CIRCUITS SYSTEMS SIGNAL PROCESS
VOL. 19, NO. 4, 2000, PP. 321–338

A WAVELET-BASED INTERPOLATION-RESTORATION METHOD FOR SUPERRESOLUTION (WAVELET SUPERRESOLUTION)*

Nhat Nguyen[1] *and Peyman Milanfar*[2]

Abstract. Superresolution produces high-quality, high-resolution images from a set of degraded, low-resolution images where relative frame-to-frame motions provide different looks at the scene. Superresolution translates data temporal bandwidth into enhanced spatial resolution. If considered together on a reference grid, given low-resolution data are nonuniformly sampled. However, data from each frame are sampled regularly on a rectangular grid. This special type of nonuniform sampling is called interlaced sampling. We propose a new wavelet-based interpolation-restoration algorithm for superresolution. Our efficient wavelet interpolation technique takes advantage of the regularity and structure inherent in interlaced data, thereby significantly reducing the computational burden. We present one- and two-dimensional superresolution experiments to demonstrate the effectiveness of our algorithm.
Key words: Superresolution, wavelet interpolation, interlaced sampling.

1. Introduction

Image superresolution refers to image processing algorithms that produce high-quality, high-resolution (HR) images from a set of low-quality, low-resolution (LR) images. There is always a demand for better-quality images. However, the level of image detail is crucial for the performance of several computer vision algorithms. Target recognition, detection, and identification systems are some of the military applications that require the highest-quality achievable images. License plate readers, surveillance monitors, and medical imaging applications are examples of civilian applications with the same requirement. In many visual

* Received December 6, 1999; revised April 15, 2000.
[1] Scientific Computing and Computational Mathematics Program, Stanford University, Stanford, California 94305-9025, USA. E-mail: nguyen@sccm.stanford.edu. This work was supported in part by the National Science Foundation Grant CCR-9984246.
[2] Department of Electrical Engineering, University of California, Santa Cruz, California 95064, USA. E-mail: milanfar@cse.ucsc.edu

applications, the imaging sensors have poor resolution outputs. When resolution cannot be improved by replacing sensors, either because of cost or hardware physical limits, we resort to using a superresolution algorithms. Even when superior equipment is available, superresolution algorithms are an inexpensive alternative.

Superresolution, at its core, is a process by which one gains spatial resolution in return for temporal bandwidth. Temporal bandwidth refers to the availability of multiple nonredundant images of the same scene. Lukosz [8], [9] was first to realize this possibility. However, superresolution cannot perform miracles. We cannot expect to extract subpixel information from a sequence of identical images; there must be nonredundant information among the images. We must be able to translate data temporal bandwidth into subpixel image content. Each LR frame provides a different "look" at the same scene. Theoretically, by providing different lighting conditions or different sensors, superresolution can be achieved without relative scene motion. This is the multichannel data fusion superresolution problem. In this paper, however, we assume one imaging device and the same lighting conditions, and we require that there be some relative motion from frame to frame. Frame-to-frame motion can be a combination of camera platform motion relative to the scene, moving objects in the scene, and camera jitters. For example, in satellite imaging, images of the ground below are captured as the camera orbits the earth, whereas in surveillance and monitoring applications, the camera is placed on a fixed platform, and observed objects move within the scene. Motion and nonredundant information are what make superresolution possible. With this information, we are able to extract subpixel content at a higher resolution than in each individual frame.

Figure 1 illustrates the problem setup. The figure shows three 4×4 pixel LR frames on an 8×8 HR grid. Each symbol (square, circle, triangle) indicates the sampling points of a frame with respect to the HR grid. We pick an arbitrary frame as a reference frame; in this case, the frame marked by the circular symbols. The sampling grid for the triangular frame is a simple translation of the reference frame grid. The motion between the sampling grid for the square frame and the reference frame grid includes translational, rotational, and magnification (zoom) components.

The forward relationship between a degraded, LR frame and the ideal HR image can be described as follows [5]:

$$\mathbf{f}_k = DCE_k\mathbf{x} + \mathbf{n}_k, \qquad 1 \leq k \leq p, \tag{1}$$

where D is the downsampling operator, C is the blurring/averaging operator, E_k's are the affine transforms that map the HR grid coordinate system to the LR grid systems, $\mathbf{x}$ is the unknown ideal HR image, and $\mathbf{n}_k$'s are the additive noise vectors. The LR frames $\mathbf{f}_k$ are given, and the decimation operator D is known. Because Charge-Coupled Device (CCD) sensors on the same array have practically identical characteristics, C is spatially linear shift invariant. In this paper, we assume that frame-to-frame motion and blurring parameters are known

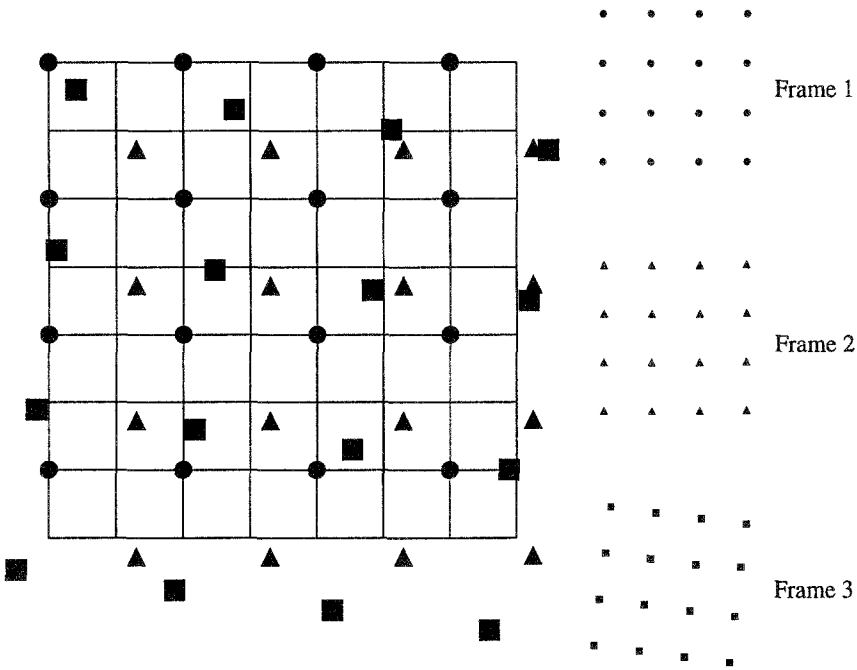

Figure 1. Low-resolution data on a high-resolution grid.

a priori or have been estimated (cf. [12], [13]) from given data and that frame-to-frame motion is purely translational or has been corrected to be so. Finally, with multiple independent sources of error, the central limit theorem allows us to assume Gaussian normal distribution for the additive noise vectors $\mathbf{n}_k$ with possibly unknown variance.

