TABLE II
MEAN SQUARE ERROR FOR MVZS, SYMMETRY, AND COMPOSITE SELECTION SCHEMES

DATA	Compression	MVZS m.s.e		Symm. Selection m.s.e		Composite selection m.s.e	
		(No extra-polator)	(with extra-polator)	(No extra-polator)	(with extra-polator)	(No extra-polator)	(with extra-polator)
2-D Gauss Markov 8 bit uniform Quantization	4:64	.04873	.0479	.0479	.04702		
	8:64	.0391	.0371	.0402	.03888		
GIRL's picture	4:64	$.123\times10^{-1}$	$.115\times10^{-1}$	$.972\times10^{-2}$	$.885\times10^{-2}$	$.972\times10^{-2}$	$.885\times10^{-2}$
	6:64	$.753\times10^{-2}$	$.645\times10^{-2}$	$.858\times10^{-2}$	$.776\times10^{-2}$	$.764\times10^{-2}$	$.602\times10^{-2}$
	8:64	$.587\times10^{-2}$	$.501\times10^{-2}$	$.729\times10^{-2}$	$.687\times10^{-2}$	$.587\times10^{-2}$	$.501\times10^{-2}$

coefficients to be used in such structures can be made very readily by using the symmetry properties to identify the coefficients with high correlation. An expression similar to (8) for the covariance when the data belong to two adjacent blocks can be easily derived. Such an expression has been used to calculate the predictor coefficient values in hybrid coding scheme [6].

It is seen from Table I that for the one-dimensional case an even (odd) symmetry term has correlation with an even (odd) symmetry term only. Hence, at the transmitter one should select components from both even and odd symmetry terms with large variances. This will avoid the possibility that all the weighting factors in the estimation equation (9) for a particular component are zero.

In the two-dimensional case, the magnitudes of the transform domain covariances calculated from (8) follow the order:

1) covariance of EE,
2) covariance of OE,
3) covariances of OO.

When one selects the transform domain components using MVZS criterion, one runs the risk of selecting a large number of even-even terms (as they have high transform domain covariances) to the exclusion of the odd-odd terms. The extrapolator in such a case may completely fail to estimate the odd-odd terms. On the other hand, if one selects transform domain components on the basis of symmetry criterion only, i.e., selecting equal number of components with large transform domain covariances from each of the four different symmetry classes; then the extrapolator efficiency also suffers, as the terms with large variances are likely to be ignored.

A composite selection scheme, where due consideration to both symmetry and transform domain variance is given, appears to be the best method of selection in transform systems using the extrapolator. In this scheme, one selects at least one component with large variance from each possible symmetry class. The rest of the components are to be selected on the basis of magnitude of the variances.

Table II gives the mean square error for different compression ratios with and without extrapolator in case of Hadamard transform of size (8 × 8). The data considered are as follows:

1) two-dimensional first-order Gauss-Markov sequence (mean zero, $\rho_x = \rho_y = 0.9$) with 8 bit uniform quantization, and

2) GIRL's [1] picture. The mean square error (MSE) is defined as $\text{MSE} = \Sigma (f - \hat{f})^2 / \Sigma f^2$ where f and $\hat{f}$ are the original and estimated data.

DISCUSSION AND CONCLUDING REMARKS

The results given in Table II show that for (4:64) compression, the use of symmetry criterion gives lower MSE than MVZS criterion of component selection, whereas for (8:64) compression, the converse is true.

This is because for 4:64 and 6:64 compression, MVZS criterion rejects OO symmetry components, while symmetry criterion for 8:64 compression selects OO terms with low transform domain covariances at the cost of EE terms. In both the cases, the extrapolator efficiency is seen to suffer. The composite selection scheme based on both MVZS and symmetry criterion ensures better extrapolator efficiency. The improvement of MSE may be slight, but there is a large impact on the subjective quality of the picture [7], [8].

