Optimal Sampling and Reconstruction of MRI Signals Resulting from Sinusoidal Gradients

Avideh Zakhor, *Member, IEEE*, Robert Weisskoff, and Richard Rzedzian

Abstract—**Fundamental operations of magnetic resonance imaging (MRI) can be formulated, for a large number of methods, as sampling the object distribution in the Fourier spatial-frequency domain, followed by processing the digitized data to produce a digital image. In these methods, controllable gradient fields determine the points in the spatial-frequency domain which are sampled at any given time during the acquisition of the free induction delay (FID) signal. Unlike the constant gradient case in which equally spaced samples of the FID signal in time correspond to uniform samples in the Fourier domain, for time-varying gradients, linear sampling in time corresponds to nonlinear sampling in the Fourier domain, and therefore straightforward inverse Fourier transformation is not sufficient for obtaining samples of the object distribution. MRI methods using time-varying gradients, such as sinusoids, are particularly important from a practical point of view, since they require considerably shorter data acquisition times. In this paper, we derive the optimum continuous time filter and its various discrete time implementations for FID signals resulting from sinusoidal gradients. In doing so we find that the estimation error associated with implementation based on linear temporal sampling, or, equivalently, nonlinear spatial frequency sampling, is smaller than that of nonlinear temporal sampling. In addition, we will show that the optimal maximum likelihood estimator for sinusoidal gradients has higher error variance than that of constant gradients. We present experimental results verifying our theoretical predictions.**

I. Introduction

During the last decade, the technology of magnetic resonance imaging (MRI) has emerged and received increased attention. In the medical area, MRI represents a noninvasive technique for obtaining high-resolution cross sectional images of the human body. Current medical applications include cancerous tumor detection [1], brain imaging [2], imaging of diseased tissues [3], and measurement of blood flow [4]. MRI also holds great potential for the study of biology at the cell level [5]. Other potential applications of MRI imaging include food science and technology, the construction industry, and the study of fuels [6].

MRI is a particularly exciting technology to the signal processing community because it presents new challenges for the development of systems which perform sophisticated data collection, processing, and display [7]–[9]. MRI is similar to other signal processing inversion problems in the sense that the design of successful signal processing algorithms for data inversion must be based on a detailed understanding of the underlying signal generated. However, MRI differs from other signal processing inversion problems due to the degree of control one has over the signal being generated [10].

Primary among the factors currently limiting the rate of growth of MRI in the diagnostic imaging marketplace are concerns relating to its limited applicability to the study of organ systems subject to significant involuntary physiological motion, especially of the heart. This problem can be alleviated by decreasing data acquisition time. By reducing the total scan time to a small fraction of the cardiac period, motion artifacts due to all physiological motions including that of the heart, are reduced.

In an effort to reduce data acquisition time, MRI methods using time-varying gradients have been developed [11]–[13]. In these methods, gradient fields are modulated during the observation of a free induction delay (FID) signal to scan the spatial frequency domain and efficiently collect the data which are necessary for image reconstruction. Since data acquisition occurs during a single FID, imaging occurs in several tens of milliseconds, thus overcoming problems such as motion artifacts mentioned above. Under these conditions, time-varying gradient fields determine the points in the spatial-frequency domain of the object which are sampled at any given time during the acquisition of the FID signal. Unlike the constant gradient case in which equally spaced samples of the FID signal in time correspond to uniform samples in the Fourier domain, for time-varying gradients, linear sampling in time corresponds to nonlinear sampling in the Fourier domain, and therefore straightforward inverse Fourier transformation is not sufficient for obtaining samples of the object distribution.

In this paper, we derive optimal sampling and reconstruction schemes for FID signals resulting from sinusoidal gradients [14]. Sinusoidal gradients are particularly important since they require considerably shorter data acquisition times due to their high-speed switching capabilities and ease of implementation. As we will see, our results can be extended to arbitrary gradients without any difficulty. The organization of the remaining part of this paper is as follows: In Section II we briefly review the

Manuscript received September 2, 1989; revised April 30, 1990. This work was supported by Advanced NMR Systems, Woburn, MA, and NSF Grant MIP-9057466. This paper was presented in part at the SPIE 1989 Symposium on Visual Communications and Image Processing.

