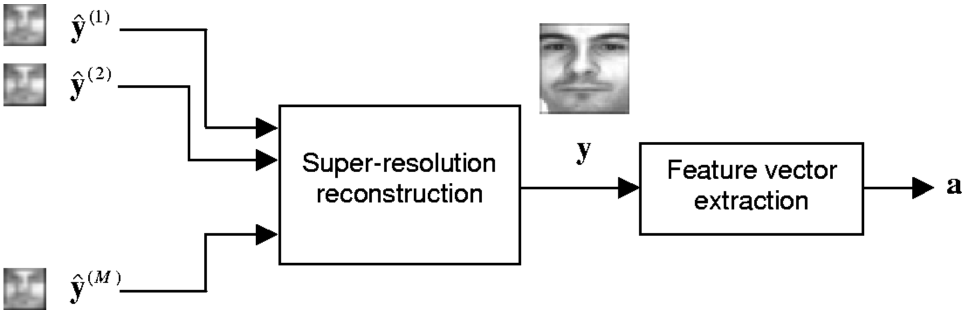


Fig. 1. Super-resolution applied as a preprocessing block to face recognition.

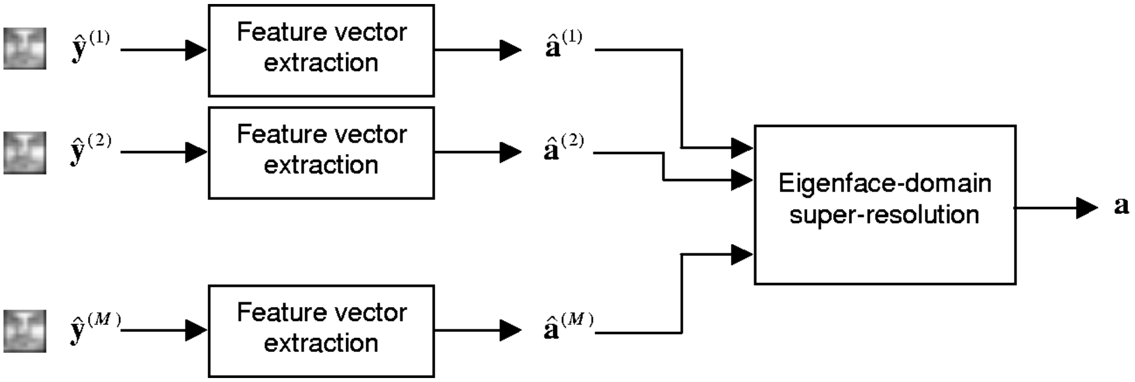


Fig. 2. Super-resolution embedded into eigenface-based face recognition.

accurate feature vector observations. This is illustrated in Fig. 2. The face observations $\mathbf{y}^{(i)}$ are first projected into the face space, and the computationally intensive super-resolution reconstruction is performed in the low-dimensional face subspace instead of in the spatial domain. A quantitative comparison of the computational complexity of these two approaches is provided in the next section.

While we are reconstructing the feature vectors in the low-dimensional subspace, we can (and will) substitute face specific information in the form of statistics of the prior distributions of the feature vectors and distributions of the noise processes. Using model-based information in regularizing the super-resolution algorithm has been shown to be successful in previous work [2]–[4]. This helps to obtain more robust results when compared to traditional super-resolution algorithms. Our experiments in this paper also confirm the advantages of using such model-based information. Our main difference, however, with respect to previous model-based algorithms is that we specifically transform all of the prior information to the low dimensional face space so that the computational complexity is kept low with little or no sacrifice in performance. This is in contrast to previous approaches that use complicated pixel-domain model-based statistical information.

IV. Reconstruction Algorithm

In this section, we present a reconstruction algorithm to solve (13) based on Bayesian estimation. The algorithm handles the observation noise and subspace representation error in the low-dimensional face subspace. The maximum *a posteriori* probability (MAP) estimator $\tilde{\mathbf{a}}$ is the argument that maximizes the product of the conditional probability $p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a})$ and the prior probability $p(\mathbf{a})$

$$\tilde{\mathbf{a}} = \arg\max_{\mathbf{a}} \left\{ p(\hat{\mathbf{a}}^{(1)}, \cdots, \hat{\mathbf{a}}^{(M)}|\mathbf{a}) p(\mathbf{a}) \right\}. \tag{14}$$

We now need to model the statistics $p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a})$ and $p(\mathbf{a})$. The prior probability $p(\mathbf{a})$ can simply be assumed to be jointly Gaussian

$$p(\mathbf{a}) = \frac{1}{Z} \exp\left(-(\mathbf{a} - \mu_{\mathbf{a}})^T \mathbf{\Lambda}^{-1} (\mathbf{a} - \mu_{\mathbf{a}})\right) \tag{15}$$

where $\mathbf{\Lambda}$ is the $L \times L$ covariance matrix, $\mu_{\mathbf{a}}$ is the $L \times 1$ mean of $\mathbf{a}$, and Z is a normalization constant.

In order to find $p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a})$, we first model the noise process in the spatial domain, and then derive its statistics in face space. We define a total noise term $\mathbf{v}^{(i)}$ that consists of the noises resulting from the subspace representation error $\mathbf{e}_{\mathbf{x}}$ and the observation noise $\mathbf{n}^{(i)}$ in the spatial domain

$$\mathbf{v}^{(i)} \triangleq \mathbf{H}^{(i)} \mathbf{e}_{\mathbf{x}} + \mathbf{n}^{(i)}. \tag{16}$$

Using this definition, we rewrite (13) for convenience

$$\hat{\mathbf{a}}^{(i)} = \mathbf{\Psi}^T \mathbf{H}^{(i)} \mathbf{\Phi} \mathbf{a} + \mathbf{\Psi}^T \mathbf{v}^{(i)}. \tag{17}$$

The reason we defined $\mathbf{H}^{(i)}\mathbf{e}_{\mathbf{x}} + \mathbf{n}^{(i)}$ as the total noise term instead of its projection onto the face subspace is because of the modeling convenience in the spatial domain. It has been demonstrated that modeling the noise [resulting from the imaging system and the estimation of $\mathbf{H}^{(i)}$] in the spatial domain as an independent identically distributed (IID) Gaussian processes is a good assumption [5], [6]. We further assume that the covariance matrix of this Gaussian process is diagonal so that the statistical parameters can be estimated easily even with the limited training data. Using these assumptions, it is easy to find the distribution of $\boldsymbol{\Psi}^T\mathbf{v}^{(i)}$ in the face space, as will be shown shortly.