REFERENCES

[1] M. Narasimhan, "Image data processing by Hadamard-Haar transforms," Ph.D. dissertation, Dep. Electron. Elec. Eng., Univ. Texas at Arlington, Arlington, TX, 1975.
[2] W. K. Pratt, "Transform image coding spectrum extrapolation," in *Proc. Seventh Hawaii Int. Conf. Syst. Sci.*, 1974, pp. 7-9.
[3] N. Ahmed and K. R. Rao, *Orthogonal Transforms for Digital Signal Processing*. New York: Springer-Verlag, 1975.
[4] N. Kumar, N. B. Chakraborti, and A. K. Mukherjee, "Use of transform domain spectrum extrapolator for efficient picture transmission," *Electron. Lett.*, vol. 16, pp. 195-196, Feb. 1980.
[5] A. N. Netravali, B. Prasad, and F. N. Mounts, "Some experiments in adaptive and predictive coding of pictures," *Bell Syst. Tech. J.*, vol. 56, pp. 1531-1547, Oct. 1977.
[6] N. B. Chakraborti, "Use of transform domain correlation for efficient picture transmission," in *Proc. Int. Conf. on Comm. Circuits and Syst.*, Jadavpur, India, Dec. 1981.
[7] J. B. O'Neal, Jr., "Differential pulse code modulation (DPCM) with entropy coding," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 169-174, Mar. 1976.
[8] J. B. O'Neal, Jr. and T. Natarajan, "Coding isotropic images," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 696-707, Nov. 1977.

Restoring Lost Samples from an Oversampled Band-Limited Signal

ROBERT J. MARKS, II

Abstract—**When a band-limited signal is sampled at its Nyquist rate, each sample is independent. When sampled in excess of that rate, the samples become dependent. Thus, if one or more samples are lost, they can be recovered from the remaining known values. Using Gerchberg's iterative algorithm applied to interpolation, we derive a closed**

Manuscript received November 25, 1981; revised November 17, 1982.
The author is with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195.

0096-3518/83/0600-0752$01.00

form expression by which a finite number of data points can be regained from the remaining data in a uniformly oversampled signal. A second restoration algorithm—based on an observation due to Howard—is also presented.

INTRODUCTION

The classical Whittaker–Shannon sampling theorem [1] dictates that a finite energy band-limited signal sampled at or above its Nyquist rate is uniquely determined by its samples. When sampled at the Nyquist rate, each signal sample is independent of every other sample. The sample values from an oversampled band-limited signal, on the other hand, are dependent. Thus, lost samples can be regained from knowledge of the remaining samples. In this paper, we present two concise closed form algorithms for restoring a finite number of lost samples from an oversampled band-limited signal.

PRELIMINARIES

An L_2 signal $f(x)$ is said to be band limited with bandwidth $2W$ if

$$f(x) = \int_{-W}^{W} F(u) \exp(j2\pi ux)\, du \tag{1}$$

where

$$F(u) = \int_{-\infty}^{\infty} f(x) \exp(-j2\pi ux)\, dx.$$

Let $2B$ be a sampling rate in excess of the Nyquist rate $2W$. Define the sampling rate parameter:

$$r \equiv \frac{W}{B} \leqslant 1.$$

Then the cardinal series

$$f(x) = \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2B}\right) \operatorname{sinc}(2Bx - m) \tag{2}$$

converges uniformly where $\operatorname{sinc} x = \sin(\pi x)/(\pi x)$. Passing both sides of (2) through a low-pass filter unity on $|u| \leqslant W$ and zero elsewhere gives

$$f(x) = r \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2B}\right) \operatorname{sinc}(2Wx - rn). \tag{3}$$

Equation (2) can be considered a special case for $r = 1$.

LOST SAMPLE RESTORATION

Let $\mathfrak{N}$ denote a finite set of M integers corresponding to various values of m. Given the set of samples $\{f(m/2B) \mid m \notin \mathfrak{N}\}$, the problem at hand is to determine $\{f(m/2B) \mid m \in \mathfrak{N}\}$. Let $\vec{f}$ denote the M-dimensional vector of the lost samples arranged in increasing order of index.