A. Zakhor is with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720.

R. Weisskoff and R. Rzedzian are with Advanced NMR Systems Inc., Woburn, MA.

IEEE Log Number 9101343.

relationship between the MRI signal and the gradient waveforms. In Section III, we derive the optimal continuous time filter and its discrete time implementation for the constant gradient case. In Section IV, the optimal filter for the sinusoidal gradients and its various digital implementation of it are proposed. Section V includes experimental results and Section VI includes conclusions.

II. Problem Formulation

Consider an object with extent $[0, L_x]$ in the x direction and $[0, L_y]$ in the y direction. If an exciting RF pulse with center at time $t = 0$ is applied to the object, the resulting MRI signal for $t > 0$ is given by [10]

$$s(t) = \int_0^{L_x} \int_0^{L_y} e^{-(t/T_2)} f(x, y) \cdot \exp\left[-j2\pi(K_x(t)x + K_y(t)y)\right] dx\, dy \quad (1)$$

where T_2 is the spin–spin relaxation time, $f(x, y)$ is the density distribution, and $K_x(t)$ and $K_y(t)$ are the integral of x and y gradients, $G_x(t)$ and $G_y(t)$, respectively. That is,

$$K_x(t) = K_x^{(0)} + \gamma \int_0^t G_x(\tau)\, d\tau \quad (2)$$

$$K_y(t) = K_y^{(0)} + \gamma \int_0^t G_y(\tau)\, d\tau \quad (3)$$

with γ being the gyromagnetic ratio. Since we only observe the signal from $t = 0$ to $t = T$, and $T_2 >> T$, we can drop the $e^{-(t/T_2)}$ term from (1) to get [10]

$$s(t) = \int_0^{L_x} \int_0^{L_y} f(x, y) \cdot \exp\left[-j2\pi(K_x(t)x + K_y(t)y)\right] dx\, dy. \quad (4)$$

The scanning pattern in an MRI experiment can be represented as a curve in the so called K space. Specifically, the curve corresponding to values of $K_x(t)$ and $K_y(t)$ in the $K_x - K_y$ plane, parameterized with respect to time, is referred to as the K space representation of the scanning pattern [10]. An example of column-by-column scanning of the K space is shown in Fig. 1. From (4), it is clear that time samples of the signal $s(t)$ correspond to samples of the Fourier transform of the object density function $f(x, y)$. Thus, for a particular time sample t_0, the quantity $s(t_0)$ corresponds to two-dimensional Fourier transform of the object at spatial frequencies $(K_x(t_0), K_y(t_0))$ in the K space. Furthermore, the rate at which the K space is scanned is related to the shape of the functions $K_x(t)$ and $K_y(t)$, or equivalently $G_x(t)$ and $G_y(t)$. For example, in the constant gradient case, the velocity at which the K space is scanned is constant and therefore uniform time samples of $s(t)$ correspond to uniformly spaced samples in the K space. On the other hand, if one of the gradients is held constant, and the other varies sinusoidally, then the scanning velocity is nonlinear, and therefore equal time samples of $s(t)$ correspond to nonuniform samples in the K space.

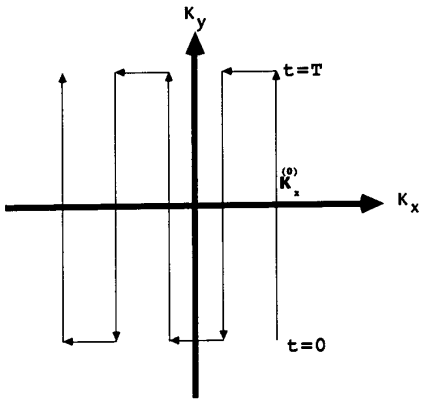


Fig. 1. Column-by-column scanning in the K space.