Defining $\mathbf{K}$ as the $s^2N^2 \times s^2N^2$ positive definite diagonal covariance matrix and $\mu_{\mathbf{v}}^{(i)}$ as the $s^2N^2 \times 1$ mean of $\mathbf{v}^{(i)}$, we can write the probability distribution of $\mathbf{v}^{(i)}$ as

$$p(\mathbf{v}^{(i)}) = \frac{1}{Z}\exp\left(-\left(\mathbf{v}^{(i)} - \mu_{\mathbf{v}}^{(i)}\right)^T \mathbf{K}^{-1}\left(\mathbf{v}^{(i)} - \mu_{\mathbf{v}}^{(i)}\right)\right) \tag{18}$$

where Z is a normalization constant.

Now, we need to derive the distribution of the projected noise, $p(\boldsymbol{\Psi}^T\mathbf{v}^{(i)})$, in order to get the conditional PDF $p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a})$. From the analysis of functions of multivariate random variables [12], it follows that $p(\boldsymbol{\Psi}^T\mathbf{v}^{(i)})$ is also jointly Gaussian since $\boldsymbol{\Psi}^T\boldsymbol{\Psi}$ is nonsingular (by construction). As a result, we have

$$\begin{aligned} p(\boldsymbol{\Psi}^T\mathbf{v}^{(i)}) = \frac{1}{Z}\exp&\left(-\left(\boldsymbol{\Psi}^T\mathbf{v}^{(i)} - \boldsymbol{\Psi}^T\mu_{\mathbf{v}}^{(i)}\right)^T\right. \\ &\left.\cdot\mathbf{Q}^{-1}\left(\boldsymbol{\Psi}^T\mathbf{v}^{(i)} - \boldsymbol{\Psi}^T\mu_{\mathbf{v}}^{(i)}\right)\right) \end{aligned} \tag{19}$$

where $\boldsymbol{\Psi}^T\mu_{\mathbf{v}}^{(i)}$ is the new mean and $\mathbf{Q}$ is the new covariance matrix computed by

$$\mathbf{Q} = \boldsymbol{\Psi}^T\mathbf{K}\boldsymbol{\Psi}. \tag{20}$$

The covariance matrix $\mathbf{Q}$ has dimension $L \times L$ while $\mathbf{K}$ is of dimension $s^2N^2 \times s^2N^2$. Using (17) and (19), we find the conditional PDF $p(\hat{\mathbf{a}}^{(i)}|\mathbf{a})$

$$\begin{aligned} p(\hat{\mathbf{a}}^{(i)}|\mathbf{a}) = \frac{1}{Z}\exp&\left(-\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \mu_{\mathbf{v}}^{(i)}\right)^T\right. \\ &\left.\cdot\mathbf{Q}^{-1}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \mu_{\mathbf{v}}^{(i)}\right)\right). \end{aligned} \tag{21}$$

Since we assumed that $\mathbf{v}^{(i)}$ is IID, it follows that the probability density function $p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a})$ is the product of $p(\hat{\mathbf{a}}^{(i)}|\mathbf{a})$ for $i = 1, \ldots, M$. Defining $\eta \triangleq \boldsymbol{\Psi}^T\mu_{\mathbf{v}}^{(i)}$ as the mean of the process $\boldsymbol{\Psi}^T\mathbf{v}^{(i)}$, we write

$$\begin{aligned} &p(\hat{\mathbf{a}}^{(1)}, \ldots, \hat{\mathbf{a}}^{(M)}|\mathbf{a}) \\ &\quad = \frac{1}{Z}\exp\left(-\sum_{i=1}^{M}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right)^T\right. \\ &\qquad\qquad \left.\cdot\mathbf{Q}^{-1}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right)\right). \end{aligned} \tag{22}$$

Substituting the conditional and prior PDFs given in (15) and (22) into (14), we obtain the MAP estimator $\tilde{\mathbf{a}}$ as follows:

$$\begin{aligned} \tilde{\mathbf{a}} = \arg\min_{\mathbf{a}}&\left\{\sum_{i=1}^{M}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right)^T\right. \\ &\cdot\mathbf{Q}^{-1}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right) \\ &\left.+\left(\mathbf{a} - \mu_{\mathbf{a}}\right)^T\boldsymbol{\Lambda}^{-1}\left(\mathbf{a} - \mu_{\mathbf{a}}\right)\right\}. \end{aligned} \tag{23}$$

So far, we have shown how to incorporate the statistics of spatial-domain noise and prior information into the low-dimensional face-space reconstruction. In the next section, we estimate the parameters for these assumed models and provide experiments analyzing the recognition performance, effects of feature vector length, sensitivity to noise and motion estimation errors, etc.

Before getting to the experimental results, we provide an algorithm to solve (23). One approach to obtain the MAP estimate $\tilde{\mathbf{a}}$ is an iterative steepest descent method. Defining $E(\mathbf{a})$ as the cost function to be minimized, the feature vector $\mathbf{a}$ can be updated in the direction of the negative gradient of $E(\mathbf{a})$. That is, at the nth iteration, the feature vector can be updated as follows:

$$\mathbf{a}_n = \mathbf{a}_{n-1} - \alpha\nabla E(\mathbf{a}_{n-1}) \tag{24}$$

where α is the step size.

From (23), a slightly generalized cost function is chosen as

$$\begin{aligned} E(\mathbf{a}) = \frac{1-\lambda}{2}&\sum_{i=1}^{M}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right)^T \\ &\cdot\mathbf{Q}^{-1}\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right) \\ &+\frac{\lambda}{2}\left(\mathbf{a} - \mu_{\mathbf{a}}\right)^T\boldsymbol{\Lambda}^{-1}\left(\mathbf{a} - \mu_{\mathbf{a}}\right) \end{aligned} \tag{25}$$

where λ is a number, $(0 \leq \lambda \leq 1)$, that controls the relative contribution of the prior information in the reconstruction. When λ is set to zero, the estimator becomes a maximum likelihood (ML) estimator. When λ is one, only the prior information is used, and the noise statistics are discarded. $\lambda = 1/2$ corresponds to the original MAP estimator.