The shift invariance property allows the operators C and E_k to commute. Hence, (1) can be rewritten as

$$\mathbf{f}_k = DE_kC\mathbf{x} + \mathbf{n}_k, \qquad 1 \leq k \leq p. \tag{2}$$

Equation (2) and Figure 1 motivate our two-step approach to superresolution. First, using the LR data frame samples $\mathbf{f}_k$, $1 \leq k \leq p$, we interpolate for $C\mathbf{x}$, the blurred version of the original HR image. Next, we deconvolve the blur C to obtain an estimate for $\mathbf{x}$. Deconvolution has been a thoroughly studied problem, and several robust techniques are available. The rest of the paper will therefore mostly address the interpolation step of the algorithm.

There is inherent structure and regularity in the grid of LR sampling points for superresolution. In Figure 1, if pixel values from all frames are considered together, the data are irregularly sampled. However, for each frame, data points are sampled on a rectangular grid. This special case of irregular sampling is called interlaced sampling [15]. Our interpolation method takes advantage of this

sampling structure. The method is based on the multiresolutional basis fitting reconstruction (MBFR) method, described in the paper by Ford and Etter [6]. In particular, we extend the work of Ford and Etter to two-dimensional (2D) interlaced sampling grids.

1.1. Previous work

Superresolution reconstruction from multiple frames is a relatively new class of restoration problems. Most techniques proposed for superresolution fall into one of three main categories: frequency domain, iterative spatial domain, and interpolation-restoration. In this paper, we will examine the last class.

Sauer and Allebach [14] were the first to consider superresolution as an interpolation problem with nonuniformly sampled data. They used a projection onto convex sets (POCS) algorithm to reconstruct the unknown values. Namely, they considered

$$\mathcal{F}^{(l+1)} = P_n \cdots P_2 P_1 \mathcal{F}^{(l)}, \tag{3}$$

where $\mathcal{F}^{(l)}$ is the lth approximate of the ideal HR image $\mathcal{F}$, and P_i are projection operators that correct for errors between $\mathcal{F}^{(l)}$ and $\mathcal{F}$ and impose band-limitedness constraints. The solution to the fixed-point iteration (3) is their estimate to $\mathcal{F}$.

Aizawa et al. [1] also modeled superresolution as an interpolation problem with nonuniform sampling and used a formula related to the Shannon sampling theorem to estimate values on an HR grid. The work of both [1] and [14] ignored the effect of sensor blurring. Tekalp et al. [18] later extended these algorithms to include blurring and sensor noise and proposed the additional restoration step for the interpolation algorithms. Ur and Gross [19] considered Papoulis' generalized multichannel sampling theorem for interpolating values on a higher-resolution grid. Because light detectors are not ideal lowpass filters, some high-frequency information about the scene is represented in the image in aliased form. Papoulis' theorem reconstructs this aliased high-frequency content by taking a properly weighted sum of the spectral information from the LR frames. Shekarforoush and Chellappa [15] extended Papoulis' theorem for merging of nonuniform samples of multiple channels into high-resolution data.

Previous work did not consider the implications of the 2D interlaced sampling structure on the computational complexity of the resulting algorithms. The computational complexity for existing algorithms for 2D data is squared that for 1D data. As we will show, by exploiting sampling regularity, the computational burden for our algorithm does not drastically increase for 2D data. In fact, computational complexity of the algorithm for 2D data is only twice that for 1D problems.

The outline of the rest of the paper is as follows. We briefly review multiresolution analysis with orthogonal wavelets in Section 2. In Section 3, we describe our 1D and 2D wavelet interpolation method for interlaced data. Section 4 discusses

implementation and complexity issues for the algorithm. Section 5 shows interpolation and superresolution experiments demonstrating the effectiveness of our techniques. We conclude with some comments in Section 6.

2. Multiresolution analysis with orthonormal wavelets

The fundamental concept behind wavelet theory is the decomposition of signals into components at different scales or resolutions. The advantage of this decomposition is that signal trends at different scales can be isolated and studied. Global trends can be examined at coarser scales, whereas local variations are better analyzed at fine scales. This section will present a brief summary of orthonormal wavelet multiresolution analysis of 1D and 2D signals. We will only review essential ideas necessary for the material in later sections. For more detailed treatments of wavelets, the reader is referred to the excellent books by Strang and Nguyen [17] and Mallat [10].

2.1. Multiresolution analysis for 1D signals

Let $L^2(\mathbf{R})$ be the vector space of square-integrable 1D signals $f(t)$. There exists a sequence of nested approximation subspaces $\mathbf{V}_j$, $j \in \mathbf{Z}$, and a scaling function $\phi(t)$ satisfying the following requirements [17]:

(i) $\mathbf{V}_j \subset \mathbf{V}_{j+1}$ and $\bigcap_{j\in\mathbf{Z}} \mathbf{V}_j = \{0\}$ and $\overline{\bigcup_{j\in\mathbf{Z}} \mathbf{V}_j} = L^2(\mathbf{R})$
(ii) $f(t) \in \mathbf{V}_j \Leftrightarrow f(2t) \in \mathbf{V}_{j+1}$
(iii) $f(t) \in \mathbf{V}_0 \Leftrightarrow f(t-k) \in \mathbf{V}_0$
(iv) $\mathbf{V}_0$ has an orthonormal basis $\{\phi(t-k)\}$

Because the set $\{\phi(t-k)\}$ is an orthonormal basis for $\mathbf{V}_0$, dilations and translations of $\phi(t)$, $\{\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)\}_{k\in\mathbf{Z}}$, form an orthonormal basis for $\mathbf{V}_j$. Furthermore, because $\mathbf{V}_0 \subset \mathbf{V}_1$, the scaling function $\phi(t)$ satisfies the following two-scale dilation equation:

$$\phi(t) = \sqrt{2}\sum_k c_k \phi(2t-k), \tag{4}$$

for some set of expansion coefficients c_k.

