Nonuniform sampling, image recovery from sparse data and the discrete sampling theorem

Leonid P. Yaroslavsky,[1] Gil Shabat,[1] Benjamin G. Salomon,[1] Ianir A. Ideses,[1] and Barak Fishbain[1,2,*]

[1]*Department of Physical Electronics, Faculty of Engineering, Tel Aviv University, Tel Aviv 69978, Israel*
[2]*Department of Industrial Engineering and Operational Research, University of California, Berkeley, Berkeley, California 94720-1777, USA*
**Corresponding author: barak@berkeley.edu*

Received August 18, 2008; accepted November 11, 2008;
posted November 24, 2008 (Doc. ID 100195); published February 19, 2009

In many applications, sampled data are collected in irregular fashion or are partly lost or unavailable. In these cases, it is necessary to convert irregularly sampled signals to regularly sampled ones or to restore missing data. We address this problem in the framework of a discrete sampling theorem for band-limited discrete signals that have a limited number of nonzero transform coefficients in a certain transform domain. Conditions for the image unique recovery, from sparse samples, are formulated and then analyzed for various transforms. Applications are demonstrated on examples of image superresolution and image reconstruction from sparse projections. © 2009 Optical Society of America

OCIS codes: 000.4430, 070.2025, 100.2000, 100.3020, 110.3010.

1. INTRODUCTION

Images and other signals are usually represented in computers in the form of their samples taken on a uniform sampling grid. However, in many applications, sampled data are collected in irregular fashion and/or it may frequently happen that some samples of the regular sampling grid are lost or unavailable. In these cases, it is necessary to convert irregularly sampled signals to regularly sampled ones or to restore missing data. Typical examples are filtering "salt-and-pepper"-type noise in images transmitted through communication channels with error detection coding, reconstruction of surface profiles in geophysics and in optical metrology, restoration of image sequences acquired in the presence of camera or object vibrations or through a turbulent medium, and image superresolution from multiple chaotically sampled frames, to name a few.

There are two approaches to treat this problem. One approach is empirical in nature and is based on simplistic numerical interpolation procedures such as, for instance, Shepard's interpolation by means of a weighted summation of known samples in close vicinity of sought samples with weights inversely proportional to the distance between them [1]. A review of these methods can be found in [2].

The second approach is based on generalizations of the classical Whittaker–Kotelnikov–Shannon sampling theory to nonuniform sampling. In this approach, it is assumed that the available signal samples are obtained from a continuous signal that belongs to a certain approximation subspace $\boldsymbol{M}$ (subspaces of band-limited signals, splines subspaces, etc.) of the parent Hilbert space (usually, L^2 Hilbert space of finite energy functions) and it is required that the interpolation procedure has to determine a continuous signal that satisfies two constraints: (1) the interpolated signal has to belong to the subspace $\boldsymbol{M}$ and (2) its available samples have to be preserved. Conditions for the existence and uniqueness of the solution are dependent on the signal model (underlying approximation subspace) and on the set of given samples. For the band-limited case, Landau proved that a necessary and sufficient condition for the unique reconstruction of a continuous band-limited 1D signal with bandwidth W from its irregularly spaced samples is that the density of its samples should exceed the Nyquist rate $1/W$ [3]. It is also shown that this condition is necessary for $\mathbf{D}$-dimensional signals with band-limited Fourier spectrum. These results have been generalized to other shift-invariant subspaces by Aldourbi and Grochenig [4]. A comprehensive presentation of this approach can be found in [5].

An attractive alternative approximation model is associated with spline subspaces [6]. However, due to their localized nature, their use for the recovery of large gaps in data is limited. A practical numerical algorithm for interpolation and approximation of 2D signals, based on multilevel B-splines, is suggested by Wolberg and colleagues [7]. The algorithm approximates 2D functions from sparse data by an iterative procedure based on lattice control points. At each iteration, the values of available samples are preserved (if possible) or approximated. At the next iteration, a denser grid of control points is created to approximate the reconstruction error, and the process continues iteratively. A similar spline-based algorithm, which uses nonuniform splines for interpolation, was suggested by Margolis and Eldar [8].

All of the mentioned methods are theoretically oriented at the approximation of continuous signals, specified by their sparse samples. There are also publications that consider discrete models. However, those publications treat only various special cases. Fereira considers discrete signal recovery from sparse data in the assumption of sig-

nal band-limitation in the discrete Fourier transform (DFT) domain [9]. Hasan and Marvasti suggest a method for recovery of discrete signals suffering from missing data during data transmission using error-detecting coding. For signal recovery, they suggest using the discrete cosine transform (DCT) -domain band-limitation assumption [10]. In [11], the problem of nonuniform sampling in the Fourier domain in multidimensional polar coordinates, is addressed in connection with image reconstruction from projections. In another publication, Averbuch and Zheludev discuss image reconstruction from projections with omissions using biorthogonal wavelet overcomplete basis functions [12].

In this paper, we suggest a general framework for recovery of discrete signals that originate from continuous signals, from incomplete sets of their samples. The basis of this framework is the following assumptions:

- Continuous signals are represented in computers by their samples. In sampling a continuous signal, say, $a(x)$, the physical coordinates of the samples are known with a certain accuracy. The ratio $N=X/\Delta x$ of the signal support interval X and the sample position accuracy Δx defines the signal's regular uniform sampling grid with N sampling positions. If all these N samples were known, they would be sufficient for representing the continuous signal.
- Available are $K<N$ samples of this signal, taken at irregular positions of the signal regular sampling grid.
- The goal of the processing is to generate, out of this incomplete set of K samples, a complete set of N signal samples in such a way as to secure the most accurate, in a certain metric, approximation of the discrete signal that corresponds to the signal that would be obtained if the continuous signal it is intended to represent were densely sampled in all N positions. For the certainty, we will use L_2 metrics.

The mathematical foundation of the framework is provided by the discrete sampling theorem for band-limited discrete signals that have only a few nonzero coefficients in their representation over a certain orthogonal basis. This theorem is introduced in Section 2. The rest of the paper is as follows. In Section 3 we discuss the validity of the assumptions put in the basis of the presented approach. In Section 4 we briefly describe algorithms for signal recovery from sparse sampled data. In Section 5, properties of certain transforms that are specifically relevant for signal recovery from sparse data are analyzed, and experimental illustrations of precise signal reconstruction from sparse data are provided. Finally, in Section 6 we discuss application issues and illustrate the discrete-sampling-theorem-based methodology of discrete signal recovery on the examples of image superresolution from multiple frames and image recovery from sparse projection data. Section 7 summarizes the paper.