Taking the derivative of $E(\mathbf{a})$ with respect to $\mathbf{a}$, the gradient of $E(\mathbf{a})$ can be calculated as

$$\begin{aligned} \nabla E(\mathbf{a}) = -(1-\lambda)&\sum_{i=1}^{M}\boldsymbol{\Phi}^T\mathbf{H}^{(i)^T}\boldsymbol{\Psi}\mathbf{Q}^{-1} \\ &\cdot\left(\hat{\mathbf{a}}^{(i)} - \boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi}\mathbf{a} - \eta\right) + \lambda\boldsymbol{\Lambda}^{-1}\left(\mathbf{a} - \mu_{\mathbf{a}}\right). \end{aligned} \tag{26}$$

Although the step size α in (24) can be chosen as fixed, a better way is to update it using the Hessian of $E(\mathbf{a})$. In this case, α is updated at each iteration using the formula

$$\alpha = \frac{\left(\nabla E(\mathbf{a}_{n-1})\right)^T\left(\nabla E(\mathbf{a}_{n-1})\right)}{\left(\nabla E(\mathbf{a}_{n-1})\right)^T H\left(\nabla E(\mathbf{a}_{n-1})\right)} \tag{27}$$

where H is the Hessian matrix found by

$$H = (1-\lambda)\sum_{i=1}^{M}\boldsymbol{\Phi}^T\mathbf{H}^{(i)^T}\boldsymbol{\Psi}\mathbf{Q}^{-1}\boldsymbol{\Psi}^T\mathbf{H}^{(i)}\boldsymbol{\Phi} + \lambda\boldsymbol{\Lambda}^{-1}. \tag{28}$$

In the reconstruction, everything but $\mathbf{a}$, $\mathbf{a}^{(i)}$, and $\mathbf{H}^{(i)}$ is known and can be computed in advance. (The details are left to the next section.) For a specific observation sequence $\mathbf{y}^{(i)}$, the feature vectors $\mathbf{a}^{(i)}$ and the blur mappings $\mathbf{H}^{(i)}$ are computed, and the true feature vector $\mathbf{a}$ is estimated. The pseude-code of the complete algorithm is as follows.

1) Choose a reference frame from the video sequence, bilinearly interpolate it, and project it onto the face space to obtain an initial estimate $\mathbf{a}_0$ for the true feature vector.
2) Obtain the feature vector $\mathbf{a}^{(i)}$ by projecting each low-resolution frame onto the face space. [That is, $\hat{\mathbf{a}}^{(i)} = \boldsymbol{\Psi}^T\mathbf{y}^{(i)}$.]

3) Estimate the motion between the reference and other frames, and compute $\mathbf{H}^{(i)}$.
4) Set the maximum number of iterations, $MaxIter$.
5) For $n = 1$ to $MaxIter$,
   a) Compute $\nabla E(\mathbf{a})$ using (26).
   b) Compute H using (28).
   c) Compute α using (27).
   d) Compute $\mathbf{a}_n$ using (24).
6) Set the MAP estimate $\tilde{\mathbf{a}}$ to $\mathbf{a}_{MaxIter}$.

We now take a look at the computational complexity of the proposed algorithm compared to a pixel-domain reconstruction. Let P be the total number of pixels in a (high-resolution) face image, Q be the length of the feature vectors, and s be the downsampling factor $(0 < s < 1)$. Excluding the motion estimation stage of the reconstruction, most of the computational cost results from the computation of $\mathbf{\Psi}^T\mathbf{H}^{(i)}\mathbf{\Phi}\mathbf{a}$. According to our image acquisition model, $\mathbf{H}^{(i)}$ can be represented as the successive application of motion warping, PSF blurring, and downsampling [see (7)]. Since the blurring and downsampling operations are time-invariant, only the motion warping operation needs to be computed for each observation separately. Denoting $\mathbf{W}^{(i)}$, $\mathbf{B}$, and $\mathbf{D}$ as the motion warping, blurring, and downsampling matrices, respectively, we need three matrix-vector multiplications to compute $\mathbf{\Psi}^T\mathbf{H}^{(i)}\mathbf{\Phi}\mathbf{a}$, where $\mathbf{H}^{(i)} = \mathbf{DBW}^{(i)}$. The first one is $\mathbf{\Phi}\mathbf{a}$, which requires approximately $2PQ$ multiplications and additions. This is then multiplied by $\mathbf{W}^{(i)}$, which requires $2P^2$ multiplications and additions. This is followed by a multiplication with the $Q \times P$ matrix $\mathbf{\Psi}^T\mathbf{BD}$, which can be precomputed and stored. The total number of multiplications and additions is approximately $2P^2 + 4PQ$. On the other hand, doing these operations in the pixel domain [using the matrix $\mathbf{H}^{(i)}$] requires $2P^2 + 4sP^2$ operations. Referring to the gradient and Hessian matrix computations [(26) and (27)], the eigenface-space reconstruction requires roughly $3M\left(4sP^2 - 4PQ\right)$ fewer operations per iteration than the spatial-domain reconstruction does. (In our experiments, P is 1600, s is 0.25, Q is 40, and M is 16.)

V. Experimental Results

We performed a set of experiments to demonstrate the efficacy of the proposed method. We investigated the effect of the face-space dimension, and sensitivity to noise and motion estimation errors. We have also performed a recognition experiment with real video sequences. We will explain each step of the experiments in detail.

A. Obtaining the Face Subspace

In these experiments, we used face images from the Yale face databases A and B [13], Harvard Robotics Laboratory database [14], AR database [15], and CMU database [16]. The images are downsampled to have a size of 40×40, and aligned according to the manually located eye and mouth locations. We selected 134 images as training data and 50 images as test data. We applied the KLT to those 134 images and chose the first 60 eigenvectors having the largest eigenvalues to form the face subspace. (These 60 eigenvectors form the columns of the matrix $\mathbf{\Phi}$.) We also downsampled the training images by four to obtain 10×10 images, applied the KLT to those images, and chose the first 60 of them to construct the eigenface space $\mathbf{\Psi}$.

B. Obtaining Low-Resolution Observations for Synthetic Video

The test images were jittered by a random amount to simulate motion, blurred, and downsampled by a factor of four to produce multiple low-resolution images for each subject. The motion vectors were saved for use in synthetic video experiments. For blurring, the images were convolved with a point spread function (PSF), which was set to a 5×5 normalized Gaussian kernel with zero mean and a standard deviation of one pixel.