In this paper, we are primarily concerned with column-by-column scanning as shown in Fig. 1. Specifically, we assume $G_x(t)$ to be zero from $t = 0$ to $t = T$, $G_y(t)$ to be nonzero during the same period, and without loss of generality set $K_y^{(0)} = 0$ in (3). Under these conditions (4) can be rewritten in the following way:

$$s(t) = \int_0^{L_y} B(y) \exp\left[-j2\pi K_y(t)y\right] dy \quad (5)$$

where

$$B(y) = \int_0^{L_x} f(x, y) \exp\left(-j2\pi K_x^{(0)}x\right) dx. \quad (6)$$

Thus samples of $s(t)$ correspond to samples of the Fourier transform of $B(y)$. In practice, the observed MRI signal $r(t)$ is contaminated with noise and is given by

$$r(t) = s(t) + n(t). \quad (7)$$

For the remaining part of this paper, we will assume $n(t)$ to be a white, Gaussian, zero mean random process with intensity σ^2. Our main goal is to find optimal ways of estimating equally spaced samples of $B(y)$ given by

$$B\left(\frac{iL_y}{N}\right) \qquad 0 \le i < N \quad (8)$$

from observing the noisy signal $r(t)$. In the next section, this problem is solved for constant y gradient, and in Section IV the solution for sinusoidal gradient is derived.

III. Constant Gradient Case

In this section, we derive the optimal continuous time filter for the constant gradient case by assuming a piecewise constant model for $B(y)$. As we will see, the resulting filter is identical to the well-known, conventional strategy used for processing signals resulting from constant gradients, i.e., inverse Fourier transforming. We use this model of $B(y)$ in Section IV to derive the optimal filter for the sinusoidal gradient case.

A. Optimal Filter

The density function $B(y)$ is commonly approximated in the following way in the MRI literature [10]:

$$B(y) \approx B\left(\frac{iL_y}{N}\right) \qquad \frac{iL_y}{N} < y < \frac{(i+1)L_y}{N}. \tag{9}$$

Thus, the staircase approximation to the integral of (5) is given by

$$s(t) \approx \frac{L_y}{N} \sum_{i=0}^{N-1} B\left(\frac{iL_y}{N}\right) \exp\left[-j2\pi K_y(t)\frac{iL_y}{N}\right]. \tag{10}$$

For the constant gradient case we have

$$G_y(t) = 1 \qquad 0 \le t < T \tag{11}$$

so that

$$s(t) = \sum_{i=0}^{N-1} B\left(\frac{iL_y}{N}\right) \exp\left(-j2\pi it\frac{\gamma L_y}{N}\right) \tag{12}$$

where for convenience, we have dropped the scale factor L_y/N. Since increasing the length of the gradient pulse results in a smaller field of view (FOV) and therefore smaller L_y, the variables T and L_y are inversely proportional to each other. We can exploit this to simplify the algebra by assuming

$$T = \frac{N}{\gamma L_y}. \tag{13}$$

In effect, we are assuming T to be small enough so that its corresponding FOV includes the object. Incorporating this assumption into (12), and taking into account (7), the observed signal can be written as

$$r(t) = n(t) + \frac{L_y}{N} \sum_{i=0}^{N-1} B\left(\frac{iL_y}{N}\right) e^{-j2\pi i(t/T)}. \tag{14}$$

Thus, our problem can be stated as that of observing $r(t)$ and estimating samples of $B(y)$. This is a problem in classical estimation theory and its solution has been known for a long time [15]. Specifically, the orthnormal basis functions for the Karhunen–Loeve (KL) expansion are given by

$$\phi_m(t) = \frac{1}{\sqrt{T}} e^{j2\pi m(t/T)} \qquad 0 \le m < N \tag{15}$$

and the sufficient statistics are given by

$$r_m = \int_0^T r(t)\phi_m(t)\,dt \qquad 0 \le m < N. \tag{16}$$

Therefore, the maximum likelihood (ML) estimate of samples of $B(y)$ is given by

$$\hat{\boldsymbol{B}} \equiv \begin{bmatrix} \hat{B}(0) \\ \hat{B}(L_y/N) \\ \vdots \\ \hat{B}(N-1)L_y/N) \end{bmatrix} = \frac{1}{\sqrt{T}} \begin{bmatrix} r_0 \\ r_1 \\ \vdots \\ r_{N-1} \end{bmatrix}. \tag{17}$$