2. DISCRETE SAMPLING THEOREM

Let A_N be a vector of N samples $\{a_k\}_{k=0,\ldots,N-1}$ that completely define a discrete signal, Φ_N be an $N\times N$ orthogonal transform matrix,

$$\Phi_N = \{\varphi_r(k)\}_{r=0,1,\ldots,N-1} \tag{2.1}$$

composed of basis functions $\varphi_r(k)$, and Γ_N be a vector of signal transform coefficients $\{\gamma_r\}_{r=0,\ldots,N-1}$ such that

$$\mathbf{A}_N = \Phi_N\Gamma_N = \left\{\sum_{r=0}^{N-1}\gamma_r\varphi_r(k)\right\}_{k=0,1,\ldots N-1} \tag{2.2}$$

Assume now that only $K<N$ signal samples $\{a_{\tilde{k}}\}_{\tilde{k}\in\tilde{\mathbf{K}}}$ are available, where $\tilde{\mathbf{K}}$ is a K-size nonempty subset of indices $\{0,1,\ldots,N-1\}$. These available K signal samples define a system of K equations:

$$\left\{a_k = \sum_{r=0}^{N-1}\gamma_r\varphi_r(k)\right\}_{k\in\tilde{\mathbf{K}}} \tag{2.3}$$

for signal transform coefficients $\{\gamma_r\}$ of certain K indices $\mathbf{r}$.

Select now a subset $\tilde{\mathbf{R}}$ of K transform coefficients indices $\{\tilde{r}\in\tilde{\mathbf{R}}\}$ and define a "**K***of***N**"-band-limited approximation $\hat{A}_N^{BL}$ to the signal A_N as

$$\hat{\mathrm{A}}_N^{BL} = \left\{\hat{a}_k = \sum_{\tilde{r}\in\tilde{R}}\gamma_{\tilde{r}}\varphi_{\tilde{r}}(k)\right\}. \tag{2.4}$$

Rewrite this equation in a more general form,

$$\hat{\mathrm{A}}_N^{BL} = \left\{\hat{a}_k = \sum_{r=0}^{N-1}\tilde{\gamma}_r\varphi_r(k)\right\}, \tag{2.5}$$

and assume that all transform coefficients with indices $r\notin\tilde{\mathbf{R}}$ are set to zero:

$$\tilde{\gamma}_r = \begin{cases}\gamma_r, & r\in\tilde{R}\\ 0, & otherwise\end{cases}. \tag{2.6}$$

Then the vector $\tilde{\mathbf{A}}_K$ of available signal samples $\{a_{\tilde{k}}\}$ can be expressed in terms of the basis functions $\{\varphi_r(k)\}$ of transform Φ_N as

$$\tilde{\mathrm{A}}_K = \mathbf{K}of\mathbf{N}_\Phi\cdot\tilde{\Gamma}_K = \left\{a_{\tilde{k}} = \sum_{\tilde{r}\in\tilde{R}}\gamma_{\tilde{r}}\varphi_{\tilde{r}}(\tilde{k})\right\}, \tag{2.7}$$

where $K\times K$ subtransform matrix $\mathbf{K}of\mathbf{N}_\Phi$ is composed of samples $\varphi_{\tilde{r}}(\tilde{k})$ of the basis functions with indices $\{\tilde{r}\in\tilde{\mathbf{R}}\}$ for signal sample indices $\tilde{k}\in\tilde{K}$, and $\tilde{\Gamma}_K$ is a vector composed of the corresponding subset $\{\gamma_{\tilde{r}}\}$ of complete signal transform coefficients $\{\gamma_r\}$. This subset of the coefficients can be found as

$$\tilde{\Gamma}_K = \{\tilde{\gamma}_r\} = \mathbf{K}of\mathbf{N}_\Phi^{-1}\cdot\tilde{\mathbf{A}}_K, \tag{2.8}$$

provided that matrix $\mathbf{K}of\mathbf{N}_\Phi^{-1}$ inverse to the matrix $\mathbf{K}of\mathbf{N}_\Phi$ exists, which in general is conditioned for a specific transform by positions $\tilde{k}\in\tilde{K}$ of available signal samples and by the selection of the subset $\{\tilde{R}\}$ of transform basis functions.

By virtue of Parceval's relationship for orthonormal transforms, the band-limited signal $\hat{A}_N^{BL}$ approximates the complete signal A_N with mean square error (MSE):

$$MSE = \|A_N - \hat{A}_N\|^2 = \sum_{k=0}^{N-1} |a_k - \hat{a}_k|^2 = \sum_{r \notin R} |\gamma_r|^2. \quad (2.9)$$

This error can be minimized by an appropriate selection of the K basis functions of the subtransform $\mathbf{K}of\mathbf{N}_\Phi$. In order to minimize the error, one must know the energy compaction ordering of basis functions of the transform Φ_N. If in addition one knows, for a class of signals, a transform that features the best energy compaction in the smallest number of transform coefficients, one can, by selection of this transform, secure the best band-limited approximation of the signal $\{a_k\}$ for the given subset $\{\tilde{a}_k\}$ of its samples.

In this way, we arrive at the following discrete sampling theorem which can be formulated in these two statements:

Statement 1. For any discrete signal of N samples defined by its $K \leqslant N$ sparse and not necessarily regularly arranged samples, its band-limited, in terms of certain transform Φ_N, approximation defined by Eq. (2.5) can be obtained with mean square error defined by Eq. (2.9) provided that positions of the samples secure the existence of the matrix $\mathbf{K}of\mathbf{N}_\Phi^{-1}$ inverse to the subtransform matrix $\mathbf{K}of\mathbf{N}_\Phi$ that corresponds to the band limitation. The approximation error can be minimized by using a transform with the best energy compaction property.

Statement 2. Any signal of N samples that is known to have only $K \leqslant N$ nonzero transform coefficients for certain transform Φ_N (Φ_N-transform band-limited signal) can be fully recovered from exactly K of its samples provided that positions of the samples secure the existence of the matrix $\mathbf{K}of\mathbf{N}_\Phi^{-1}$ inverse to the subtransform matrix $\mathbf{K}of\mathbf{N}_\Phi$ that corresponds to the band limitation.

3. VALIDTY OF THE ASSUMPTIONS

The applicability of the above results depends on the validity of the assumption that band-limited, in a certain basis, approximation of signals is an acceptable solution in image recovery. We believe that this assumption is validated by a consensus in the signal processing and image processing community regarding signal compression, where such transforms as DCT and certain wavelets are known for their very good energy compaction properties for wide variety of signals in image and audio processing. These transforms are used successfully for compression by means of replacement of signals by their band-limited approximations. Recent advances in compressive sensing [13] also are based on the signal's "band-limitedness" assumption. Haar transform and Walsh transform were found to have good energy compaction properties for bi-level images such as drawings and documents. An important application, in which the assumption of image band-limitedness is supported by the physical reality, is computed tomography, in which slice projections can very frequently be regarded in the inverse Radon transform domain as band-limited signals, because outer parts of slices are usually known to be empty.

4. ALGORITHMS FOR SIGNAL RECOVERY FROM SPARSE NONUNIFORMLY SAMPLED DATA

Implementation of signal recovery from sparse nonuniformly sampled data according to Eq.(2.8) requires matrix inversion, which is generally a very computationally demanding procedure, though for some transforms, such as DFT, DCT, Walsh, Haar, and others that feature fast Fourier transform (FFT)-type algorithms, pruned versions of these algorithms can be used [14–16]. In applications, one can also be satisfied with signal reconstruction with a certain limited accuracy and apply for the reconstruction of a simple iterative reconstruction procedure of the Gershberg–Papoulis [17] type shown in the flow diagram of Fig. 1. We used this algorithm in the experiments reported in this paper. One can find a review of other iterative and noniterative algorithmic options in [9].