Theorem: If $r < 1$, then

$$\vec{f} = [\boldsymbol{I} - \boldsymbol{S}]^{-1} \vec{h}_0 \tag{4}$$

where $\boldsymbol{S}$ is an $M \times M$ Toeplitz matrix with elements

$$\{r \operatorname{sinc} r(m - n) \mid (m, n) \in \mathfrak{N} \times \mathfrak{N}\}.$$

$\boldsymbol{I}$ denotes the identity matrix and $\vec{h}_0$ is a vector of linear combinations of the known samples. For $m \in \mathfrak{N}$, the elements of $\vec{h}_0$ are $h_0(m/2B)$ where

$$h_0(x) = r \sum_{n \notin \mathfrak{N}} f\left(\frac{n}{2B}\right) \operatorname{sinc}(2Wx - rn). \tag{5}$$

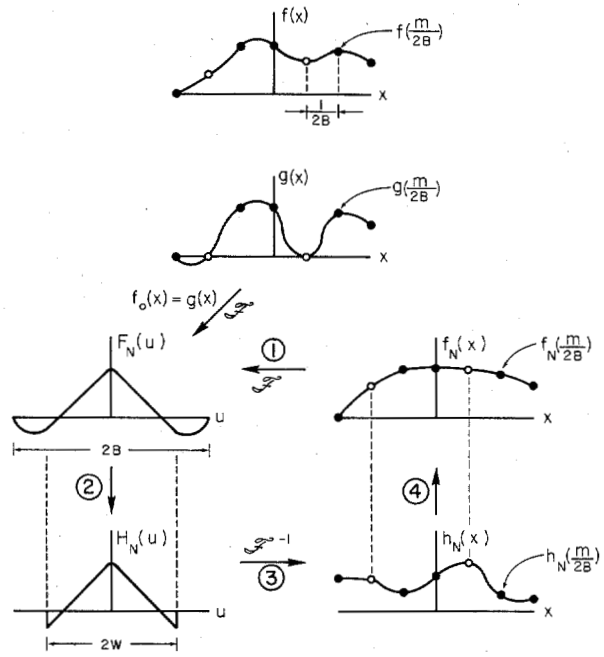

Fig. 1. Illustration of Gerchberg's iterative algorithm applied to the restoration of lost sample values.

Corollary: If $r < 1$, then

$$f(0) = \frac{r}{1 - r} \sum_{m \neq 0} f\left(\frac{m}{2B}\right) \operatorname{sinc}(rm). \tag{6}$$

The signal can then be interpolated without use of the sample at the origin:

$$f(x) = \sum_{m \neq 0} f\left(\frac{m}{2B}\right) \left[\operatorname{sinc}(2Bx - m) + \frac{r}{1 - r} \operatorname{sinc}(rm) \operatorname{sinc}(2Bx)\right] \tag{7a}$$

$$= r \sum_{m \neq 0} f\left(\frac{m}{2B}\right) \left[\operatorname{sinc}(2Wx - rm) + \frac{r}{1 - r} \operatorname{sinc}(rm) \operatorname{sinc}(2Wx)\right] \tag{7b}$$

Proof of (6) follows immediately from (4) for one lost sample at the origin. Equations (7a) and (7b) then follow from (2) and (3), respectively.

Proof via Gerchberg's Algorithm: Gerchberg's iterative algorithm [2] is pictured in Fig. 1. A band-limited function $f(x)$ is oversampled at a rate $2B$. M samples, illustrated by hollow dots, are lost. From the remaining data, we form the function

$$g(x) = \sum_{m \notin \mathfrak{N}} f\left(\frac{m}{2B}\right) \operatorname{sinc}(2Bx - m). \tag{8}$$

Note that this is equivalent to setting the unknown sample values to zero as is shown in the sketch of $g(x)$ in Fig. 1.

The first step in the algorithm is Fourier transformation. Let the result of the Nth iteration be the band-limited signal $f_N(x)$ with initialization $f_0(x) = g(x)$. Since

$$f_N(x) = \sum_{m=-\infty} f_N\left(\frac{m}{2B}\right) \operatorname{sinc}(2Bx - m)$$

the corresponding transform is given by the Fourier series

$$F_N(u) = \frac{1}{2B} \sum_{m=-\infty}^{\infty} f_N\left(\frac{m}{2B}\right) \exp(-j\pi mu/B) \operatorname{rect}\left(\frac{u}{2B}\right) \tag{9}$$

where rect (ξ) is unity for $|\xi| \leq \frac{1}{2}$ and is zero elsewhere. This is shown as step 1 in Fig. 1.