From (17) it is clear that the estimate for the ith sample of $B(y)$ is given by

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{T}\int_0^T r(t)e^{j2\pi i(t/T)}\,dt. \tag{18}$$

The estimator of (17) is unbiased and its error covariance is given by

$$E[(\hat{\boldsymbol{B}} - \boldsymbol{B})(\hat{\boldsymbol{B}} - \boldsymbol{B})^*] = \frac{\sigma^2}{T} I \tag{19}$$

where x^* denotes the complex conjugate of the transpose of the vector x.

B. Discrete Time Implementation

In this section, we derive a discrete time implementation of the optimal continuous time filter of Section III-A, and show that its MSE performance characteristics are identical to those of the continuous time filter. The structure of the implementation we propose in this section is heavily influenced by various implementations discussed in the Appendix, which are shown to be either suboptimal, or impractical from an implementation viewpoint.

We begin our discussion by assuming $s(t)$ to be band limited so that passing it through an ideal low-pass filter of the form

$$H(f) = \begin{cases} 1 & |f| \ge W_c \\ 0 & \text{elsewhere} \end{cases} \tag{20}$$

leaves it unchanged. Our proposed estimator is

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{2\Delta PN} \sum_{k=0}^{PN-1} \left(\int_{Tk/PN-\Delta}^{Tk/PN+\Delta} \tilde{r}(t)\,dt\right) e^{j(2\pi ik/PN)} \tag{21}$$

where $\tilde{r}(t)$ is the low-pass version of $r(t)$ using the above filter, PN is the number of samples of small integrals of $\tilde{r}(t)$, and Δ is the integration width for each sample.[1] The estimator of (21) can be shown to be unbiased, and its error covariance is given by

$$\begin{aligned} &\text{MSE}\,(i) \\ &= \left(\frac{1}{2\Delta PN}\right)^2 \sum_{k_1=0}^{PN-1}\sum_{k_2=0}^{PN-1} \exp\left[j\frac{2\pi i(k_1-k_2)}{PN}\right] \\ &\quad \cdot \left(\int_{Tk_1/PN-\Delta}^{Tk_1/PN+\Delta}\int_{Tk_2/PN-\Delta}^{Tk_2/PN+\Delta} E[\tilde{n}(\tau_1)\tilde{n}^*(\tau_2)]\,d\tau_1\,d\tau_2\right) \end{aligned} \tag{22}$$

where $\tilde{n}(t)$ is the low-pass version of the zero mean, white, Gaussian noise $n(t)$. Using linear system theory and theory of random processes, we conclude that $\tilde{n}(t)$ is also zero mean and its power spectrum of $\tilde{n}(t)$ is given by

[1]The reason behind small integrals of $\tilde{r}(t)$ as opposed to its instantaneous samples is discussed at length in the Appendix. Roughly speaking, it has to do with the fact that variance of instantaneous samples of a white process is infinite.

$$P(f) = \begin{cases} \sigma^2 & |f| \le W_c \\ 0 & \text{elsewhere.} \end{cases} \tag{23}$$

Taking the inverse Fourier transform of $P(f)$, we can find the autocorrelation function of $\tilde{n}(t)$:

$$E[\tilde{n}(t)\tilde{n}^*(t-\tau)] = \int_{-W_c}^{W_c} \sigma^2 e^{j2\pi f\tau}\, df = \sigma^2 \frac{\sin(2\pi W_c \tau)}{\pi\tau}. \tag{24}$$