5. ANALYSIS OF TRANSFORMS

A. Discrete Fourier Transform

Consider the $\mathbf{K}of\mathbf{N}_{DFT}^{LP}$-trimmed $\mathbf{DFT}_N$ matrix:

$$\mathbf{K}of\mathbf{N}_{DFT}^{LP} = \left\{ \exp\left(i2\pi \frac{\tilde{k}\tilde{r}_{LP}}{N} \right) \right\} \quad (5.1)$$

that corresponds to a DFT $\mathbf{K}of\mathbf{N}$ low-pass band-limited signal. Due to the complex conjugate symmetry of DFT of real signals, $\boldsymbol{K}$ has to be an odd number, and the set of frequency domain indices of $\mathbf{K}of\mathbf{N}_{DFT}$-low-pass band-limited signals in Eq. (5.1) is defined as

$$\tilde{r}_{LP} \in \tilde{R}_{LP} = \{[0, 1, \ldots, (K-1)/2, N-(K-1)/2, \ldots, N-1]\}. \quad (5.2)$$

For such a case, the following theorems hold:

Theorem 1. $\mathbf{K}of\mathbf{N}$-low-pass DFT band-limited signals of $\boldsymbol{N}$ samples with only $\boldsymbol{K}$ nonzero low-frequency DFT coefficients can be precisely recovered from exactly $\boldsymbol{K}$ of their samples taken in arbitrary positions.

Proof. As follows from Eqs. (2.3)–(2.8), the theorem is proved if the matrix $\mathbf{K}of\mathbf{N}_{DFT}^{LP}$ is invertible. A matrix is invertible if its determinant is nonzero. In order to check whether the determinant of the matrix $\mathbf{K}of\mathbf{N}_{DFT}$ is nonzero, permute the order of columns of the matrix as follows:

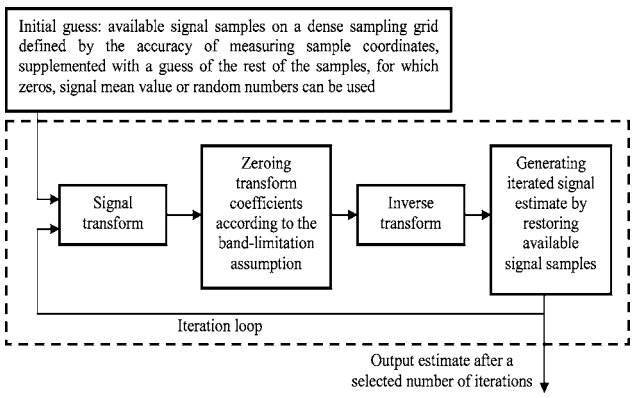

Fig. 1. Flow diagram of the iterative signal recovery procedure.

$$\tilde{\tilde{r}} \in \tilde{\tilde{R}} = \{[N-(K-1)/2, \ldots, N-1, 0, 1, \ldots, (K-1)/2]\} \tag{5.3}$$

and obtain the matrix

$$\mathbf{K}of\mathbf{N}_{DFT}^{DFTsh} = \left\{ \exp\left[i2\pi \frac{\tilde{k}\tilde{\tilde{r}}}{N} \right] \right\} = \left\{ \exp\left[i2\pi \frac{N-(K-1)/2}{N} \tilde{k} \right] \delta(\tilde{k} - \tilde{\tilde{\tilde{r}}}) \right\} \times \left\{ \exp\left[i2\pi \frac{\tilde{\tilde{\tilde{k}}}\tilde{\tilde{r}}}{N} \right] \right\}, \tag{5.4}$$

where

$$\tilde{\tilde{\tilde{r}}} \in \tilde{\tilde{\tilde{R}}} = \{[0, \ldots, K-1]\}. \tag{5.5}$$

The first matrix in this product of matrices is a diagonal matrix, which is obviously invertible. The second one is a version of Vandermonde matrices, which are also known to have a nonzero determinant if, as in our case, its ratios for each row are distinct [18].

As permutation of the matrix columns does not change the absolute value of its determinant, Eq. (5.4) implies that the determinant of **K***of***N**-trimmed $\boldsymbol{DFT}_N$ matrix of Eq. (5.1) is also nonzero for arbitrary set $\tilde{K}=\{\tilde{k}\}$ of positions of $\boldsymbol{K}$ available signal samples.

One can easily see that for DFT **K***of***N**-high-pass band-limited signals for which

$$\mathbf{K}of\mathbf{N}_{DFT}^{HP} = \left\{ \exp\left(i2\pi \frac{\tilde{k}\tilde{r}_{HP}}{N} \right) \right\}, \tag{5.6}$$

where

$$\tilde{r}_{HP} \in \tilde{R}_{HP} = \{[(N-K+1)/2, (N-K+3)/2, \ldots, (N+K-1)/2]\}, \tag{5.7}$$

a similar theorem holds

Theorem 2. **K***of***N**-high-pass DFT band-limited signals of $\boldsymbol{N}$ samples with only $\boldsymbol{K}$ nonzero high-frequency DFT coefficients can be precisely recovered from exactly $\boldsymbol{K}$ of their arbitrarily taken samples.

Note that due to the complex conjugate symmetry of DFT of real signals, $\boldsymbol{K}$ in this case has to be odd whatever $\boldsymbol{N}$ is.

Obviously, the above Theorems 1 and 2 can be extended to a more general case of signal DFT band limitation when indices $\{\tilde{r}\}$ of nonzero DFT spectral coefficients form arithmetic progressions with common difference other than one, such as, for instance,

$$\tilde{r}_{mLP} \in \tilde{R}_{mLP} = \left\{ 0, m, \ldots, m\frac{(K-1)}{2}, N-m\frac{(K-1)}{2}, \ldots, N-m\frac{(K-1)}{2}+\frac{(K+1)}{2} \right\}. \tag{5.8}$$

B. Discrete Cosine Transform

$\boldsymbol{N}$-point DCT of a signal is equivalent to 2N-point shifted discrete Fourier Transform (SDFT) with shift parameters (1/2,0) of a $2\boldsymbol{N}$-sample signal obtained from the initial one by its mirror reflection [16]. **K***of***N**-trimmed matrix of SDFT (1/2,0)

$$\mathbf{K}of\mathbf{N}_{SDFT} = \left\{ \exp\left(i2\pi \frac{(\tilde{k}+1/2)\tilde{r}}{2N} \right) \right\} \tag{5.9}$$

can be represented as a product

$$\mathbf{K}of\mathbf{N}_{SDFT} = \left\{ \exp\left(i2\pi \frac{\tilde{k}\tilde{r}}{2N} \right) \left\{ \exp\left(i\pi \frac{\tilde{r}}{2N} \right) \delta(k-r) \right\} \right\} = \mathbf{K}of\mathbf{N}_{DFT} \left\{ \exp\left(i\pi \frac{\tilde{r}}{2N} \right) \delta(k-r) \right\} \tag{5.10}$$

of a $2\boldsymbol{N}$-point DFT matrix and a diagonal matrix $\{\exp(i\pi\tilde{r}/2N)\delta(k-r)\}$. The latter one is invertible, and the invertibility of the **K***of***N**-trimmed $\boldsymbol{DFT}_{2N}$ matrix $\mathbf{K}of\mathbf{N}_{DFT}$ can be proved, for the above-described band limitations, as was done above for the DFT case. Therefore, theorems for DCT similar to those for DFT hold.