We know from (1) that the spectrum of $f(x)$ is identically zero for $|u| > W$. This is enforced in step 2 where we form the function

$$H_N(u) \equiv F_N(u) \operatorname{rect}\left(\frac{u}{2W}\right). \tag{10}$$

Step 3 is inverse Fourier transformation to the function $h_N(x)$:

$$h_N(x) = \int_{-\infty}^{\infty} H_N(u) \exp(j2\pi ux)\, du. \tag{11}$$

This estimate of $f(x)$ is sampled at a rate of $2B$. The values of $f(m/2B)$, however, are known exactly for all $m \notin \mathfrak{M}$. These replace the samples of $h_N(m/2B)$ for $m \notin \mathfrak{M}$ and we obtain the sample values:

$$f_{N+1}\left(\frac{m}{2B}\right) = \begin{cases} h_N\left(\frac{m}{2B}\right); & m \in \mathfrak{M} \\ f\left(\frac{m}{2B}\right); & m \notin \mathfrak{M} \end{cases} \tag{12}$$

This is step 4. Step 1 is then performed and, in the limit

$$\lim_{N\to\infty} f_N\left(\frac{m}{2B}\right) = f\left(\frac{m}{2B}\right); \quad m \in \mathfrak{M}$$

and our lost samples are regained. Convergence of Gerchberg's algorithm in a similar context has been proven in three distinct ways [1]-[4]. The problem we are addressing is one of interpolation. Such problems are well-posed [4]-[5].

Let's review the description of the previous section with the goal of placing one iteration of the algorithm in a single concise expression. Since only values of $m \in \mathfrak{M}$ are of interest, we have from (12)

$$\begin{aligned} f_{N+1}\left(\frac{m}{2B}\right) &= h_N\left(\frac{m}{2B}\right); \quad m \in \mathfrak{M} \\ &= \frac{1}{2B} \sum_{n=-\infty}^{\infty} f_N\left(\frac{n}{2B}\right) \\ &\quad \cdot \int_{-W}^{W} \exp[j\pi u(m-n)/B]\, du \end{aligned}$$

where, in the second step, we have substituted (9) into (10) into (11) into (12) with $x = m/2B$. Evaluating gives

$$f_{N+1}\left(\frac{m}{2B}\right) = r \sum_{n=-\infty}^{\infty} f_N\left(\frac{n}{2B}\right) \operatorname{sinc} r(m-n); \quad m \in \mathfrak{M}$$

Dividing the sum into $n \in \mathfrak{M}$ and $n \notin \mathfrak{M}$, we obtain

$$\begin{aligned} f_{N+1}\left(\frac{m}{2B}\right) &= h_0\left(\frac{m}{2B}\right) + r \sum_{n\in\mathfrak{M}} f_N\left(\frac{n}{2B}\right) \\ &\quad \cdot \operatorname{sinc} r(m-n); \quad m \in \mathfrak{M}. \end{aligned} \tag{13}$$

The samples of $h_0(x)$ contain the totality of the contribution of the known sample points.

Let $\vec{f}_N$ denote the M-dimensional vector consisting of the Nth estimate of the M lost sample values. Then (13) can be written in matrix form as

$$\vec{f}_{N+1} = \vec{h}_0 + \boldsymbol{S}\vec{f}_N. \tag{14}$$

This is a concise iterative form for restoring lost sample values with a zero vector initialization.

We now consider evaluation of (14) in the limit. Define the unilateral Z transform of the vector sequence $\{\vec{f}_N | N = 0, 1, 2, \cdots\}$ by

$$\vec{f}_N \longleftrightarrow \vec{F}(z) = \sum_{N=0}^{\infty} z^{-N} \vec{f}_N.$$

Transforming (14) and solving for $\vec{F}(z)$ gives

$$\vec{F}(z) = \frac{[\boldsymbol{I}z - \boldsymbol{S}]^{-1}\vec{h}_0}{1 - z^{-1}}.$$

From the final value theorem for Z transforms, it follows that

$$\vec{f} = \lim_{N\to\infty} \vec{f}_N = [\boldsymbol{I} - \boldsymbol{S}]^{-1}\vec{h}_0 \tag{15}$$

and the proof is complete. For $r = 1$, note that the $\boldsymbol{I} - \boldsymbol{S}$ matrix is singular. An alternate proof, based on an observation due to Howard [6], [7] is offered in the Appendix.