For a function $f(t) \in L^2(\mathbf{R})$, the projection $f_j(t)$ of $f(t)$ onto the subspace $\mathbf{V}_j$ represents an approximation of that function at scale j. The approximation becomes more accurate as j increases. The difference in successive approximations $g_j(t) = f_{j+1}(t) - f_j(t)$ is a detail signal that lives in a wavelet subspace $\mathbf{W}_j$. In fact, we can decompose the approximation space $\mathbf{V}_{j+1}$ as

$$\mathbf{V}_{j+1} = \mathbf{V}_j \oplus \mathbf{W}_j. \tag{5}$$

Equation (5) shows us one of the fundamental reasons why wavelets have been used so successfully in signal processing. Signals can be neatly broken down into

a coarse approximation signal and a fine detail signal. Any $f(t) \in L^2(\mathbf{R})$ can be written as a sum of its approximate at some scale J along with the subsequent detail components at scale J and higher. Hence,

$$L^2(\mathbf{R}) = \mathbf{V}_J \oplus \bigoplus_{j \geq J} \mathbf{W}_j. \tag{6}$$

As in the approximation spaces case, the wavelet spaces $\mathbf{W}_j$ are spanned by a set of orthonormal basis functions $\{\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)\}_{k \in \mathbf{Z}}$, which are dilations and translations of a single wavelet function $\psi(t)$. Furthermore, the wavelet function satisfies the wavelet equation

$$\psi(t) = \sqrt{2} \sum_k d_k \phi(2t - k), \tag{7}$$

for some set of expansion coefficients d_k.

By equation (6), we can expand any function $f(t) \in L^2(\mathbf{R})$ as follows:

$$f(t) = \sum_{k \in \mathbf{Z}} a_{J,k}\phi_{J,k}(t) + \sum_{j \geq J} \sum_{k \in \mathbf{Z}} b_{j,k}\psi_{j,k}(t), \tag{8}$$

where

$$a_{J,k} = \int f(t)\phi_{J,k}(t)\,dt$$

$$b_{j,k} = \int f(t)\psi_{j,k}(t)\,dt$$

are the expansion coefficients for $f(t)$.

2.2. *Multiresolution analysis for 2D images*

The wavelet model in the previous section for 1D signals can be extended to 2D images. We describe in this subsection a separable multiresolution analysis of $L^2(\mathbf{R}^2)$ studied by Meyer [11] and Mallat [10]. Given a multiresolution analysis $(\mathbf{V}_j^{(1)})_{j \in \mathbf{Z}}$ of $L^2(\mathbf{R})$, a set of nested subspaces $(\mathbf{V}_j^{(2)})_{j \in \mathbf{Z}}$ forms a multiresolution approximation of $L^2(\mathbf{R}^2)$ with each vector space $\mathbf{V}_j^{(2)}$ being a tensor product of identical 1D approximation spaces

$$\mathbf{V}_j^{(2)} = \mathbf{V}_j^{(1)} \otimes \mathbf{V}_j^{(1)}. \tag{9}$$

Furthermore, the scaling function $\Phi(t, s)$ for the 2D multiresolution subspaces can be decomposed as

$$\Phi(t, s) = \phi(t)\phi(s), \tag{10}$$

where $\phi(t)$ is the 1D scaling function of the multiresolution analysis $(\mathbf{V}_j^{(1)})_{j \in \mathbf{Z}}$. The set of functions

$$\Phi_{j,k,l}(t, s) = \phi_{j,k}(t)\phi_{j,l}(s), \qquad j, k, l \in \mathbf{Z} \tag{11}$$

is an orthonormal basis for $\mathbf{V}_j^{(2)}$. The 2D wavelet subspaces $\mathbf{W}_j^{(2)}$ are generated by three wavelets to capture detail information in the horizontal, vertical, and diagonal directions:

$$\Psi^h(t,s) = \psi(t)\phi(s) \tag{12}$$

$$\Psi^v(t,s) = \phi(t)\psi(s) \tag{13}$$

$$\Psi^d(t,s) = \psi(t)\psi(s). \tag{14}$$

The corresponding orthonormal wavelet basis for $\mathbf{W}_j^{(2)}$ is the set

$$\Psi^h_{j,k,l}(t,s) = \psi_{j,k}(t)\phi_{j,l}(s), \tag{15}$$

$$\Psi^v_{j,k,l}(t,s) = \phi_{j,k}(t)\psi_{j,l}(s), \tag{16}$$

$$\Psi^d_{j,k,l}(t,s) = \psi_{j,k}(t)\psi_{j,l}(s), \qquad j,k,l \in \mathbf{Z}. \tag{17}$$

Analogous to the 1D case, any image $f(t,s) \in L^2(\mathbf{R}^2)$ can be expanded as a sum of its approximate image at some scale J in $\mathbf{V}_J^{(2)}$ along with subsequent detail components at scale J and higher.

$$\begin{aligned} f(t,s) = &\sum_{k,l\in\mathbf{Z}} a_{J,k,l}\Phi_{J,k,l}(t,s) + \sum_{j\geq J}\sum_{k,l\in\mathbf{Z}} b^h_{j,k,l}\Psi^h_{j,k,l}(t,s) \\ &+ \sum_{j\geq J}\sum_{k,l\in\mathbf{Z}} b^v_{j,k,l}\Psi^v_{j,k,l}(t,s) + \sum_{j\geq J}\sum_{k,l\in\mathbf{Z}} b^d_{j,k,l}\Psi^d_{j,k,l}(t,s), \end{aligned} \tag{18}$$

with

$$a_{J,k,l} = \int\int f(t,s)\Phi_{J,k,l}(t,s)\,dt\,ds$$

$$b^h_{J,k,l} = \int\int f(t,s)\Psi^h_{J,k,l}(t,s)\,dt\,ds$$

$$b^v_{J,k,l} = \int\int f(t,s)\Psi^v_{J,k,l}(t,s)\,dt\,ds$$

$$b^d_{J,k,l} = \int\int f(t,s)\Psi^d_{J,k,l}(t,s)\,dt\,ds$$

The first term on the right-hand side of (18) represents the coarse-scale approximation to $f(t,s)$. The second term represents the detail component in the horizontal direction, the third and fourth the detail components in the vertical and diagonal directions, respectively.

3. Wavelet interpolation of interlaced data

This section will describe our interpolation technique for interlaced data. We use the expansion formulas (8) and (18) to first estimate for the wavelet coefficients. Using these estimates, we interpolate for the function values at the HR grid points.

3.1. *Interpolation for nonuniformly sampled 1D signals*

We first consider the case of nonuniformly sampled 1D signals. Suppose that we have a function $f(t)$ for which we would like to compute M uniformly distributed values, say, at $t = 0, 1, \ldots, M-1$. We are given P nonuniformly sampled data points of $f(t)$ at $t = t_0, t_1, \ldots, t_{P-1}$, $0 \le t_i < M$, where typically, $P < M$.