C. Estimating the Statistics of Noise and Feature Vectors

From the training image set $\mathbf{I}_1, \ldots, \mathbf{I}_K$, $(K = 134)$, we estimate the statistics of $\mathbf{a}$ and $\mathbf{v}^{(i)}$. The unbiased estimates for the mean and covariance matrix of $\mathbf{a}$ are simply obtained from the sample mean and variances

$$\mu_{\mathbf{a}} \simeq \frac{1}{K}\sum_{j=1}^{K}\left(\mathbf{\Phi}^T\mathbf{I}_j\right) \tag{29}$$

and

$$\mathbf{\Lambda} \simeq \frac{1}{K}\sum_{j=1}^{K}\left(\mathbf{\Phi}^T\mathbf{I}_j - \mu_{\mathbf{a}}\right)\left(\mathbf{\Phi}^T\mathbf{I}_j - \mu_{\mathbf{a}}\right)^T. \tag{30}$$

Because of the limited number of training images, for more reliable estimation, we assume a diagonal covariance matrix, so the off-diagonal elements of the matrix $\mathbf{\Lambda}$ are set to zero.

The mean and covariance matrices of $\mathbf{v}^{(i)}$ are found similarly. Letting $\mathbf{y}_j^{(i)}$ be the ith observation of the jth training image, $(i = 1, \ldots, M$ and $j = 1, \ldots, K)$, we estimate the mean and covariance matrices as follows:

$$\mu_{\mathbf{v}} \simeq \frac{1}{KM}\sum_{j=1}^{K}\sum_{i=1}^{M}\left(\mathbf{y}_j^{(i)} - \mathbf{H}^{(i)}\mathbf{\Phi}\mathbf{\Phi}^T\mathbf{I}_j\right) \tag{31}$$

and

$$\mathbf{K} \simeq \frac{1}{KM}\sum_{j=1}^{K}\sum_{i=1}^{M}\left(\mathbf{y}_j^{(i)} - \mathbf{H}^{(i)}\mathbf{\Phi}\mathbf{\Phi}^T\mathbf{I}_j - \mu_{\mathbf{v}}\right) \cdot \left(\mathbf{y}_j^{(i)} - \mathbf{H}^{(i)}\mathbf{\Phi}\mathbf{\Phi}^T\mathbf{I}_j - \mu_{\mathbf{v}}\right)^T. \tag{32}$$

Again, the off-diagonals of $\mathbf{K}$ are set to zero. The mean η and covariance matrix $\mathbf{Q}$ for $\mathbf{\Psi}^T\mathbf{v}^{(i)}$ are found using $\eta = \mathbf{\Psi}^T\mu_{\mathbf{v}}$ and $\mathbf{Q} = \mathbf{\Psi}^T\mathbf{K}\mathbf{\Psi}$.

D. Reconstruction for Synthetic Video

One of the frames for each video sequence is chosen as the reference frame, bilinearly interpolated by four, and projected onto the face space $\mathbf{\Phi}$ to obtain the initial estimate for the true feature vector. It is then updated using the algorithm proposed in the previous section. The mapping $\mathbf{H}^{(i)}$ is computed from the known motion vectors and PSF, and 16 low-resolution images are used in the reconstruction. The model parameters $\mu_{\mathbf{a}}$, $\mathbf{\Lambda}$, η, and $\mathbf{Q}$ computed in Step C are used in the reconstruction with λ set to 0.5. The number of iterations $MaxIter$ is set to seven for each sequence.

We also wanted to compare the results of this eigenface-domain super-resolution algorithm with a traditional pixel-domain super-resolution. We applied the pixel-domain super-resolution algorithm given in [11] to the low-resolution video sequences again using the same 16 low-resolution images and setting the number iterations to seven. After the high-resolution images are

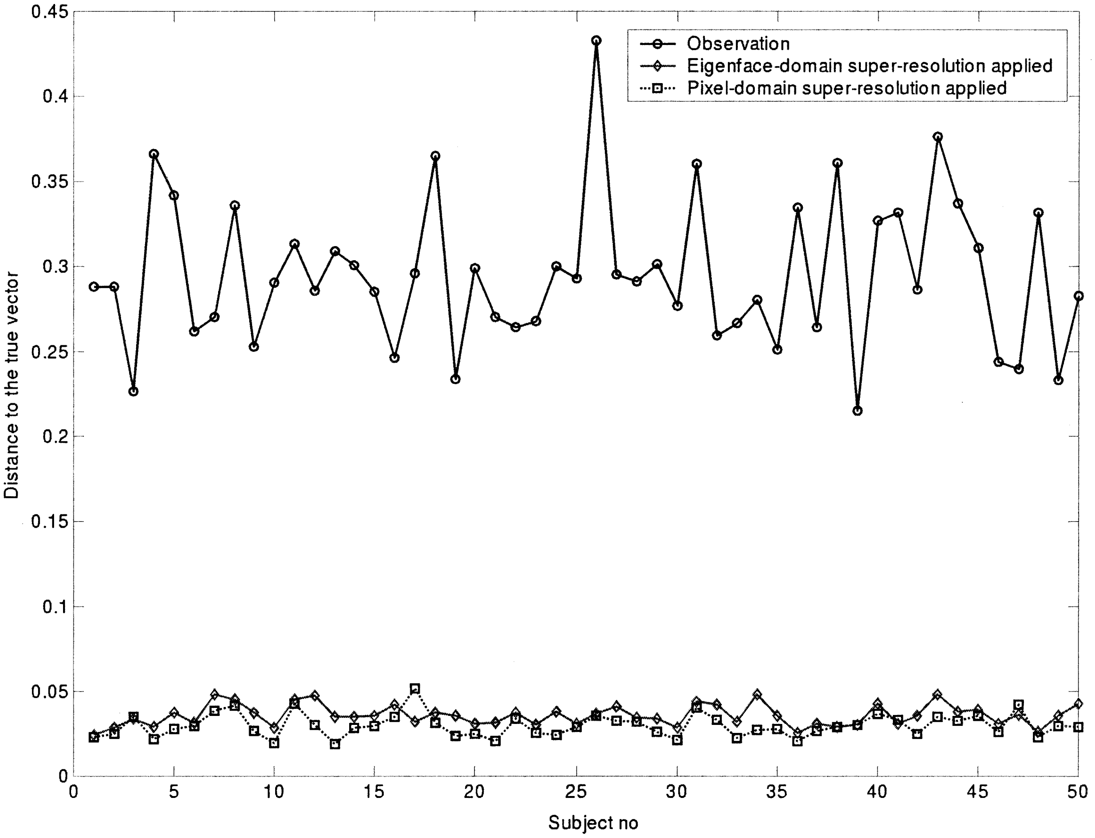


Fig. 3. Error in feature vector computation.

reconstructed, they are projected onto the face space $\mathbf{\Phi}$ to obtain the feature vectors.