An Alternate Restoration Formula

An alternate method to restore lost samples follows.

Theorem: If $r < 1$, then

$$\vec{f} = \boldsymbol{E}^{-1}\vec{G} \tag{16}$$

where $\boldsymbol{E}$ has elements

$$\left\{\frac{-1}{2B} \exp(-j\pi mu_p/B) \,\middle|\, (m, p) \in \mathfrak{M} \times \mathfrak{M}\right\}$$

and the points $\{u_p | p = 1, 2, \cdots, \rho \in \mathfrak{M}\}$ are nonequal but otherwise arbitrary values chosen from the interval $W < u < B$. The vector $\vec{G}$ contains elements $G(u_p)$ where, from (8)

$$G(u) = \frac{1}{2B} \sum_{m\notin\mathfrak{M}} f\left(\frac{m}{2B}\right) \exp(-j\pi mu/B) \operatorname{rect}\left(\frac{u}{2B}\right) \tag{17}$$

Corollary: If $r < 1$,

$$f(0) = -\sum_{m\neq 0} f\left(\frac{m}{2B}\right) \exp[-j\pi m(r+1)/2].$$

This follows from (16) for a single sample with $u_0 = (B + W)/2$. From (2) and (3) it then follows that, for $f(x)$ real,

$$\begin{aligned} f(x) &= \sum_{m\neq 0} f\left(\frac{m}{2B}\right) [\operatorname{sinc}(2Bx - m) \\ &\quad - \cos\{\pi m(r+1)/2\} \operatorname{sinc}(2Bx)] \\ &= r \sum_{m\neq 0} f\left(\frac{m}{2B}\right) [\operatorname{sinc}(2Wx - rm) \\ &\quad - \cos\{\pi m(r+1)/2\} \operatorname{sinc}(2Wx)]. \end{aligned}$$

Note that (7a) and (7b) converge faster due to the $1/m$ factor in the sinc not present in the corresponding cos term above.

Proof: The function $F(u)$ is zero for $|u| > W$ and $G(u)$ is zero for $|u| > B$. Note that over the intervals $W < |u| < B$ we have $F(u) = 0$. Fourier transforming (2) and separating the

sum gives

$$F(u) = \frac{1}{2B}\left[\sum_{m\in\mathcal{M}} + \sum_{m\notin\mathcal{M}}\right] f\left(\frac{m}{2B}\right) \exp(-j\pi mu/B) \cdot \text{rect}\left(\frac{u}{2B}\right). \tag{18}$$

We thus have for $W < |u| < B$

$$G(u) = \frac{-1}{2B}\sum_{m\in\mathcal{M}} f\left(\frac{m}{2B}\right) \exp(-j\pi mu/B); \quad W < |u| < B \tag{19}$$

where we have substituted (17) for the $\Sigma_{m\notin\mathcal{M}}$ term in (18). Viewing $\{\exp(-j\pi mu/B) | m \in \mathcal{M}\}$ as a basis set for $G(u)$ on this interval, it is clear that the coefficients can be uniquely determined. This observation was made in a similar context by Howard [6], [7].

One method of solving for the lost samples is to sample (19) at M points within the interval $W < u < B$ and form the matrix equation $\vec{G} = \boldsymbol{E}\vec{f}$. Equation (16) follows immediately assuming $\boldsymbol{E}$ is not singular.

APPENDIX

Here a second proof of (4) based on Howard's observation is offered. We begin by inverse transforming (19) over the intervals $W < |u| < B$. Define

$$k(x) \equiv \left[\int_{-B}^{-W} + \int_{W}^{B}\right] G(u) \exp(j2\pi ux)\, du \tag{A1}$$

and

$$\psi(x) \equiv \frac{1}{2B}\left[\int_{-B}^{-W} + \int_{W}^{B}\right] e^{j2\pi ux}\, du = (1 - r)\,\text{sinc}[(B - W)x]\cos[\pi(B + W)x].$$

Substituting (17) into (A1) gives

$$k(x) = \sum_{m\notin\mathcal{M}} f\left(\frac{m}{2B}\right) \psi\left(x - \frac{m}{2B}\right). \tag{A2}$$

This expression is in terms of the known samples. An equivalent expression in terms of the lost samples follows from application of (19) to (A1):

$$k(x) = -\sum_{m\in\mathcal{M}} f\left(\frac{m}{2B}\right) \psi\left(x - \frac{m}{2B}\right). \tag{A3}$$

As before, we equate (A2) and (A3) and compute the lost samples.