Assuming

$$\frac{1}{2W_c} \le \frac{T}{PN} \tag{25}$$

the double summation in (22) can be replaced by a single summation:

$$\text{MSE}(i) = \left(\frac{1}{2\Delta PN}\right)^2 \sum_{k=0}^{PN-1} \cdot \left(\int_{Tk/PN-\Delta}^{Tk/PN+\Delta} \int_{Tk/PN-\Delta}^{Tk/PN+\Delta} E[\tilde{n}(\tau_1)\tilde{n}^*(\tau_2)]\, d\tau_1\, d\tau_2\right). \tag{26}$$

Since $\tilde{n}(t)$ is stationary, utilizing (24), the double integral of (22) can be changed into a single integral:

$$\text{MSE}(i) = \left(\frac{1}{2\Delta PN}\right)^2 \sum_{k=0}^{PN-1} \cdot \left(\int_{-2\Delta}^{2\Delta} \sigma^2(2\Delta - \tau) \frac{\sin(2\pi W_c\tau)}{\pi\tau}\, d\tau\right). \tag{27}$$

Assuming $1/2W_c >> \Delta$, we can consider $\sin(2\pi W_c\tau)/\pi\tau$ to be almost constant for $-2\Delta < \tau < 2\Delta$. Thus we get

$$\text{MSE}(i) = \sigma^2 \frac{2W_c}{PN}. \tag{28}$$

This seems to imply that decreasing the bandwidth of the filter W_c results in lower error covariance. However, by the assumption of (25), W_c must be larger than or equal to $PN/2T$. This makes the error covariance shown in (28) larger than or equal to σ^2/T, which is the lowest possible estimation error achieved by the continuous time optimal filter of (18). Furthermore, W_c must be larger than the bandwidth of the MRI signal $s(t)$. If the frequency content of $s(t)$ with constant gradient is only nonzero in the range $[-B_c, B_c]$, then the strategy resulting in least error covariance and smallest required sampling rate is to choose $W_c = B_c$, and $PN = 2TW_c$. Under these conditions, the sampling rate of $s(t)$ is at the Nyquist rate, i.e., $2B_c$ and the error covariance of the estimator is

$$\text{MSE}(i) = \frac{\sigma^2}{T}. \tag{29}$$

IV. Sinusoidal Gradient Case

In this section, we will derive the optimal filter for the case where the y gradient is sinusoidal and is of the form:

$$G_y(t) = \frac{\pi}{2} \sin\left(\frac{\pi t}{T}\right) \qquad 0 < t < T \tag{30}$$

and the observed signal is

$$r(t) = n(t) + \sum_{i=0}^{N-1} B\left(\frac{iL_y}{N}\right) \exp\left\{-j\pi i\left[1 - \cos\left(\frac{\pi t}{T}\right)\right]\right\}. \tag{31}$$

Considering the above equation, a reasonable choice of the basis functions would be

$$\phi_k(t) = \frac{1}{T} \exp\left\{j\pi k\left[1 - \cos\left(\frac{\pi t}{T}\right)\right]\right\} \qquad 0 \le k < N. \tag{32}$$

However, unlike the basis functions of the constant gradient case shown in (15), the above set are not orthonormal or othogonal. In fact, their inner product is given by

$$\int_0^T \phi_m(t)\phi_n^*(t)\, dt = \frac{(-1)^{(m-n)}}{T} J_0(\pi(m-n)) \tag{33}$$

where J_0 denotes the zeroth-order Bessel function. Multiplying both sides of (31) by $\phi_k(t)$ and integrating from $t = 0$ to $t = T$, we get

$$\begin{bmatrix} \int_0^T r(t)\phi_0(t)\, dt \\ \int_0^T r(t)\phi_0(t)\, dt \\ \vdots \\ \int_0^T r(t)\phi_{N-1}(t)\, dt \end{bmatrix} = Q \begin{bmatrix} B(0) \\ B\left(\frac{L_y}{N}\right) \\ \vdots \\ B\left(\frac{(N-1)L_y}{N}\right) \end{bmatrix} + \begin{bmatrix} \int_0^T n(t)\phi_0(t)\, dt \\ \int_0^T n(t)\phi_1(t)\, dt \\ \vdots \\ \int_0^T n(t)\phi_{N-1}(t)\, dt \end{bmatrix} \tag{34}$$