The feature vectors obtained from these algorithms are compared with the true feature vectors (which are computed using the 40×40 original high-resolution images). For each subject (image sequence), we computed the normalized distance between the true feature vector $\mathbf{a}$ and the estimated feature vector $\tilde{\mathbf{a}}$. The normalized distance $D(\tilde{\mathbf{a}}, \mathbf{a})$ is defined as

$$D(\tilde{\mathbf{a}}, \mathbf{a}) = \frac{||\tilde{\mathbf{a}} - \mathbf{a}||}{||\mathbf{a}||} \times \frac{1}{Length(\mathbf{a})} \times 100 \qquad (33)$$

where $Length(\mathbf{a})$ is the length of vector $\mathbf{a}$.

Fig. 3 shows the results for three cases: i) Feature vectors computed from a single observation (no super-resolution applied). ii) Feature vectors computed after pixel-domain super-resolution applied. iii) Feature vectors reconstructed using the proposed eigenface-domain super-resolution. As seen in the figure, eigenface-domain super-resolution achieves a similar performance to the pixel-domain super-resolution at less computation.

We also provide an example from the face database. Fig. 4 shows the results for *Subject 1* in the test data. In that figure, (a) is the original 40 × 40 image, (b) is one of the observations interpolated using nearest neighbor interpolation, (c) is the bilinearly interpolated observation, which is the initial estimate in reconstruction, (d) is the result of the pixel-domain super-resolution, (e) is the projection of the result in (d) into the face space, and (f) is the representation of the reconstructed feature-vector from the eigenface-domain super-resolution algorithm. As seen, (e) and (f) are almost identical, but (f) is obtained at a lower computational burden.

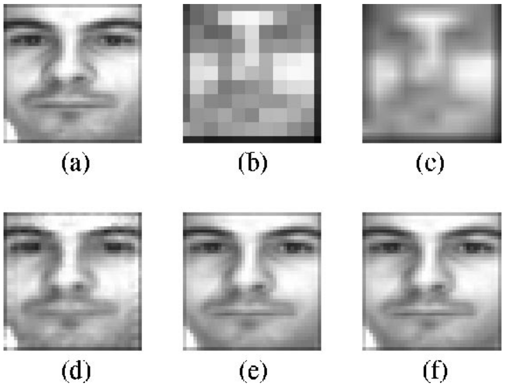


Fig. 4. (a) Original 40 × 40 image. (b) 10 × 10 low-resolution observation is interpolated using nearest neighbor interpolation. (c) 10 × 10 low-resolution observation is interpolated using bilinear interpolation. (d) Pixel-domain super-resolution applied. (e) The result of pixel-domain super-resolution reconstruction is projected into the face subspace. (f) Representation of the feature vector reconstructed using the eigenface-domain super-resolution in the face subspace.

This experiment was done for a face-space dimension of 60, which brings up the question of how the feature vector length (i.e., dimension of the face space) affects the performance. This question is addressed in the next experiment. We will also demonstrate that eigenface-domain super-resolution is more robust to noise and motion estimation errors than the pixel-domain super-resolution.

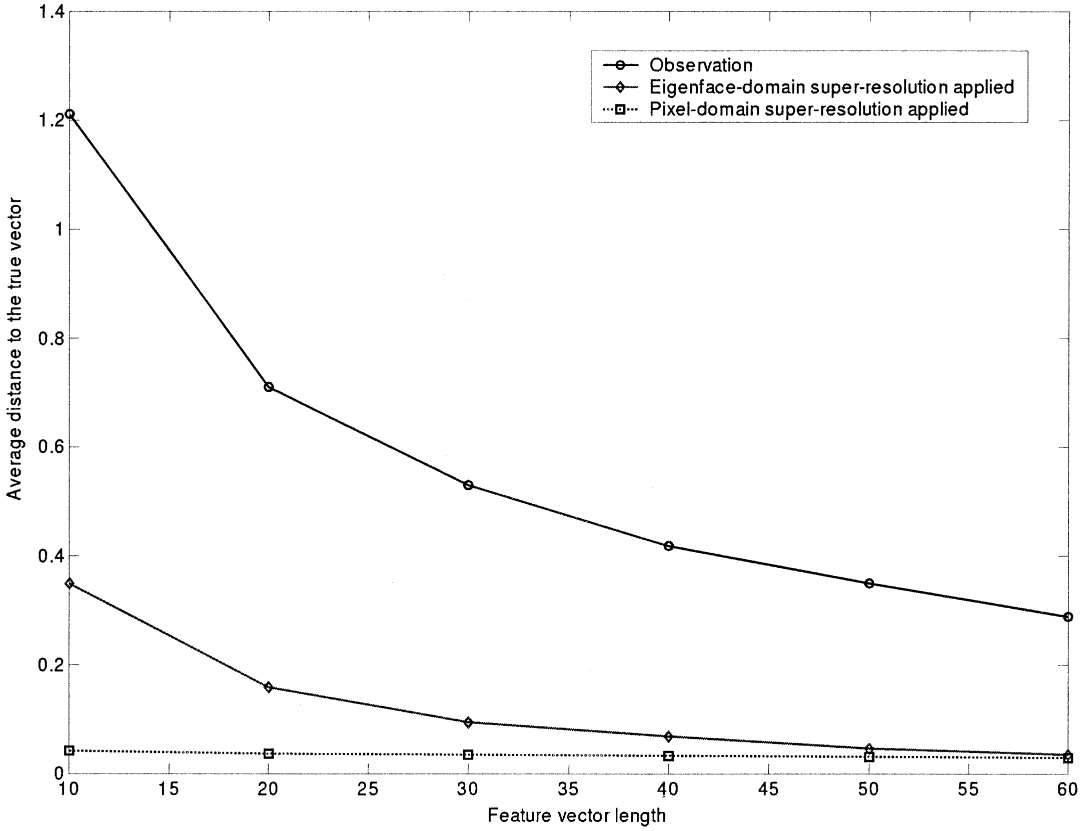


Fig. 5. Effect of feature vector length on performance.