One method is to sample $k(x)$ at $x = p/2B$; $p \in \mathcal{M}$. One can easily show that

$$\psi\left(\frac{n}{2B}\right) = \delta_n - r\,\text{sinc}\,rn$$

where δ_n denotes the Kronecker delta. Thus, from (A3),

$$k\left(\frac{p}{2B}\right) = -\sum_{m\in\mathcal{M}} f\left(\frac{m}{2B}\right)[\delta_{p-m} - r\,\text{sinc}\,r(p - m)]$$

or, in matrix form

$$\vec{k} = -[\boldsymbol{I} - \boldsymbol{S}]\,\vec{f} \tag{A4}$$

where $\vec{k}$ denotes the M vector of samples computed from (A2):

$$k\left(\frac{p}{2B}\right) = \sum_{m\notin\mathcal{M}} f\left(\frac{m}{2B}\right)[\delta_{m-p} - r\,\text{sinc}\,r(m - p)] = -r\sum_{m\notin\mathcal{M}} f\left(\frac{m}{2B}\right)\text{sinc}\,r(m - p)$$

and we have recognized that, since $p \in \mathcal{M}$ and $m \notin \mathcal{M}$, the Kronecker delta here is always zero. Comparing with (5),

$$k\left(\frac{p}{2B}\right) = -h_0\left(\frac{p}{2B}\right); \quad p \in \mathcal{M}$$

Solving for $\vec{f}$ in (A4) thus gives (4) and the second proof is complete.

ACKNOWLEDGMENT

The author expresses his deep appreciation to D. Radbel for valuable discussions.

REFERENCES

[1] A. Papoulis, *Signal Analysis.* New York: McGraw-Hill, 1977.
[2] R. W. Gerchberg, "Super-resolution through error energy reduction," *Optica Acta*, vol. 21, pp. 709–720, 1974.
[3] P. DeSantis and F. Gori, "On an iterative method for super-resolution," *Optica Acta*, vol. 22, pp. 691–695, 1975.
[4] D. C. Youla, "Generalized image restoration by method of alternating orthogonal projections," *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 694–702, 1978.
[5] H. Stark, D. Cahana, and H. Webb, "Restoration of arbitrary finite-energy optical objects from limited spatial and spectral information," *J. Opt. Soc. Amer.*, vol. 71, p. 635, 1981.
[6] S. J. Howard, "Method for continuing Fourier spectra given by the fast Fourier transform," *J. Opt. Soc. Amer.*, vol. 71, p. 95, 1981; "Erratum," vol. 71, p. 614, 1981.
[7] S. J. Howard, "Continuation of discrete Fourier spectra using minimum-negativity constraint," *J. Opt. Soc. Amer.*, vol. 71, p. 819, 1981.

Comparison of the Characteristics of Linear Least Squares and Orthonormal Expansion in Estimation

J. E. MCFEE AND Y. DAS

Abstract—**It is frequently assumed in signal processing applications that expansion of a sequence by a weighted sum of mutually orthonormal sequences yields weighting coefficients that are identical to the estimates of the parameters which maximize the likelihood function if the linear sum of sequences is chosen as a model. Although this may be a valid approximation when signal-to-noise ratios are large, it is not generally the case and may lead to erroneous results when substantial noise exists. This paper explores the relationship between orthonormal expansion and linear least squares estimation. In doing so, the conditions under which orthonormal expansion coefficients are maximum likelihood estimates are identified. Several interesting properties related to both techniques are also revealed. The results are relevant to a wide range of signal processing applications such as the discrete Fourier**

Manuscript received January 28, 1982; revised November 16, 1982 and January 13, 1983.

The authors are with the Mines and Range Clearance Group, Defence Research Establishment, Suffield, Ralston, Alta., Canada TOJ 2NO.