These theorems hold also for 2D DFT and DCT transforms provided that the band-limitation conditions are separable. The case of nonseparable band limitation requires further study. In the discussion of experiments that follows, we will compare separable and nonseparable band limitation in the DCT domain. Note that working in the DFT or the DCT domain results, in the case of low-pass band limitation, in signal discrete sinc-interpolation [19].

We illustrate the above reasoning by some simulation examples. The plots in Fig. 2 illustrate exact reconstruction of a DFT-band-limited signal (solid curue) for two cases, when all the signal samples are randomly placed within signal support [Fig. 2(a)] and when all available signal samples form a compact group [Fig. 2(b)]. Figure 2(c) illustrates restoration of the same signal with randomly placed samples by means of the iterative algorithm, while Fig. 2(d) shows standard deviation of signal restoration error as a function of the number of iterations. Note that the speed of convergence of the iterative algorithm depends heavily on the realization of sample positions and, for some samples, realizations of sample positions might be very slow.

Figures 3 and 4 illustrate precise restoration from sparse data of images band-limited in the DCT domain by a square (separable band limitation) and by 90° circle sector (a pie piece, inseparable band limitation). In these experiments, image restoration using a multilevel B-spline interpolation algorithm was used as a benchmark [7]. For the implementation of the algorithm, a code kindly provided by G. Wolberg was used.

The image presented in Fig. 3 is a 64×64 pixel test image, low-pass band-limited in the DCT domain by a square of 14×14 samples [Fig. 3(b)]. It has only $14\times14=196$ nonzero DCT spectral components out of the 64×64 ones. This image was sampled at 196 random positions obtained from a standard Matlab pseudorandom number generator. One can see from the figure that the

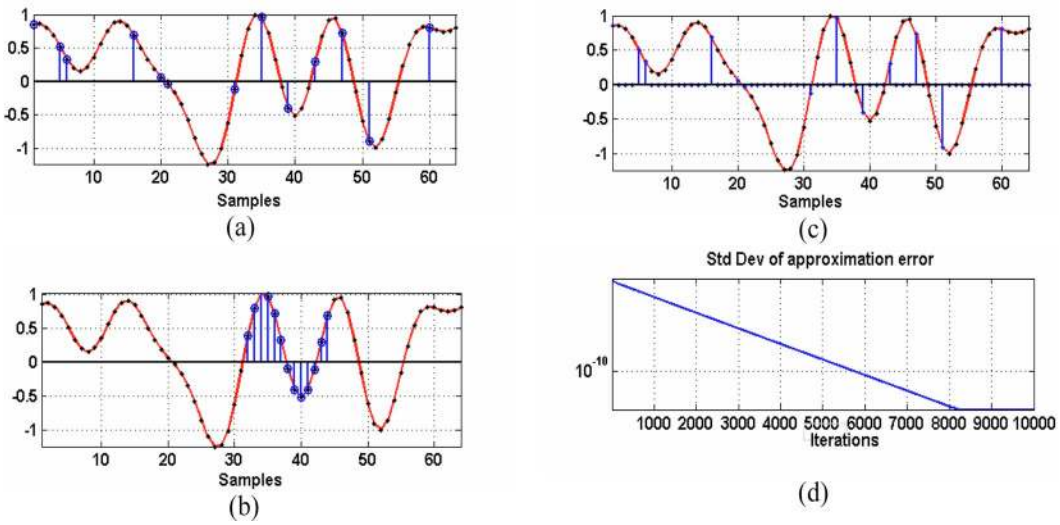

Fig. 2. (Color online) Restoration of a DFT low-pass band-limited signal by matrix inversion for the cases of (a) random and (b) compactly placed signal samples and (c) restoration by the iterative algorithm. Plot (d) shows the standard deviation of the signal restoration error as a function of the number of iterations. The experiment was conducted for test signal length of 64 samples; bandwidth of 13 frequency samples (~1/5 of the signal base band).

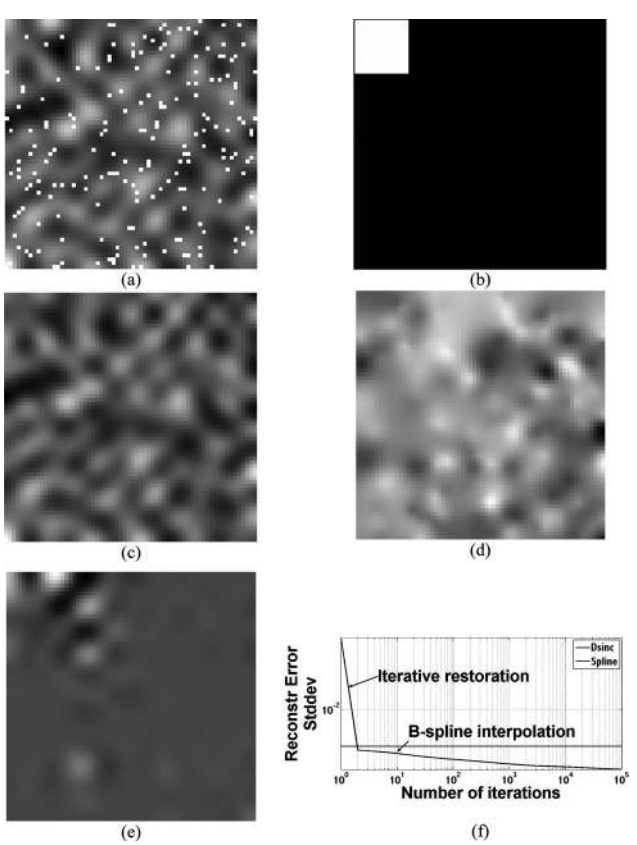

Fig. 3. Recovery of an image band limited in the DCT domain by a square: (a) initial image with 3136 "randomly" placed samples in positions shown by white dots; (b) the shape of the image spectrum in the DCT domain; (c) image restored by the iterative algorithm after 100,000 iterations with restoration peak signal-to-error standard deviation (PSNR) 4230; (d) image restored by B-spline interpolation with restoration PSNR 966; (e) iterative algorithm restoration error (white, large errors; black, small errors); (f) restoration error standard deviation versus number of iterations for the iterative algorithm and that for the B-spline interpolation.