where Q is a Toeplitz matrix with its ijth element given by

$$q_{ij} = (-1)^{(i-j)} J_0(\pi(i-j)). \tag{35}$$

Although the random variables $\int_0^T n(t)\phi_m(t)\,dt$ are jointly Gaussian, unlike the constant gradient case, they are not independent of each other. Specifically, we have

$$E[n_k n_m] = \frac{\sigma^2}{T}(-1)^{(k-m)} J_0(\pi(k-m)) \tag{36}$$

where n_k is defined to be

$$n_k \equiv \int_0^T n(t)\phi_k(t)\,dt. \tag{37}$$

Thus, the covariance matrix of the jointly Gaussian random variables n_k is

$$R = \frac{\sigma^2}{T} Q. \tag{38}$$

Having transformed the waveform estimation problem into a discrete one, the maximum likelihood solution can be easily found:

$$\begin{bmatrix} \hat{B}(0) \\ \hat{B}\left(\frac{L_y}{N}\right) \\ \vdots \\ \hat{B}\left(\frac{(N-1)L_y}{N}\right) \end{bmatrix} = Q^{-1} \begin{bmatrix} \int_0^T r(t)\phi_0(t)\,dt \\ \int_0^T r(t)\phi_1(t)\,dt \\ \vdots \\ \int_0^T r(t)\phi_{N-1}(t)\,dt \end{bmatrix}. \tag{39}$$

The above estimator is unbiased and its error covariance matrix is given by

$$\begin{aligned} E[(\hat{\boldsymbol{B}} - \boldsymbol{B})(\hat{\boldsymbol{B}} - \boldsymbol{B})^*] &= (Q^* R^{-1} Q)^{-1} \\ &= \frac{\sigma^2}{T}(Q^*)^{-1} \\ &= \frac{\sigma^2}{T} Q^{-1}. \end{aligned} \tag{40}$$

We have found numerically that the matrix Q^{-1} is diagonally dominant, and that its diagonal elements for $N = 128$ are approximately equal to 1.23. Thus, the mean-squared error in optimal estimation of the ith sample of $B(y)$ is given by

$$\text{MSE}\ (i) \approx 1.23\,\frac{\sigma^2}{T}. \tag{41}$$

Note that the estimation error of the optimal filter for sinusoidal gradients is 1.23 times larger than that of constant gradients. In the remaining parts of this section, we propose various digital implementations of the estimator of (39).

A. Discrete Time Implementation

Recall that the "clean" MRI signal $s(t)$ for the sinusoidal gradient case is given by

$$s(t) = \sum_{i=0}^{N-1} B\left(\frac{iL_y}{N}\right) \exp\left\{-j\pi i\left[1 - \cos\left(\frac{\pi t}{T}\right)\right]\right\}. \tag{42}$$

This implies that uniform time samples of $s(t)$ correspond to nonuniform samples of the discrete time Fourier transform of the sequence $B(iL_y/N)$. On the other hand, PN nonuniform time samples of $s(t)$ at times t_k given by

$$\begin{aligned} K_y(t_k) &= \frac{\gamma T}{2}\left[1 - \cos\left(\frac{\pi t_k}{T}\right)\right] \\ &= \frac{kT}{PN} \end{aligned} \tag{43}$$

correspond to uniform samples of the discrete time Fourier transform of the sequence $B(iL_y/N)$.[2] A pictorial representation of such a sampling is shown in Fig. 2. Therefore, the natural question which arises has to do with relative signal-to-noise ratio (SNR) performance of uniform and nonuniform sampling of the observed signal in the presence of noise $n(t)$. In the next two sections, we evaluate the error characteristics of these two schemes, and compare their performance with the optimal continuous time filter of Section IV.