E. Effect of Feature Vector Size on Reconstruction

We repeated the experiments for various feature vector sizes to examine the effect of the face-space dimension on reconstruction. The results are given in Fig. 5. In that figure, the x-axis is the dimension of the face space, and the y-axis is the normalized distance averaged over 50 subjects. Due to the face-space representation error, pixel-domain super-resolution performs better than the eigenface-domain super-resolution at very low face-space dimensions. As expected, as the feature vector size is increased, the performance of the eigenface-domain super-resolution approaches that of the pixel-domain super-resolution. Note that this is the result for the case where there is no observation noise or motion estimation error. As will be shown shortly, when there is noise or motion estimation error, eigenface-domain super-resolution becomes better than the pixel-domain super-resolution even at the low face-space dimensions. This is because the solution obtained in eigenface domain is constrained by face-specific priors.

F. Effect of Noise on Reconstruction

In order to examine the effects of observation noise, we added zero-mean Gaussian IID noise to each low-resolution video frame. The experiment is done for a feature vector size of 40, and repeated for each of the 50 video sequences. Fig. 6 shows the results for different noise powers. (The x-axis is the variance of the noise, and the y-axis is the average normalized distance.) As seen in that figure, when the noise power is zero, the pixel-domain super-resolution is better than the eigenface-domain super-resolution. However, as the noise power increases, eigenface-domain super-resolution outperforms pixel-domain super-resolution. The reason is that eigenface-domain super-resolution constrains the solution to lie in the face space, and therefore, it is more robust to noise.

G. Effect of Motion Estimation Error on Reconstruction

In addition to the robustness to observation noise, eigenface-domain super-resolution is also more robust to motion estimation errors than pixel-domain super-resolution. This time, we perturbed each true motion vector with a zero-mean Gaussian IID random vector to simulate the motion estimation error. The face dimension for the experiment is again 40. As seen in Fig. 7, as the motion estimation error increases, the pixel-domain super-resolution becomes worse than eigenface-domain super-resolution immediately. It is also observed that the pixel-domain super-resolution becomes even worse than using only one image to get the feature vector. Again eigenface-domain super-resolution is less sensitive to motion estimation errors because of the face-space regularization.

H. Recognition Experiment With Real Video Sequences

Finally, we tested the proposed algorithm with real video sequences from the CMU database. We performed a recognition experiment with a database of 68 people. For each person, we selected the neutral face image from the facial expression part of the database as the training image. We manually located the positions of the eyes and the mouth in those images, cropped them according to those locations, downsampled them to a size of 40 × 40, and projected them into the eigenspace to get the training feature vector for each person.

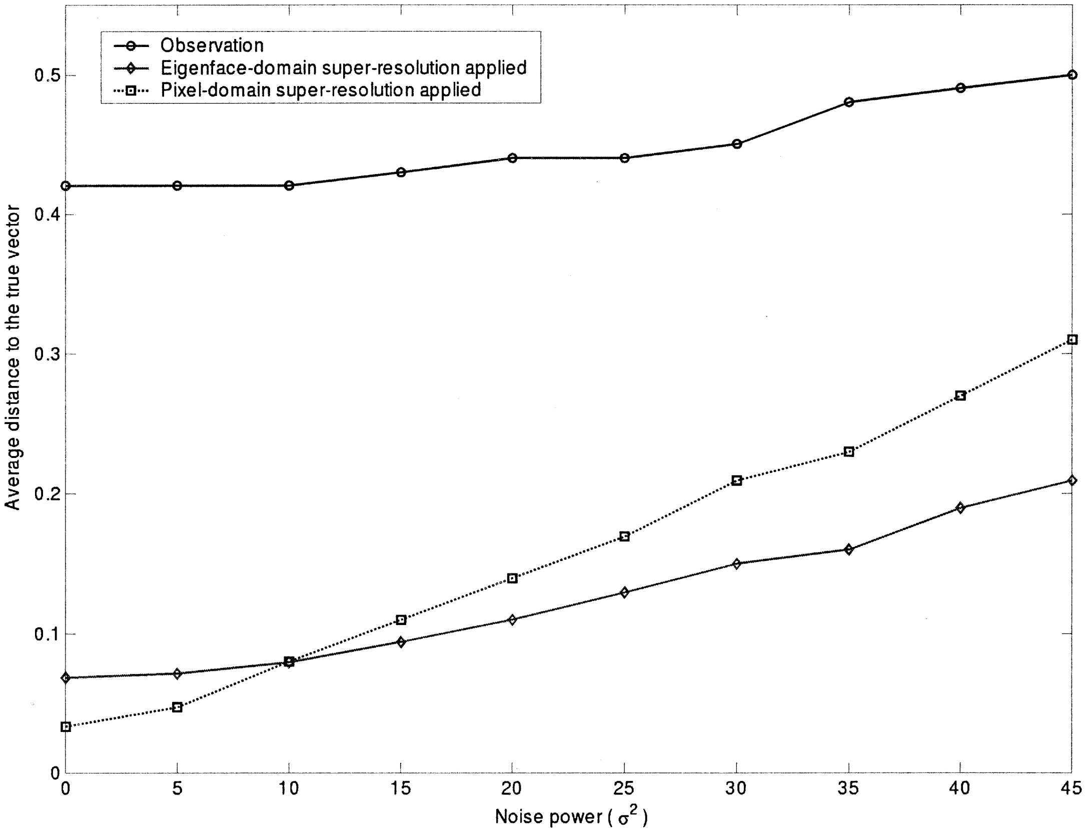


Fig. 6. Effect of observation noise on performance.

To perform recognition, we used the talking video sequences provided in the CMU database. Each sequence contains a single person talking for 2 s. We had a total of 68 such sequences, one for each person in our database. The goal of our recognition experiment is to identify the person who appears in the video. We used 16 consecutive images from each video sequence. The original sequences are very high resolution, so we downsampled them so that the face is around 40 pixels wide. The resulting sequences form our high-resolution face image sequences, and we use them as the ground truth to evaluate the success of our experiments. We then blurred these face image sequences (using the PSF given in Step B) and downsampled them (by four) to form low-resolution observations. These low-resolution image sequences are the input images for the recognition experiment. We manually located the positions of the eyes and the mouth in the first frame of these image sequences.