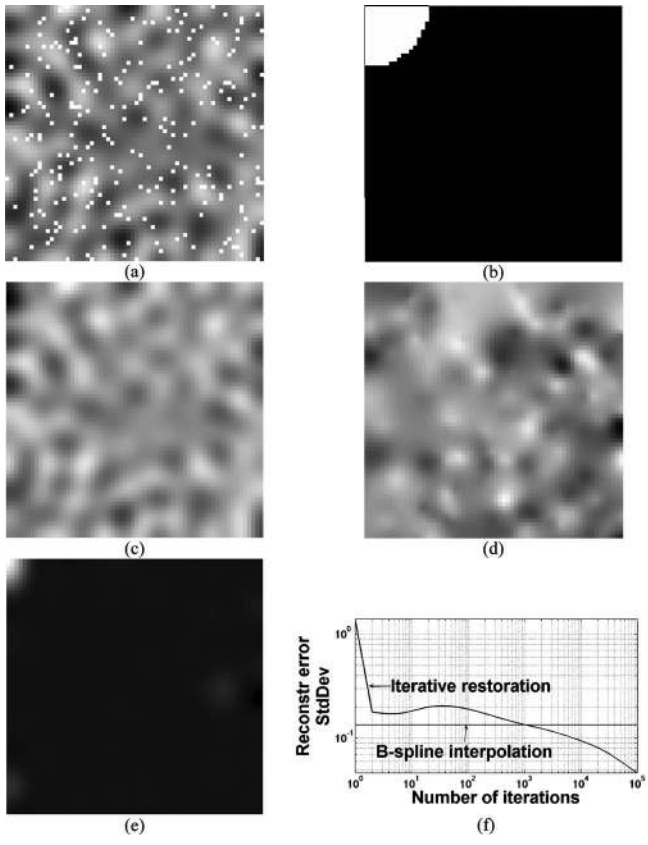

Fig. 4. Recovery of an image band limited in the DCT domain by a circle sector: (a) initial image with 3964 "randomly" placed samples in positions shown by white dots; (b) the shape of the image spectrum in the DCT domain; (c) image restored by the iterative algorithm after 100,000 iterations with restoration PSNR 21.5; (d) image restored by B-spline interpolation with restoration PSNR 7.42; (e) iterative algorithm restoration error (white, large errors; black, small errors); (f) restoration error standard deviation versus number of iterations of the iterative algorithm for the iterative algorithm and that for the B-spline interpolation.

iterative algorithm provides a quite accurate restoration of the initial image, though precise restoration may require a quite large number of iterations. An important peculiarity of the 2D case is that the convergence of the iteration is very nonuniform within the image. Usually, the restoration error rapidly becomes very small almost everywhere in the image, and only in some parts, where sample density happens to be low, do the restoration errors remain substantial and converge to zero quite slowly.

Image band-limitation by a square is separable, and, as was shown earlier, it does not impose any limitations on the positions of sparse samples. However, it is not isotropic. In the case of isotropic band limitation in the DCT domain by a circle sector (a pie piece), the situation is quite different. Experiments show that the speed of convergence of the iterative algorithm drops dramatically in this case. Hundreds of thousands of iterations are needed to make the overall standard deviation of the restoration error low enough, though again, restoration error remains substantial only in limited areas of the image. B-spline interpolation error is also high in this case, though it is more uniform over the image. The convergence speed of the iterative algorithm in the case of isotropic circle sector band limitation can be substantially improved if the number of available image samples exceeds the number of nonzero DCT spectral coefficients, which is redundant from the point of view of the discrete sampling theorem. This is illustrated in Fig. 4. The image presented in Fig. 4 is a 64×64 pixel test image low-pass band-limited in the DCT domain by a circle sector. It has 196 nonzero DCT spectral components, out of the 64×64 signal's samples, all located within a circle sector shown in white in Fig. 4(b). In contradistinction to the image in Fig. 3, this one was sampled at 248 random positions. The redundancy $248/196=1.27$ in the number of samples with respect to the number of nonzero spectral coefficients is approximately equal to the ratio of the area of the square to the area of the circle sector inscribed into this square. As one can see from Fig. 4(f), with such a redundancy iterative restoration converges much faster, though overall restoration error even after 100,000 iterations remains higher than that for the separable band limitation by a square illustrated in Fig. 3. The same holds for B-spline interpolation restoration, shown in Fig. 4(d). Once again, one can see that the convergence of the iterative algorithm is nonuniform over the image and that relatively large restoration error occurs only in a small area of the image where the density of available samples happens to be low.

In some applications, there is a natural and substantial redundancy in the number of available image samples with respect to the image's bandwidth. One such case is illustrated in Fig. 5, where an example of image restoration from its level lines is given. A 256×256 pixels image, shown in the figure, is band limited in the DCT domain by a circle sector and contains 302 nonzero spectral coefficients. The image was sampled in 6644 samples on a set of its level lines (8 levels), which resulted in 22-fold redundancy with respect to the image spectrum. As one can see from the figure, such a redundancy accelerated the convergence of the iterative algorithm very substantially and, after a few tens of iterations, enabled restoration

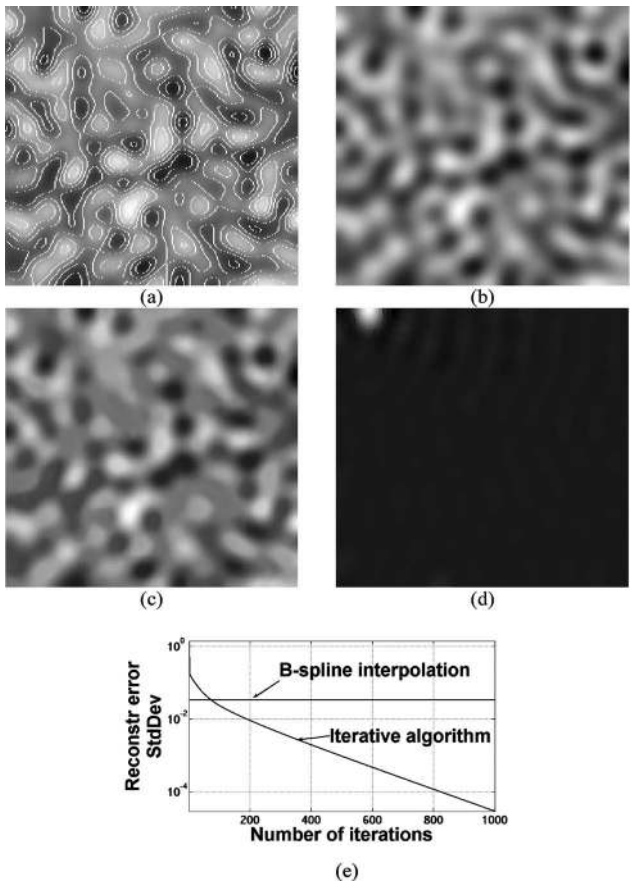

Fig. 5. Recovery of an image band-limited in the DCT domain by a circle sector from its level lines: (a) initial image with level lines (shown by white lines); (b) image restored by the iterative algorithm after 1,000 iterations with restoration PSNR 35,000 (note that the restoration error is concentrated in a small area of the image); (c) image restored by B-spline interpolation with restoration PSNR 29.4; (d) iterative algorithm restoration error (white, large errors; black, small errors); (e) restoration error standard deviation versus number of iterations for the iterative algorithm and that for the B-spline interpolation.

that is much superior to that provided by the B-spline interpolation.