We take the unit-time spacing grid to be resolution level $\mathbf{V}_0$. By repeated application of (5), we can decompose $\mathbf{V}_0$ in the following fashion:

$$\mathbf{V}_0 = \mathbf{V}_J \oplus \bigoplus_{j=J}^{-1} \mathbf{W}_j, \qquad J \le -1. \tag{19}$$

Hence, we can separate $f(t) \in \mathbf{V}_0$ into its approximation and detail components and further expand these components in the orthonormal bases of $\mathbf{V}_J$ and $\{\mathbf{W}_j\}_{-1 \ge j \ge J}$:

$$f(t) = f_J(t) + \sum_{j=J}^{-1} g_j(t), \qquad f_J(t) \in \mathbf{V}_J, \quad g_j(t) \in \mathbf{W}_j \tag{20}$$

$$= \sum_k a_{J,k}\phi_{J,k}(t) + \sum_{j=J}^{-1}\sum_k b_{j,k}\psi_{j,k}(t), \tag{21}$$

$$a_{J,k} = \int f(t)\phi_{J,k}(t)\,dt$$

$$b_{j,k} = \int f(t)\psi_{j,k}(t)\,dt.$$

Substituting in the values of the sampled data, we have a set of P linear equations

$$f(t_i) = \sum_k a_{J,k}\phi_{J,k}(t_i) + \sum_{j=J}^{-1}\sum_k b_{j,k}\psi_{j,k}(t_i), \qquad i = 0, \ldots, P-1. \tag{22}$$

Suppose that $[0, N]$ is the support interval for $\phi(t)$ and let $t_{\max} = \max_i t_i$ and $t_{\min} = \min_i t_i$. In the first summation on the right-hand side of (22), only finitely many terms are nonzero because $\phi_{J,k}(t_i) = \phi(2^J t_i - k)$ is nonzero if and only if $2^J t_i - k$ is in the support interval for $\phi(t)$, i.e.,

$$0 \le 2^J t_i - k \le N. \tag{23}$$

Therefore,

$$-N + \lceil 2^J t_{\min} \rceil \le k \le \lfloor 2^J t_{\max} \rfloor. \tag{24}$$

Similar arguments can be made for the wavelet basis functions $\psi_{j,k}(t_i)$. Let $S_J = \{-N + \lceil 2^J t_{\min} \rceil, \ldots, \lfloor 2^J t_{\max} \rfloor\}$ be the set of shifts with nonzero contribution on the right-hand side of (22). We can now rewrite (22) as

$$f(t_i) = \sum_{k \in S_J} a_{J,k}\phi_{J,k}(t_i) + \sum_{j=J}^{-1}\sum_{k \in S_j} b_{j,k}\psi_{j,k}(t_i), \qquad i = 0, \ldots, P-1, \tag{25}$$

which, in vector form, becomes

$$\mathbf{f} = G_J \mathbf{a}_J + \sum_{j=J}^{-1} H_j \mathbf{b}_j, \tag{26}$$

where

$$\mathbf{f} = (f(t_i))_{i=0,\ldots,P-1}, \qquad \mathbf{a}_J = (a_{J,k})_{k\in S_J}, \qquad \mathbf{b}_j = (b_{j,k})_{k\in S_j},$$
$$G_J = (\phi_{J,k}(t_i))_{i=0,\ldots,P-1}^{k\in S_J}, \qquad H_j = (\psi_{j,k}(t_i))_{i=0,\ldots,P-1}^{k\in S_j}.$$

To construct G_J and H_j, we need to know basis function values at sampling points $\{t_i\}$. For most wavelet bases, there are no closed-form expressions for basis functions $\{\phi_{J,k}(t), \psi_{j,k}(t)\}$. However, basis function values at dyadic points can be calculated efficiently by recursion. At scale K and for scaling function with support on $[0, N]$, the set of dyadic points is defined to be $\mathcal{D}_K^N = \{0, 1/2^K, \ldots, N - 1/2^K\}$. We choose K large enough so that the set of sampling points $\{2^J t_i - k, 2^j t_i - k\}$ can be well approximated by a subset of $\mathcal{D}_K^N$.

From Equation (26), we can approximate the coarse-scale approximation coefficients $\mathbf{a}_J$ by ignoring the detail components and considering just

$$\mathbf{f} \approx G_J \mathbf{a}_J. \tag{27}$$

Ford and Etter [6] recommend that J be chosen so that the system above can be solved in a least-squares sense; that is, $P > \lfloor 2^J t_{\max} \rfloor - \lceil 2^J t_{\min} \rceil + N + 1$. The scale chosen is dependent on the total number of sample points, the interval spanned by these points, and the support size of the scaling and wavelet functions. Because (27) is an approximation, we solve for a *regularized* least-squares estimate in the wavelet domain

$$\hat{\mathbf{a}}_J = \operatorname{argmin}_{\mathbf{a}_J} \|\mathbf{f} - G_J \mathbf{a}_J\|_2^2 + \lambda \|\mathbf{a}_J\|_2^2, \tag{28}$$

or equivalently,

$$\hat{\mathbf{a}}_J = (G_J^T G_J + \lambda I)^{-1} G_J^T \mathbf{f} \tag{29}$$

for some regularization parameter λ. The regularization parameter λ plays a balancing role in (28) and (29). If λ is too large, the solution obtained will be too far away from the original system we wish to solve. If λ is too small, noise effects are exacerbated in the solution of the under-regularized system. The least-squares estimate $\hat{\mathbf{a}}_J$ of the coefficients yields a coarse-scale estimate of $\mathbf{f}$, denoted by $\hat{\mathbf{f}}_J$,

$$\hat{\mathbf{f}}_J = G_J \hat{\mathbf{a}}_J. \tag{30}$$

The difference between $\mathbf{f}$ and $\hat{\mathbf{f}}_J$ can then be used to approximate the wavelet coeffcients $\mathbf{b}_J$

$$\begin{aligned} \mathbf{g}_J &= \mathbf{f} - \hat{\mathbf{f}}_J \\ &= \mathbf{f} - G_J \hat{\mathbf{a}}_J \\ &\approx H_J \mathbf{b}_J. \end{aligned} \tag{31}$$

Because the number of nonzero coefficients $b_{J,n}$ is the same as the number of nonzero coefficients $a_{J,n}$, equation (31) can also be solved in the least-squares sense for a regularized estimate of $\mathbf{b}_J$. In general, the regularization parameter is chosen to be small for the coarse-scale approximation and large for the fine-scale detail because the signal-to-noise ratio tends to be smaller in the fine scale. In fact, an estimate for λ can be computed using prior information or a statistical model for the wavelet coefficients (cf. [16], [3]). The desired values of $f(t)$ at the HR grid points $t = 0, 1, \ldots, M-1$ can then be computed using the estimated coefficients:

$$f(t) \approx \sum_{k \in S_J} \hat{a}_{J,k} \phi_{J,k}(t) + \sum_{k \in S_J} \hat{b}_{J,k} \psi_{J,k}(t), \qquad t = 0, 1, \ldots, M-1. \quad (32)$$

For wavelet superresolution, the data is sampled nonuniformly but in a recurring manner. This type of sampling is called nonuniform recurring sampling or interlaced sampling [15]. More specifically, we are given sampled data on an LR grid in terms of "frames," which are sets of data points separated by a uniform shift. Let r be the resolution enhancement factor, m be the number of data points per frame, and n be the number of given frames. The available samples are

$$\{f(\epsilon_i), f(r+\epsilon_i), f(2r+\epsilon_i), \ldots, f((m-1)r+\epsilon_i)\},$$
$$0 \le \epsilon_i < r, \qquad i = 1, \ldots, n.$$

Given these mn sample points, we would like to reconstruct values of $f(t)$ for the HR grid points $t = 0, 1, \ldots, mr-1$.