Then, we ran three different recognition experiments. In the first experiment, we used the first image from each low-resolution image sequence for recognition. We cropped the faces from the frames according to the locations of the eyes and the mouth, projected them into the eigenspace, and performed minimum distance classification with the $L2$ norm. The recognition rate in this case was 44%. In the second experiment, we again cropped the faces from the first frames of the low resolution image sequences according to the locations of the eyes and the mouth. Then, we used block based motion estimation to get the motion vectors from one image frame to the other. In motion estimation, we computed the motion vectors for each pixel with quarter-pixel accuracy. We set the block size and the search range to 8 and ±8, respectively, and we found motion vectors for each pixel by performing a full search with mean absolute difference being the matching criteria. Then, we projected all low resolution face images into the eigenface space, and performed eigenface space super-resolution to construct an accurate feature vector for each person. The recognition experiment in this case provided a recognition rate of 74%. In the third experiment, we used the first frame of each high-resolution video sequence to perform recognition. The recognition rate with these high-resolution images was 79%.

The results we reported above show that the decrease in the resolution of the face image decreases the recognition rate significantly. (In our experiments, the decrease was from 79% to 44%.) With the super-resolution reconstruction, the recognition rate improved significantly, and got close to the high-resolution recognition rate.

VI. Conclusions

The performance of face recognition systems decreases significantly if the resolution of the face image falls below a certain level. For video sequences, super-resolution techniques can be used to obtain a high-resolution face image by combining the information from multiple low-resolution images. Although super-resolution can be applied as a separate preprocessing block, in this paper, we propose to apply super-resolution after dimensionality reduction in a face recognition system.

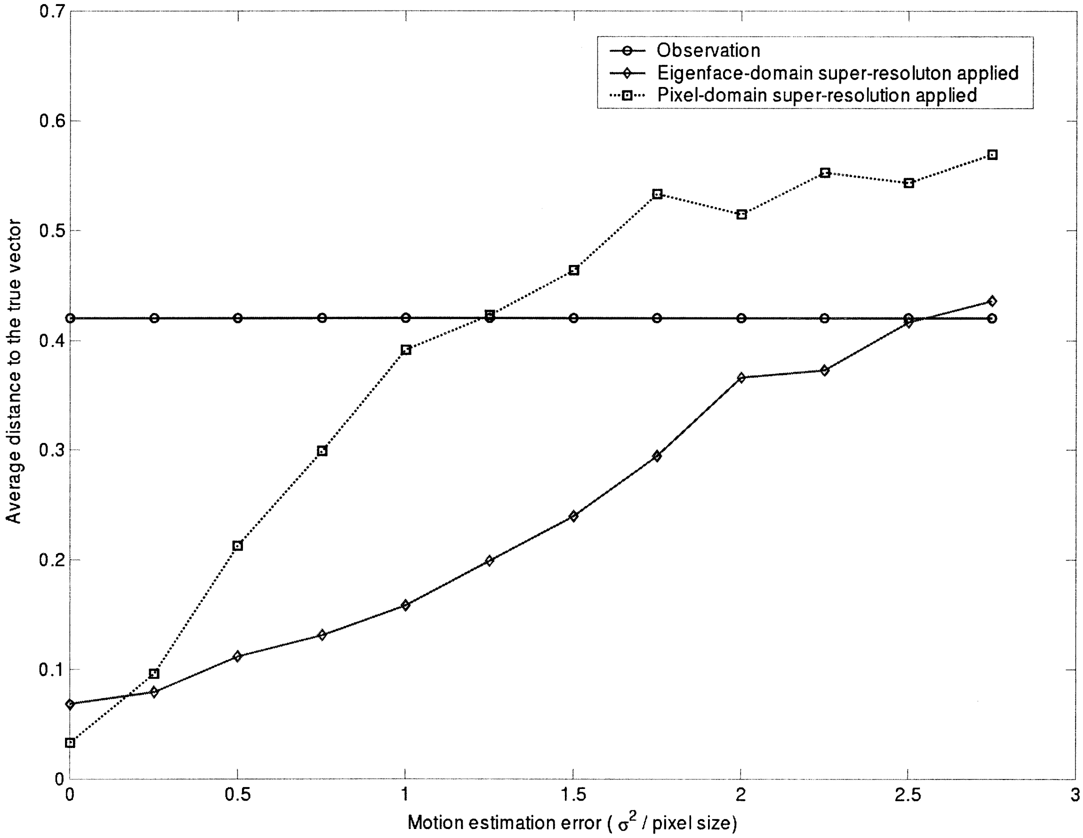


Fig. 7. Effect of motion estimation error on performance.

In this way, only the necessary information for recognition is reconstructed. We have also shown how to incorporate the model-based information into the face-space reconstruction algorithm. This helps to obtain more robust results when compared to the traditional super-resolution algorithms. In the experiments, we demonstrated robustness to noise and motion estimation error. We have investigated the effect of face-space dimension on the reconstruction, and provided recognition results for real video sequences.

This paper only examines the case for face images; however, the idea can be extended to other pattern recognition problems easily. One such application is the recognition of car license plates from video.

References

[1] T. E. Boult, M.-C. Chiang, and R. J. Micheals, "Super-resolution via image warping," in *Super-Resolution Imaging*, E. S. Chaudhuri, Ed. Norwell, MA: Kluwer, 2001, pp. 131–169.

[2] S. Baker and T. Kanade, "Hallucinating faces," in *Proc. 4th Int. Conf. Automatic Face and Gesture Recognition*, Mar. 2000.

[3] C. Liu, H.-Y. Shum, and C.-S. Zhang, "A two-step approach to hallucinating faces: Global parametric model and local nonparametric model," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Dec. 2001.

[4] D. Capel and A. Zisserman, "Super-resolution from multiple views using learnt image models," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Dec. 2001.

[5] R. R. Schultz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Processing*, vol. 5, pp. 996–1011, June 1996.

[6] M. Elad and A. Feuer, "Restoration of a single superresolution image from several blurred, noisy and undersampled measured images," *IEEE Trans. Image Processing*, vol. 6, pp. 1646–1658, Dec. 1997.

[7] B. C. Tom, A. K. Katsaggelos, and N. P. Galatsanos, "Reconstruction of a high resolution image from registration and restoration of low resolution images," in *Proc. IEEE Int. Conf. Image Processing*, Austin, TX, Nov. 1994, pp. 13–16.

[8] R. C. Hardie, K. J. Barnard, and E. E. Armstrong, "Joint MAP registration and high-resolution image estimation using a sequence of undersampled images," *IEEE Trans. Image Processing*, vol. 6, pp. 1621–1633, Dec. 1997.

[9] L. Sirovich and M. Kirby, "'Low dimensional procedure for the characterization of human faces," *J. Opt. Soc. Amer. A*, vol. 4, no. 3, pp. 519–524, 1987.