C. Wavelets and Other Bases

The main peculiarity of wavelet bases is that their basis functions are most naturally ordered in terms of two components: scale and position within the scale. Scale index is analogous to the frequency index for DFT. Position index tells only of the shift of the same basis function within the signal on each scale. Therefore, band limitation for DFT translates to scale limitation for wavelets. Limitation in terms of position is trivial: it simply means that some parts of the signal are not relevant. Commonly, discrete wavelets are designed for signals whose length is an integer power of 2 ($N=2^n$). For such signals, there are $s\leqslant n$ scales and possible band limitations.

The simplest special case of wavelet bases is Haar basis. Signals with $N=2^n$ samples and with only K lower index nonzero Haar transform (the transform coefficients $\{K,\ldots,N-1\}$ are zero) are $(\tilde{s}=(\lfloor\log_2(K-1)\rfloor+1))$ band-limited, where $\lfloor x\rfloor$ is an integer part of x. Such signals are

piecewise constant within intervals between basis function zero crossings. The shortest intervals of the signal constancy contain $2^{n-\tilde{s}}$ samples. As one can see from Fig. 6(a), for any two samples that are located on the same interval, all Haar basis functions on this and lower scales have the same value. Therefore, having more than one sample per constant interval will not change the rank of the matrix **K***of***N**. The condition for perfect reconstruction is therefore to have at least one sample on each of those intervals.

For other wavelets as well as for other bases a general necessary, sufficient, and easily verified condition for the invertibility of **K***of***N** -trimmed transform submatrix is not known to the present authors. Standard linear algebra procedures for determining matrix rank can be used for testing invertibility of the matrix.

For Walsh basis functions, the index corresponds to the "sequency," or to the number of zero crossings of the basis function. The sequency carries a certain analogy to the signal frequency. Basis functions ordering according to their sequency, which is characteristic of Walsh transform, preserves for many real-life signals the property of decaying transform coefficients' energy with their index. Therefore, for Walsh transform the notion of low-pass band-limited signal approximation, similar to the one described in Subsection 5.A for DFT can be used. On the other hand, as one can see from Fig. 6(b), Walsh basis functions, similarly to the Haar basis function, can be characterized by the scale index, which specifies the shortest interval of signal constancy. Signals with $N=2^n$ samples and band limitation of K Walsh transform coefficients have the shortest signal-constancy intervals of $2^{n-\tilde{s}}$ samples, where $\tilde{s}=(\lfloor \log_2(K-1) \rfloor +1)$. A necessary condition for perfect reconstruction is to have K signal samples taken on different intervals. Unlike the Haar transform case, not all the intervals are needed to be sampled but only K intervals out of the total number of intervals. For a special case of K equal to a power of 2, there are K intervals, each of which has to be sampled to secure perfect reconstruction. This is the case when the reconstruction condition for Walsh transform is identical to that for Haar transform.

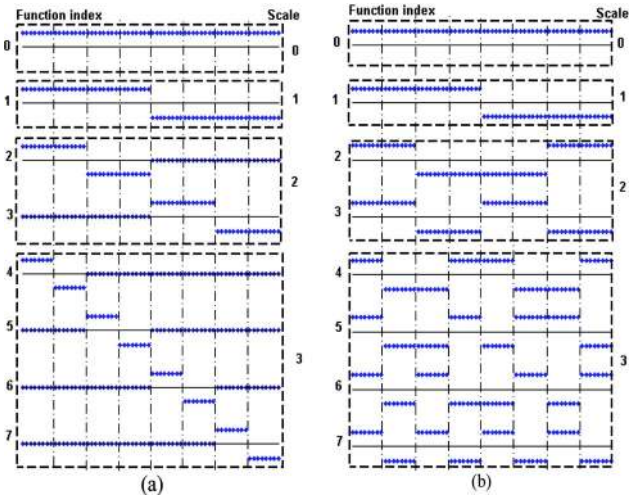

Fig. 6. (Color online) First eight basis functions of the 64-point (a) Haar and (b) Walsh transforms. Intervals of function constancy are outlined by dashed–dotted lines. Functions that belong to the same scale are outlined by dashed boxes.

Figure 7 illustrates the case of recovery of an image band limited in the Haar transform domain. Two examples are shown: arrangement of sparse samples, for which signal recovery is possible (a) and that for which signal is not recoverable (b). Note that when the Haar reconstruction is possible, it is reduced to the trivial nearest-neighbor interpolation.

An example of perfect reconstruction of Walsh-transform-domain band-limited signal of $N=512$ and band limitation $K=5$ is illustrated in Fig. 8. In this example, the resulting **K***of***N** Walsh matrix is

$$\mathbf{K}of\mathbf{N}^{Walsh}|_{K=5}=\begin{bmatrix} 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & 1 & 1 & -1 \end{bmatrix}, \quad (5.11)$$

and its rank is equal to 5. One should note that in this particular example, perfect reconstruction in the Haar transform domain is not possible since one of the shortest intervals of the signal constancy contains no samples.

6. APPLICATION EXAMPLES

A. Image Superresolution from Multiple Differently Sampled Video Frames

One of the potential applications of the above signal recovery technique is image superresolution from multiple video frames with chaotic pixel displacements due to atmospheric turbulence, camera instability, or similar random factors [20]. By means of elastic registration of a sequence of frames of the same scene, one can determine, for each image frame and with subpixel accuracy, pixel displacements caused by random acquisition factors. Using these data, a synthetic fused image can be generated by placing pixels from all available video frames in their proper positions on the correspondingly denser sampling grid according to their found displacements. In this process, some pixel positions on the denser sampling grid will remain unoccupied, especially when a limited number of

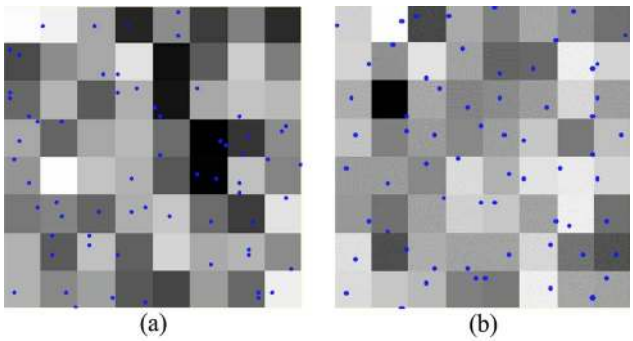

Fig. 7. (Color online) Two cases of sparse sampling of an image band-limited in the Haar transform: (a) not-recoverable case, (b) recoverable case (sample points are marked with dots). Image size 64×64 pixels; band limitation 8×8 (scale 3).

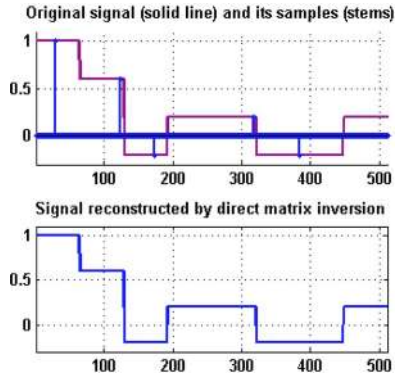

Fig. 8. (Color online) Example of perfect reconstruction in the Walsh domain.

image frames is fused. These missing pixels can then be restored using the above-described iterative band-limited interpolation algorithm.