Following (27) and putting the data in vector form, we get the following set of equations to solve for the coarse-scale coefficients $\mathbf{a}_J$:

$$\mathbf{f}^{(i)} \approx G_J^{(i)} \mathbf{a}_J, \qquad i = 1, \ldots, n, \quad (33)$$

where

$$\mathbf{f}^{(i)} = (f(pr+\epsilon_i))_{p=0,\ldots,m-1}, \qquad G_J^{(i)} = \left(\phi_{J,k}(pr+\epsilon_i)\right)_{p=0,\ldots,m-1}^{k \in S_J}.$$

We estimate the wavelet coefficients $\mathbf{b}_J$ and the values of $f(t)$ on the HR grid in the same fashion described previously.

3.2. *Interpolation for interlaced 2D images*

In image superresolution, the data frames are LR rectangular grids of sample points. Let h, w denote the height and width (in units of pixels) of an LR frame. The set of available data is then

$$\left\{f(pr+\epsilon_{i_t}, qr+\epsilon_{i_s})\right\}, \qquad 0 \le \epsilon_{i_t}, \qquad \epsilon_{i_s} < r,$$
$$p = 0, \ldots, h-1, \qquad q = 0, \ldots, w-1, \qquad i = 1, \ldots, n.$$

From these nhw sample points on LR grids, we would like to reconstruct values of $f(t,s)$ on HR grid points $\{(t,s) | t = 0, \ldots, hr-1, s = 0, \ldots, wr-1\}$. Analogous

to the 1D case, we substitute in sample values of $f(t,s)$ to obtain a set of linear equations and solve a least-squares system for the coarse-scale coefficients;

$$f(pr+\epsilon_{i_t}, qr+\epsilon_{i_s}) \approx \sum_{k\in S_{J_t}} \sum_{l\in S_{J_s}} a_{J,k,l}\Phi_{J,k,l}(pr+\epsilon_{i_t}, qr+\epsilon_{i_s}) \tag{34}$$

$$= \sum_{k\in S_{J_t}} \sum_{l\in S_{J_s}} a_{J,k,l}\phi_{J,k}(pr+\epsilon_{i_t})\phi_{J,l}(qr+\epsilon_{i_s}). \tag{35}$$

In matrix form, the double sum above can be written as a Kronecker product of 1D wavelet transform matrices

$$\mathbf{f}^{(i)} \approx (G_{J_t}^{(i)} \otimes G_{J_s}^{(i)})\mathbf{a}_J, \tag{36}$$

where $\mathbf{f}^{(i)}$ is the vector with the pixel values of the ith frame reordered rowwise, $\mathbf{a}_J$ is the vector of unknown coarse-scale coefficients, and the entries $G_{J_t}^{(i)}$, $G_{J_s}^{(i)}$ are basis function values at sampling points of frame i along the horizontal and vertical directions, respectively. Proceeding as in the 1D case, we solve (36) for a regularized least-squares estimate $\hat{\mathbf{a}}_J$ of $\mathbf{a}_J$. The difference between $\mathbf{f}^{(i)}$ and its coarse-scale estimate $(G_{J_t}^{(i)} \otimes G_{J_s}^{(i)})\hat{\mathbf{a}}_J$ can next be used to estimate the horizontal detail coefficients $\mathbf{b}_J^h$:

$$\mathbf{g}_J^{(i)} = \mathbf{f}^{(i)} - (G_{J_t}^{(i)} \otimes G_{J_s}^{(i)})\hat{\mathbf{a}}_J \tag{37}$$

$$\approx (G_{J_t}^{(i)} \otimes H_{J_s}^{(i)})\mathbf{b}_J^h. \tag{38}$$

Continuing as before, the residual is then used to calculate $\mathbf{b}_J^v$ and $\mathbf{b}_J^d$. The choice of scale J makes a crucial difference in reconstruction quality. We pick the finest scale J so that the number of sample values is more than the number of unknown coefficients in (36) and (38).

4. Implementation and computational complexity

We discuss an efficient implementation and the computational complexity of our interpolation approach for interlaced data in this section. We first consider the 1D interlaced frames equation (33). The regularized least-squares solution can be expressed as

$$\hat{\mathbf{a}}_J = \left(\begin{bmatrix} G_J^{(1)T} \cdots G_J^{(n)T} \end{bmatrix} \begin{bmatrix} G_J^{(1)} \\ \vdots \\ G_J^{(n)} \end{bmatrix} + \lambda I \right)^{-1} \begin{bmatrix} G_J^{(1)T} \cdots G_J^{(n)T} \end{bmatrix} \begin{bmatrix} \mathbf{f}^{(1)} \\ \vdots \\ \mathbf{f}^{(n)} \end{bmatrix} \tag{39}$$

$$= \left(\sum_{i=1}^{n} G_J^{(i)T} G_J^{(i)} + \lambda I \right)^{-1} \sum_{i=1}^{n} G_J^{(i)T} \mathbf{f}^{(i)}. \tag{40}$$