[10] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1991, pp. 586–591.

[11] A. J. Patti, M. I. Sezan, and A. M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Trans. Image Processing*, vol. 6, pp. 1064–1076, Aug. 1997.

[12] H. Stark and J. W. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[13] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Generative models for recognition under variable pose and illumination," in *Proc. IEEE Conf. Face and Gesture Recognition*, 2000, pp. 277–284.

[14] P. Hallinan, "A low dimensional representation of human faces for arbitrary lighting conditions," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1994, pp. 995–999.

[15] A. M. Martinez and R. Benavente, "The AR face database," in *Proc. CVC Tech. Rep. 24*, 1998.

[16] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression (PIE) database of human faces," The Robotics Institute, Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep. CMU-RI-TR-01-02, 2001.

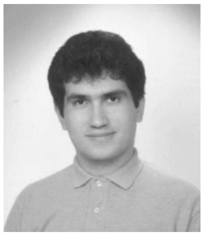
Bahadir K. Gunturk (S'01) received the B.S. degree in electrical engineering from Bilkent University, Ankara, Turkey, in 1999, and the M.S. degree in electrical engineering from Georgia Institute of Technology, Atlanta, in 2001. He is currently pursuing the Ph.D. degree at the Georgia Institute of Technology.

His research interests include image/video processing, multimedia communications, and computer vision.

Mr. Gunturk received the Outstanding Research Award from the Center for Signal and Image Processing, Georgia Institute of Technology, in 2001.

Aziz U. Batur (S'01) received his B.S. degree in electrical engineering from Bilkent University, Turkey, in 1998, and his M.S. degree in electrical engineering from Georgia Institute of Technology, Atlanta, in 2000. He is currently pursuing the Ph.D. degree in the Center for Signal and Image Processing, Georgia Institute of Technology.

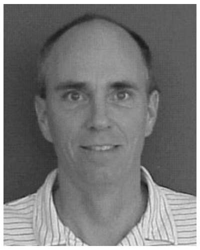
His research interests include image processing and computer vision.



Yucel Altunbasak (S'94–M'97–SM'01) received the B.S. degree from Bilkent University, Ankara, Turkey, in 1992 with highest honors. He received the M.S. and Ph.D. degrees from the University of Rochester, Rochester, NY, in 1993 and 1996, respectively.

He joined Hewlett-Packard Research Laboratories (HPL), Palo Alto, CA, in July 1996. His position at HPL provided him with the opportunity to work on a diverse set of research topics, such as video processing, coding and communications, multimedia streaming and networking. He also taught digital video and signal processing courses at Stanford University, Stanford, CA, and San Jose State University, San Jose, CA, as a Consulting Assistant Professor. He joined the School of Electrical and Computer Engineering, Georgia Institute of Technology, in 1999 as an Assistant Professor. He is currently working on industrial- and government-sponsored projects related to video and multimedia signal processing, inverse problems in imaging, and network distribution of compressed multimedia content. His research efforts resulted in over 75 publications and 12 patents/patent applications.

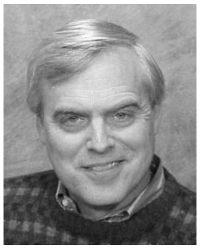
Dr. Altunbasak is an area/associate editor for IEEE TRANSACTIONS ON IMAGE PROCESSING, SIGNAL PROCESSING: IMAGE COMMUNICATIONS, and for the *Journal of Circuits, Systems and Signal Processing*. He is a member of the IEEE Signal Processing Society's IMDSP Technical Committee. He serves as a co-chair for "Advanced Signal Processing for Communications" Symposia at ICC'03. He also serves as a session chair in technical conferences, as a panel reviewer for government funding agencies, and as a technical reviewer for various journals and conferences in the field of signal processing and communications. He received the National Science Foundation (NSF) CAREER Award in 2002.



Monson H. Hayes, III (F'92) received the Sc.D. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, in 1981.

He then joined the faculty at Georgia Tech, Atlanta, where he is currently a Professor of electrical and computer engineering. He has published over 100 papers and is the author of two textbooks. His research interests are in the areas of face recognition, multimedia signal processing, image and video processing, adaptive signal processing, and DSP education.

Dr. Hayes is a recipient of the Presidential Young Investigator Award and the the IEEE Senior Award. He has been a member of the DSP Technical Committee (1984–1989), and Chairman (1995–1997). He was Associate Editor for the TRANSACTIONS on ACOUSTICS, SPEECH AND SIGNAL PROCESSING (1984–1988), Secretary-Treasurer of the ASSP Publications Board (1986–1988), Member of the ASSP Administrative Committee (1987–1989), General Chairman of the 1988 DSP Workshop, Member of the Signal Processing Society Standing Committee on Constitution and Bylaws (1988–1994), Chairman of the ASSP Publications Board (1988–1994), Member of the Technical Directions Committee (1992–1994), and General Chairman of ICASSP-96. Currently, he is Associate Editor for the IEEE TRANSACTIONS ON EDUCATION, member of the *Signal Processing Magazine* editorial board, a member of the MMSP Technical Committee, and General Chair for ICIP-06 in Atlanta.



Russell M. Mersereau (F'83) received the S.B. and S.M. degrees in 1969 and the Sc.D. degree in 1973 from the Massachusetts Institute of Technology, Cambridge.

He joined the School of Electrical and Computer Engineering at the Georgia Institute of Technology, Atlanta, in 1975. His current research interests are in the development of algorithms for the enhancement, modeling, and coding of computerized images, synthesis aperture radar, and computer vision. In the past this research has been directed to problems of distorted signals from partial information of those signals, computer image processing and coding, the effect of image coders on human perception of images, and applications of digital signal processing methods in speech processing, digital communications, and pattern recognition. He is the coauthor of the text *Multidimensional Digital Signal Processing*.

Dr. Mersereau has served on the editorial board of the PROCEEDINGS OF THE IEEE and as Associate Editor for Signal Processing of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, and the IEEE SIGNAL PROCESSING LETTERS. He is the corecipient of the 1976 Bowder J. Thompson Memorial Prize of the IEEE for the best technical paper by an author under the age of 30, a recipient of the 1977 Research Unit Award of the Southeastern Section of the ASEE, and three teaching awards. He was awarded the 1990 Society Award of the Signal Processing Society. He is currently the Vice President for Awards and Membership of the Signal Processing Society.