1) Nonlinear Sampling: Similar to Section III-B, we begin with processing the observed signal $r(t)$ through an ideal low-pass filter of the form

$$H(f) = \begin{cases} 1 & |f| < W \\ 0 & \text{elsewhere.} \end{cases} \tag{44}$$

In doing so, we assume the bandwidth of the filter to be larger than the highest frequency in $s(t)$. If $\tilde{r}(t)$ denotes the output of the above filter due to $r(t)$, then our proposed estimator operates on small time integrals of $\tilde{r}(t)$ around time instants t_k given by (43). Specifically, the estimator is given by

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{2\Delta PN} \sum_{k=0}^{PN-1} \left(\int_{t_k-\Delta}^{t_k+\Delta} \tilde{r}(t)\,dt\right) e^{j(2\pi ik/PN)}. \tag{45}$$

Note that the above estimator is similar to the one shown in (21), except that unlike (21) it uses nonuniform samples at times t_k as opposed to uniform samples at times TK/PN. The estimator of (45) can be shown to be unciased, and its error covariance is given by

$$\begin{aligned} \text{MSE}\ (i) = \left(\frac{1}{2\Delta PN}\right)^2 \sum_{k_1=0}^{PN-1} \sum_{k_2=0}^{PN-1} \exp\left(j\,\frac{2\pi i(k_1 - k_2)}{PN}\right) \\ \cdot \left(\int_{t_{k_1}-\Delta}^{t_{k_1}+\Delta} \int_{t_{k_2}-\Delta}^{t_{k_2}+\Delta} E[\tilde{n}(\tau_1)\tilde{n}^*(\tau_2)]\,d\tau_1\,d\tau_2\right) \end{aligned} \tag{46}$$

[2]Equation (43) shows an example of the so-called implicit sampling in a sense that sampling coordinates are determined by the structure of the signal under consideration [16]. Another example of implicit sampling is level crossing based representation of signals [17], [18]. In contrast with implicit sampling, there is explicit sampling, in which the sampling times are predefined and signal independent. Nyquist sampling is an example of such a scheme.

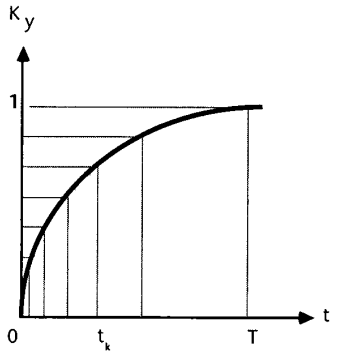


Fig. 2. Pictorial representation of nonlinear sampling described by (43).

where $\tilde{n}(t)$ is the low-passed version of $n(t)$. Assuming

$$\frac{1}{2W} < |t_l - t_{l-1}| \qquad 0 \le l < PN \tag{47}$$

the double summation in (46) can be replaced with a single one. Utilizing the autocorrelation function of $\tilde{n}(t)$ given by (24), and assuming $1/2W >> \Delta$, (46) can be rewritten as

$$\text{MSE } (i) = \sigma^2 \frac{2W}{PN}. \tag{48}$$

Thus, to minimize the error covariance of our estimate, we must choose W as small as possible. However, for the above analysis to hold, it must be chosen larger than the bandwidth of the "clean" MRI signal $s(t)$. If B_s denotes the highest frequency component of $s(t)$ with sinusoidal gradient, to minimize the error covariance, we must choose $W = B_s$. From (48) it is clear that to improve the error covariance, the number of samples must be chosen as large as possible. However, due to assumption of (47), with W fixed to B_s, PN cannot become too large. Specifically, for PN samples, the minimum spacing between two neighboring samples is

$$|t_l - t_{l-1}|_{\min} = \frac{2T}{\pi} \sin^{-1}\left(\frac{1}{PN}\right) \approx \frac{2T}{\pi PN}. \tag{49}$$

Taking into account (47), the largest allowable number of samples is

$$PN_{\max} = \frac{4WT}{\pi} = \frac{4B_sT}{\pi}. \tag{50}$$

Substituting the above equation into (48) we get

$$\text{MSE } (i) = \frac{\pi}{2}\frac{\sigma^2}{T}. \tag{51}$$

Comparing the above equation with (41), we conclude that nonlinear estimation results in mean-squared error which is approximately 28% larger than that of the optimal continuous time filter of (39).