The equation above can be solved most efficiently by an iterative method such as the conjugate gradient method (cf. [2]). We only need to compute matrix-vector

products involving $\sum_{i=1}^{n} G_J^{(i)^T} G_J^{(i)} + \lambda I$ and not its explicit inverse. Furthermore, because of the finite support of $\phi(t)$, $G_J^{(i)}$ have a banded structure that may be further exploited. We can derive similar expressions for wavelet detail coefficients. For 2D interlaced images, the regularized least-squares estimate for the coarse-scale wavelet coefficients from equation (36) is

$$\hat{\mathbf{a}}_J = \left(\left[\left(G_{J_t}^{(1)} \otimes G_{J_s}^{(1)}\right)^T \cdots \left(G_{J_t}^{(n)} \otimes G_{J_s}^{(n)}\right)^T \right] \begin{bmatrix} G_{J_t}^{(1)} \otimes G_{J_s}^{(1)} \\ \vdots \\ G_{J_t}^{(n)} \otimes G_{J_s}^{(n)} \end{bmatrix} + \lambda I \right)^{-1} \cdot \left[\left(G_{J_t}^{(1)} \otimes G_{J_s}^{(1)}\right)^T \cdots \left(G_{J_t}^{(n)} \otimes G_{J_s}^{(n)}\right)^T \right] \begin{bmatrix} \mathbf{f}^{(1)} \\ \vdots \\ \mathbf{f}^{(n)} \end{bmatrix}. \tag{41}$$

Recalling the following properties of the Kronecker product:

(i) $(A \otimes B)^T = A^T \otimes B^T$.
(ii) $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$.
(iii) $(A \otimes B)\text{reshape}(V) = \text{reshape}(AVB^T)$, where reshape$(\cdot)$ reorders the entries of a matrix in rowwise order into vector format,

and applying these properties to equation (41), we have

$$\hat{\mathbf{a}}_J = \left(\sum_{i=1}^{n} \left(G_{J_t}^{(i)^T} G_{J_t}^{(i)}\right) \otimes \left(G_{J_s}^{(i)^T} G_{J_s}^{(i)}\right) + \lambda I \right)^{-1} \sum_{i=1}^{n} \left(G_{J_t}^{(i)^T} \otimes G_{J_s}^{(i)^T}\right) \mathbf{f}^{(i)} \tag{42}$$

$$= \left(\sum_{i=1}^{n} \left(G_{J_t}^{(i)^T} G_{J_t}^{(i)}\right) \otimes \left(G_{J_s}^{(i)^T} G_{J_s}^{(i)}\right) + \lambda I \right)^{-1} \sum_{i=1}^{n} \left(G_{J_t}^{(i)^T} F^{(i)} G_{J_s}^{(i)}\right), \tag{43}$$

where $F^{(i)}$ denotes the ith frame in matrix form. Similar forms for wavelet directional detail coefficients estimates can be derived in the same manner. Analogous to the 1D case, equation (43) can be solved most efficiently by an iterative method. Matrix-vector products involving the system matrix $\sum_{i=1}^{n} \left(G_{J_t}^{(i)^T} G_{J_t}^{(i)}\right) \otimes \left(G_{J_s}^{(i)^T} G_{J_s}^{(i)}\right) + \lambda I$ can take advantage of computational properties of the Kronecker product. We quantify the computational complexity of solving (40) and (43) in more detail in the following discussion.

The computational burden for the method is comprised of two main components. The first component is the cost of constructing the matrices of wavelet basis functions evaluated at the sampling points. In the 1D case, at scale K and with wavelet support of size N, the cost for generating the matrices of sampled wavelet basis functions is $O(N^2 2^{(K+1)})$. Our 2D wavelet basis is a separable basis, so the construction of the matrices of sampled wavelet basis functions only costs twice as much as in the 1D case.

The second component of computational burden is the least-squares solution of (40) and (43). We use conjugate gradient for an iterative solution. Complexity

per iteration of conjugate gradient can be based on the cost of a matrix-vector product with the system matrix. In the 1D case, the system matrix is

$$\sum_{i=1}^{n} {G_J^{(i)}}^T G_J^{(i)} + \lambda I,$$

with $G_J^{(i)}$ a matrix approximately $m \times (2^J mr + N + 1)$ in dimensions, where m is the number of samples per frame. The computational complexity for this is $\mathcal{O}(nm(2^J mr + N))$. Analogously, the system matrix for interlaced interpolation in two dimensions is

$$\sum_{i=1}^{n} \left({G_{J_t}^{(i)}}^T G_{J_t}^{(i)}\right) \otimes \left({G_{J_s}^{(i)}}^T G_{J_s}^{(i)}\right) + \lambda I,$$

where $G_{J_t}^{(i)}$ and $G_{J_s}^{(i)}$ are matrices approximately $h \times (2^J hr + N + 1)$ and $w \times (2^J wr + N + 1)$ in dimensions, respectively. The variables h and w denote the height and width of an LR frame, respectively, in units of samples or pixels. The computational complexity for a matrix vector product with the system matrix is $\mathcal{O}(nh(2^J hr + N) + nw(2^J wr + N))$. By taking advantage of the interlacing structure and the Kronecker product representation, the computational cost for our interpolation approach only doubles for the 2D case as compared to the 1D case.

5. Numerical experiments

This section presents numerical results for interpolation experiments with 1D signals and superresolution experiments for 2D images.

5.1. Wavelet interpolation experiments for 1D signals

In this first set of experiments, we will use the wavelet techniques described above to interpolate values of a 1D signal. We start with an original signal of length 168. The signal is then blurred with a Gaussian point spread function with variance 1 and downsampled by a factor of 3 to generate three LR frames, each with 56 sample points. We keep only one of those LR frames, leading to a severely underdetermined interpolation problem. The given frame has sample values of $f(t)$ at $t = 0, 3, \ldots, 165$. The algorithm attempts to reconstruct the signal at time $t = 0, 1, \ldots, 167$. Figure 2 displays the result of wavelet reconstruction using a Daubechies DB6 filter [4].

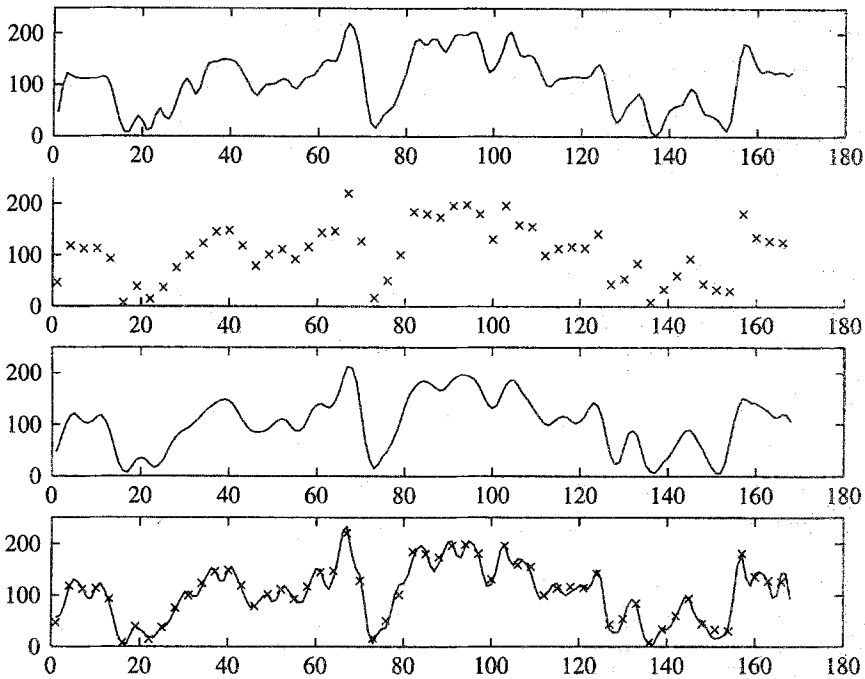

Figure 2. Top to bottom. The first plot shows the original signal. The second contains the available data samples. The third plot displays the result of a coarse-scale approximation at scale $J = -2$. The last graph plots the final result against the given samples.