In the implementation of the algorithm, the denser sampling grid of the fused image is formed accordingly to the subpixel accuracy, with which positions of pixels are measured in the sequence of turbulent frames. In our experiment, the size of the fused image sampling grid was 8× that of initial frames. The bandwidth limitation of the superresolved image depends on the spread of image samples acquired in the process of fusion and on the number of frames used for fusion. In our experiments, we set the final size of the fused-image sampling grid to be twice that of the original frames. The simulation result of iterative recovery of unavailable image samples is presented in Fig. 9, which shows (a) one of the low-resolution turbulent frames, (b) image fused from 50 frames, and (c) the result of iterative interpolation achieved after 50 iterations. The figure clearly demonstrates a substantial improvement of image resolution and quality.

B. Image Reconstruction from Sparse Projections in Computed Tomography

The discussed sparse-data recovery algorithm can find a straightforward application in tomography, where it frequently happens that a substantial part of the slices that surrounds the body slice is known to be an empty field. This means that slice projections (sinograms) are Radon transform band-limited functions. Therefore, whatever number of projections is available, a certain number of additional projections that are commensurable, according to the discrete sampling theorem, with the size of the slices empty zone can be obtained, and a corresponding

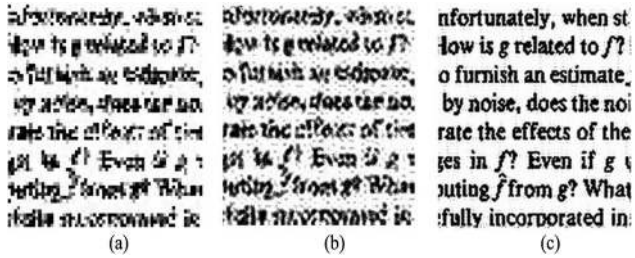

Fig. 9. Iterative image interpolation in the superresolution process: (a) a low-resolution frame, (b) image fused by elastic image registration from 50 frames, (c) the result of iterative interpolation of image (b) after 50 iterations.

resolution increase in the reconstructed images can be achieved using the described iterative band-limited reconstruction algorithm. Another option is recovery of projection data that might be missing because of sensor faults or for other reasons.

In order to demonstrate the applicability of the discrete sampling theorem for image recovery from sparse projections, one needs a discrete Radon transform and its algebraically exact inverse. While the theory defines the continuous Radon integral transform and its inverse, the discrete equivalent is not a trivial problem. In our experiments we used a stable forward and inverse Radon transform described in [21] and the code found in [22]. The applicability of the suggested framework for image reconstruction from sparse projections is illustrated in Fig. 10. By simple segmentation of the initial image shown in Fig. 10(a) it was found that the outer 55% of the image area is empty. Then the same percentage of projection samples selected randomly using the Matlab random number generator were zeroed, after which the iterative reconstruction algorithm was run. The results, shown in Fig. 10(c)–10(f), show that while direct image reconstruction with missing samples fails completely [Fig. 10(c)],

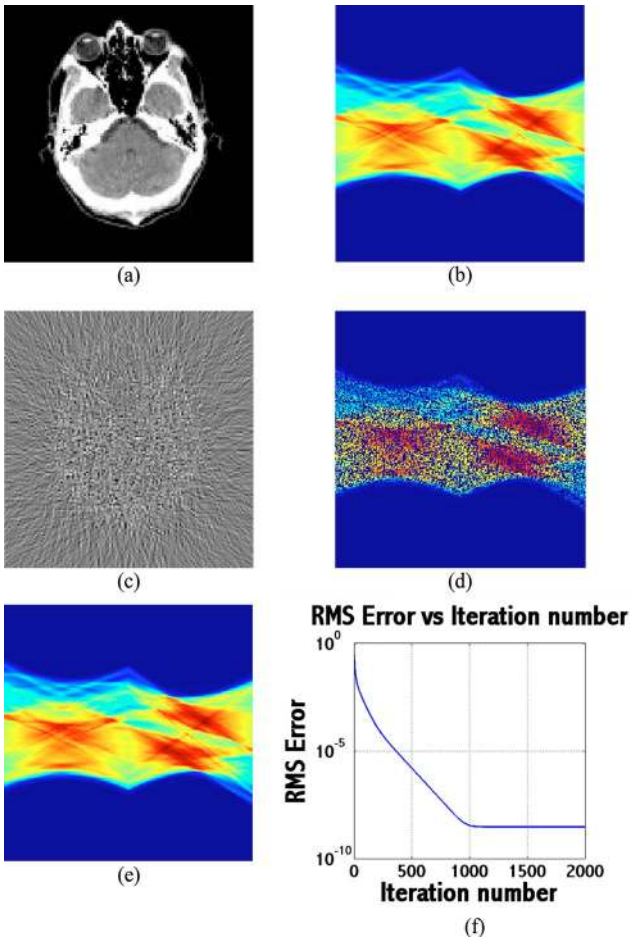

Fig. 10. (Color online) Recovery of missing samples of a sinogram: (a) original image and (b) its Radon transform (sinogram), (c) image reconstructed from the sinogram, (d) image corrupted by the loss of 55% of its randomly selected samples; (e) a sinogram recovered from (d) using the iterative band-limited interpolation algorithm, and (f) plot of the standard deviation of the slice reconstruction error as a function of the iteration number.

virtually perfect recovery of 55% of the missing samples of sinograms is possible with the iterative reconstruction algorithm after several hundreds of iterations.

Figure 11 illustrates that recovery of completely missing projections is also possible. Every second projection of the image shown in Fig. 10(a) was removed, and then all initial projections were recovered by the iterative algorithm, which made use of the fact that the outer 55% part of the image area is known to be empty. In this case the standard deviation of the reconstruction error is not as low as in the previous case, which perhaps can be attributed to the incomplete reversibility of the truncated Radon transforms. However, the achieved low reconstruction error of about 10^{-3} allows us to suggest that for such cases, when half or more of the image area is known to be empty, one can achieve image reconstruction with superresolution that corresponds to double the number of available image projections.

7. CONCLUSION

The paper addresses the problem of reconstruction of discrete signals from their irregular samples and recovery of missing data. Considering that positions of available signal samples are always specified with a certain accuracy that defines the maximal number of signal samples sufficient for signal representation, we suggest a new approach to optimal recovery of discrete signals from irregularly sampled or sparse data based on the discrete sampling theorem introduced in Section 2. The discrete sampling theorem refers to discrete signals band limited in the domain of a certain transform and states that **K***of***N** band-limited discrete signals of $\boldsymbol{N}$ samples, which have only $K \leqslant N$ nonzero transform coefficients, can be precisely recovered from their $\boldsymbol{K}$ sparse samples provided that positions of the available samples satisfy certain limitations depending on the transform. This theorem also provides a tool for optimal, in terms of root-mean-squared error, approximation of arbitrary discrete signals specified by their sparse samples with **K***of***N**-band-limited signals, given appropriate selection of the signal representation transform.