5.2. Wavelet superresolution experiments for 2D images

The setup for the first 2D superresolution experiment is similar to the 1D experiments. A 172×172 pixel HR image (upper left corner of Figure 3) is blurred with a Gaussian point spread function of variance 1 and downsampled by a factor of 4 to simulate 16 LR frames. We randomly choose 10 of those LR frames, each of size 43×43 pixels, again leading to a severely undersampled superresolution problem. The resolution enhancement factor is 4. The wavelet interpolation-restoration process first interpolates for blurred values at the HR grid points. An estimate for the original HR image is obtained by deconvolving the interpolated values with the known blur. Figure 3 shows the result of wavelet superresolution for our test 2D sequence using Daubechies DB4 filter interpolation in combination with Tikhonov regularized restoration (cf. [12]).

The LR forward looking infrared (FLIR) images in our second superresolution experiment are provided courtesy of Brian Yasuda and the FLIR research group in the Sensors Technology Branch, Wright Laboratory, WPAFB, Ohio. Each image is 64×64 pixels, and a resolution enhancement factor of 5 is sought. The objects in the scene are stationary, and 16 frames are acquired by controlled movements

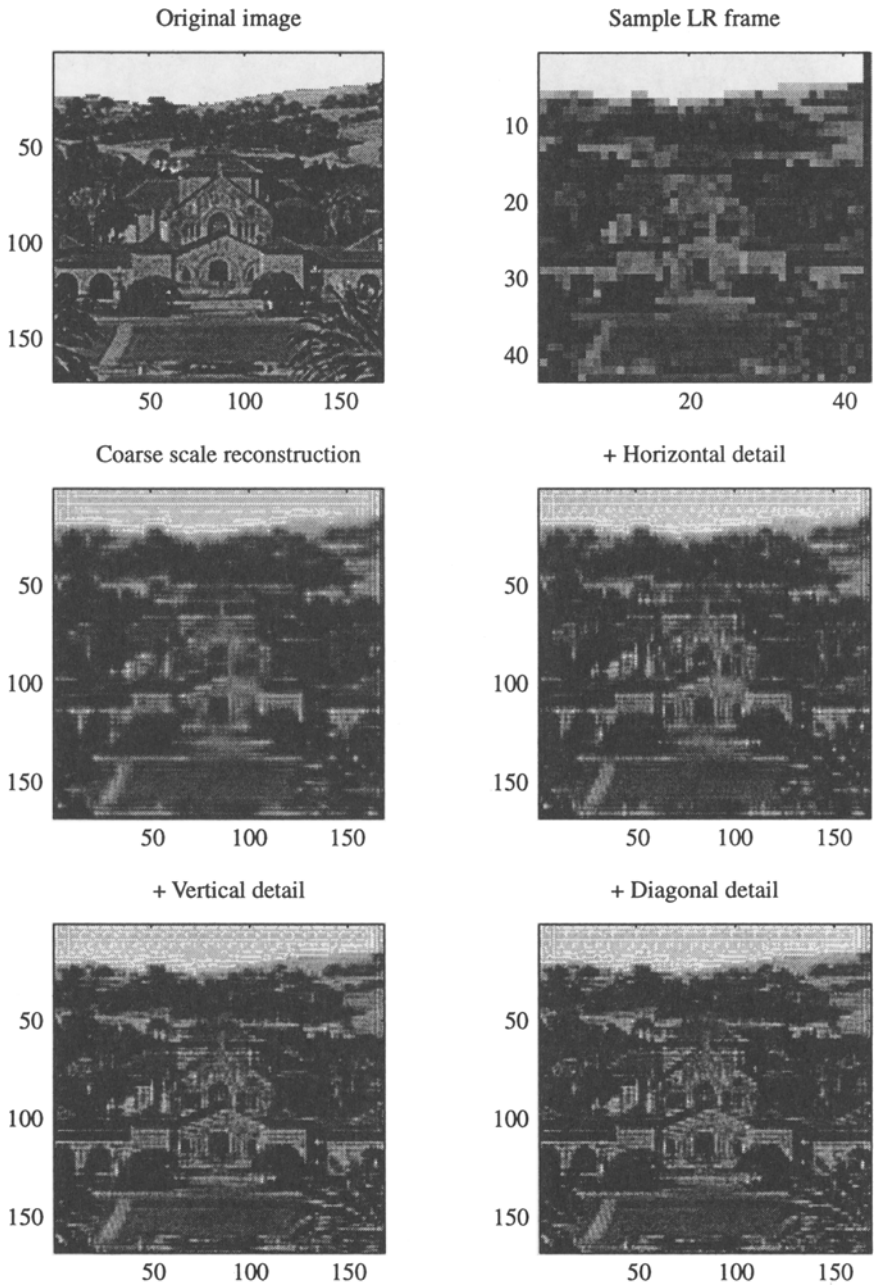

Figure 3. The first (upper left) display is the original Stanford HR image. The second (upper right) shows a sample LR frame. Subsequent images are the coarse-scale approximation plus various incremental levels of detail refinements.