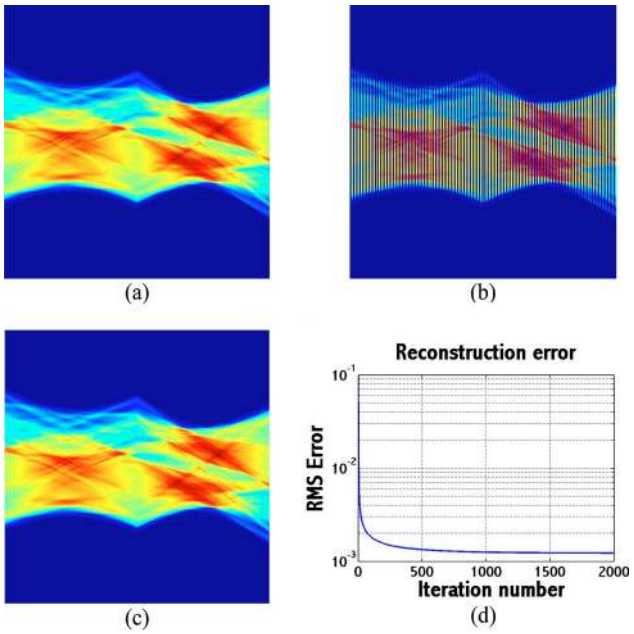

Fig. 11. (Color online) Recovery of missing image projections: (a) original projections (sinogram) of the test image of Fig. 10(a), (b) sinogram with every second projection removed, (c) sinogram recovered from (b) using the iterative interpolation algorithm, and (d) plot of the standard deviation of the image reconstruction error as a function of the iteration number.

Two algorithms for discrete sampling theorem based signal reconstruction are considered: direct matrix inversion and Gershberg–Papoulis iterative algorithm.

Properties of different transforms, such as discrete Fourier, discrete cosine, Haar, Walsh, and wavelet transforms, that are relevant to application of the discrete sampling theorem are discussed and, in particular, it is shown that precise reconstruction of 1D **K***of***N**-DFT band-limited and **K***of***N**-DCT band-limited signals is always possible from sparse samples regardless of sample positions on the signal dense grid. The same holds for two-dimensional signals, given separable band-limitation conditions. For nonseparable band limitation, such as limitation by a circle sector in the DCT domain, experimental evidence is obtained that exact image recovery may not be possible for arbitrarily placed samples and that a redundant number of samples is required.

Applications of the discrete-sampling-theorem-based approach to image recovery from sparse data are illustrated on examples of image superresolution from multiple randomly sampled frames and image reconstruction from sparsely sampled projections. For the latter case, it is shown that in applications where object slices contain areas that *a priori* are known to be empty, reconstruction of slice images from a given set of projections is possible with superresolution.

ACKNOWLEDGMENTS

The authors thank G. Wolberg, Department of Computer Science, City College of New York, USA, for his help with respect to the B-spline interpolation and A. Averbuch, School of Computer Science, Tel-Aviv University, Israel, for his help with respect to the discrete stable forward and inverse Radon transforms. This work was partially supported by Tampere International Center for Signal Processing, Tampere University of Technology (Tampere, Finland), by the Ran Naor Highway Safety Research Center Foundation at the Technion, Israel, and by the Domestic Nuclear Detection Office in the U.S. Department of Homeland Security (DHS) grant CBET-0736232.

REFERENCES

1. D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data," in *Proceedings of the 1968 ACM 23rd National Conference* (ACM, 1968), pp. 517–523.
2. S. K. Lodha and R. Franke, "Scattered data techniques for surfaces," in *Proceedings of the IEEE Conference on Scientific Visualization* (IEEE, 1997), Vol. 38, No. 157, pp. 181–200.
3. H. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," Acta Math. **117**, 37–52 (1967).

4. A. Aldroubi and K. Grochenig, "Non-uniform sampling and reconstruction in shift-invariant spaces," SIAM Rev. **43**, 585–620 (2001).
5. F. Marvasti, ed., *Nonuniform Sampling* (Kluwer Academic/Plenum, 2001).
6. M. Unser, "Splines: a perfect fit for signal and image processing," IEEE Signal Process. Mag. June 1999, pp. 22–38.
7. S. Lee, G. Wolberg, and S. Y. Shin, "Scattered data interpolation with multilevel B-splines," IEEE Trans. Visualization. Comput. Graphics **3**, 228–244 (1997).
8. E. Margolis and Y. C. Eldar, "Interpolation with non-uniform B-splines," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE, 2004), Vol. 2, pp. 577–580.
9. P. J. S. G. Ferreira, "Iterative and noniterative recovery of missing samples for 1-D band-limited signals," in *Nonuniform Sampling*, F. Marvasti, ed. (Kluwer Academic/Plenum, 2001), pp. 235–278.
10. M. Hasan and F. Marvasti, "Application of nonuniform sampling to error concealment," in *Nonuniform Sampling*, F. Marvasti, ed. (Kluwer Academic/Plenum, 2001), pp. 619–646.
11. A. Averbuch, R. Coifman, M. Israeli, I. Sidelnikov, and Y. Shkolinsky, "Irregular sampling for multi-dimensional polar processing of integral transforms," in *Advances in Signal Transforms: Theory and Applications*, J. Astola and L. Yaroslavsky, eds. (Hindawi, 2007), pp. 143–198.
12. A. Averbuch and V. Zheludev, "Wavelet and frame transforms originated from continuous and discrete splines," in *Advances in Signal Transforms: Theory and Applications*, J. Astola and L. Yaroslavsky, eds. (Hindawi, 2007), pp. 1–54.
13. Compressed sensing resources: http://www.dsp.ece.rice.edu/cs/. Last accessed December 18, 2008.
14. Y. Katiyi and L. Yaroslavsky, "Regular matrix methods for synthesis of fast transforms: general pruned and integer-to-integer transforms," in *Proceedings of IEEE International Workshop on Spectral Methods and Multirate Signal Processing* (IEEE, 2001), pp. 17–24.
15. Y. Katiyi and L. Yaroslavsky, "V/HS structure for transforms and their fast algorithms," in *Proceedings of the IEEE 3rd International Symposium on Signal Processing and Analysis* (IEEE, 2003), Vol. 1, pp. 482—487.
16. L. Yaroslavsky, *Digital Holography and Digital Signal Processing* (Kluwer Academic, 2004).
17. A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," IEEE Trans. Circuits Syst. **22**, pp. 735–742 (1975).
18. R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis* (Cambridge U. Press, 1991).
19. L. Yaroslavsky, "Fast discrete sinc-interpolation: a gold standard for image resampling," in *Advances in Signal Transforms: Theory and Application*, J. Astola and L. Yaroslavsky, eds., EURASIP Book Series on Signal Processing and Communications (Hindawi, 2007), pp. 337–405.
20. L. P. Yaroslavsky, B. Fishbain, G. Shabat, and I. Ideses, "Super-resolution in turbulent videos: making profit from damage," Opt. Lett. **32**, 3038–3040 (2007).
21. A. Averbuch, R. Coifman, D. Donoho, M. Israeli, and Y. Shkolinsky, "A framework for discrete integral transformations II—the 2D discrete Radon transform," SIAM J. Sci. Comput. (USA) **30**, 764–784 (2008).
22. Standford University, Statistics Department, David Donoho's homepage, http://www.stat.stanford.edu/~donoho/.