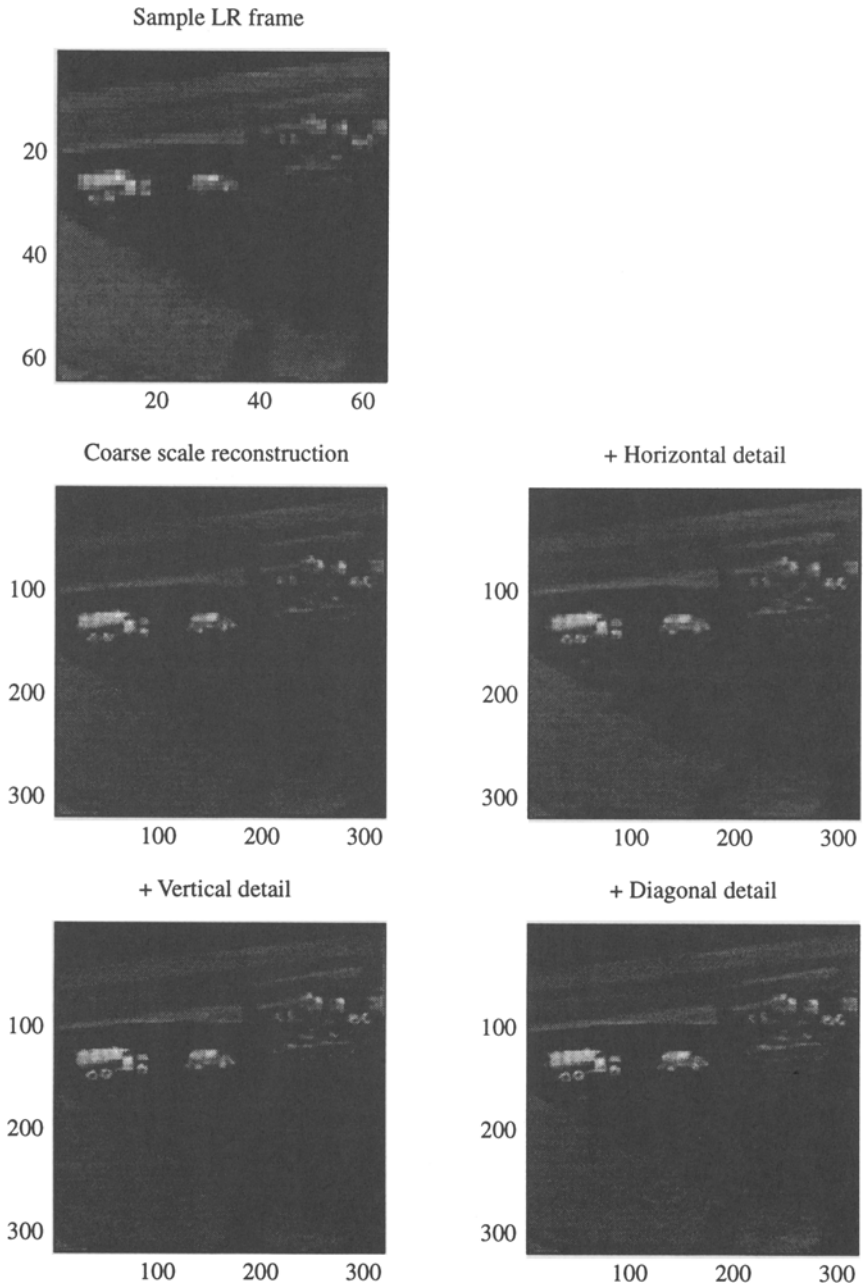

Figure 4. The first (upper left) display is a sample FLIR LR frame. The subsequent images are the coarse-scale approximation plus various incremental levels of detail refinements.

of an FLIR imager described in [7]. Figure 4 contains the results of wavelet superresolution for the FLIR test sequence using Daubechies DB4 filter interpolation, along with regularized restoration.

6. Summary

This paper presents a new wavelet interpolation-restoration method for image superresolution. In contrast to previous interpolation-restoration approaches, our method exploits the interlacing structure of the sampling grid in superresolution. Using a separable orthonormal wavelet basis for 2D images, we derive a wavelet decomposition using Kronecker products. As a result, the computational properties of the Kronecker products allow efficient calculation of the wavelet coefficients. Computational complexity of our method applied to 2D interlaced data increases only by a factor of 2 compared to that for 1D data.

References

[1] K. Aizawa, T. Komatsu, and T. Saito, Acquisition of very high resolution images using stereo cameras, in *Proceedings SPIE Visual Communications and Image Processing '91*, Boston, MA, pp. 318–328, November 1991.

[2] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, New York, 1994.

[3] M. Crouse, R. Nowak, and R. Baraniuk, Wavelet-based statistical signal processing using hidden Markov models, *IEEE Trans. Signal Process.*, 46(4), 886–902, April 1998.

[4] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, New York, 1992.

[5] M. Elad, Super-resolution reconstruction of images, Ph. D. thesis, The Technion-Israel Institute of Technology, December 1996.

[6] C. Ford and D. Etter, Wavelet basis reconstruction of nonuniformly sampled data, *IEEE Trans. Circuits and Systems II*, 45(8), 1165–1168, August 1998.

[7] R. Hardie, K. Barnard, and E. Armstrong, Joint MAP registration and high-resolution image estimation using a sequence of undersampled images, *IEEE Trans. Image Process.*, 6(12), December 1997.

[8] W. Lukosz, Optical systems with resolving power exceeding the classical limit, *J. Opt. Soc. Am.*, 56(11), 1463–1472, 1966.

[9] W. Lukosz, Optical systems with resolving power exceeding the classical limit II, *J. Opt. Soc. Am.*, 57(7), 932–941, 1967.

[10] S. Mallat, A theory for multiresolution in signal decomposition: The wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intellig.*, 11(7), 674–683, July 1989.

[11] Y. Meyer, Principe d'incertitude, bases Hilbertiennes et algebres d'operateurs, in *Bourbaki Seminar*, vol. 662, Paris, pp. 1985–1986.

[12] N. Nguyen, Numerical techniques for image superresolution, Ph. D. thesis, Stanford University, May 2000.

[13] N. Nguyen, P. Milanfar, and G. Golub, Blind superresolution with generalized cross-validation using Gauss-type quadrature rules, in *Proceedings of the 33rd Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, October 1999.

[14] K. Sauer and J. Allebach, Iterative reconstruction of band-limited images from non-uniformly spaced samples, *IEEE Trans. Circuits and Systems*, CAS-34, 1497–1505, 1987.

[15] H. Shekarforoush and R. Chellappa, Data-driven multi-channel super-resolution with application to video sequences, *J. Opt. Soc. Am. A*, 16(30), 481–492, March 1999.

[16] E. Simoncelli, Statistical models for images: Compression, restoration and synthesis, in *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, November 1997.

[17] G. Strang and T. Nguyen, *Wavelet and Filter Banks*, Wellesley-Cambridge Press, Wellesley, MA, 1997.

[18] A. Tekalp, M. Ozkan, and M. Sezan, High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration, in *Proceedings ICASSP '92*, vol. 3, San Francisco, CA, pp. 169–172, March 1992.

[19] H. Ur and D. Gross, Improved resolution from subpixel shifted pictures, *CVGIP: Graphical Models and Image Processing*, 54(2), 181–186, March 1992.