2) Linear Sampling: In this section, we will propose an estimator which uses linear samples of the low-pass version of the observed signal. Again, we are assuming that the low-pass filter given by

$$H(f) = \begin{cases} 1 & |f| < W \\ 0 & \text{elsewhere} \end{cases} \tag{52}$$

has a larger bandwidth than the highest frequency in $s(t)$. From (31) we can find the expression for linear samples of the low-pass version of the observed signal:

$$\int_{Tm/PN-\Delta}^{Tm/PN+\Delta} \tilde{r}(t)\,dt = \int_{Tm/PN-\Delta}^{Tm/PN+\Delta} \tilde{n}(t)\,dt + 2\Delta \sum_{k=0}^{N-1} B\left(\frac{kL_y}{N}\right) \cdot \exp\left\{-j\pi k\left[1 - \cos\left(\frac{\pi m}{PN}\right)\right]\right\}. \tag{53}$$

$\tilde{r}(t)$ and $\tilde{n}(t)$ in the above equation are the low-pass version of $r(t)$ and $n(t)$, respectively. In addition, we have assumed Δ to be small enough so that we can make the following approximation:

$$\int_{Tm/PN-\Delta}^{Tm/PN+\Delta} \exp\left\{-j\pi k\left[1 - \cos\left(\frac{\pi t}{T}\right)\right]\right\} dt = 2\Delta \exp\left\{-j\pi k\left[1 - \cos\left(\frac{\pi m}{PN}\right)\right]\right\}. \tag{54}$$

Writing (53) in a vector format, we get

$$\begin{bmatrix} \int_{-\Delta}^{\Delta} \tilde{r}(t)\,dt \\ \int_{T/PN-\Delta}^{T/PN+\Delta} \tilde{r}(t)\,dt \\ \vdots \\ \int_{T(PN-1)/PN-\Delta}^{T(PN-1)/PN+\Delta} \tilde{r}(t)\,dt \end{bmatrix} = H \begin{bmatrix} B(0) \\ B\left(\frac{L_y}{N}\right) \\ \vdots \\ B\left(\frac{(N-1)L_y}{N}\right) \end{bmatrix} + \begin{bmatrix} \int_{-\Delta}^{\Delta} \tilde{n}(t)\,dt \\ \int_{T/PN-\Delta}^{T/PN+\Delta} \tilde{n}(t)\,dt \\ \vdots \\ \int_{T(PN-1)/PN-\Delta}^{T(PN-1)/PN+\Delta} \tilde{n}(t)\,dt \end{bmatrix} \tag{55}$$

where H is a $PN \times N$ matrix whose mkth element is given by

$$h_{mk} = 2\Delta \exp\left\{-j\pi k\left[1 - \cos\left(\frac{\pi m}{PN}\right)\right]\right\}. \quad (56)$$

Thus, the linear least squares estimator based on linear samples of $\tilde{r}(t)$ is given by

$$\hat{\boldsymbol{B}} = (H^*H)^{-1}H^*\boldsymbol{r} \quad (57)$$

where $\boldsymbol{r}$ denotes the vector on the left-hand side of (55). The terms of the above equation can be related to those of the continuous time optimal filter shown in (39). Specifically, $H^*\boldsymbol{r}$ in (57) corresponds to the discrete time version of the vector on the right-hand side of (39) which denotes the projection of the observed signal $r(t)$ along various basis functions. Furthermore, the square matrix H^*H in (57) corresponds to discrete time implementation of the matrix Q in (39).

We now examine the performance characteristics of the discrete time estimator of (57). Assuming

$$\frac{1}{2W} < \frac{T}{PN} \quad (58)$$

and $\Delta << 1/2W$, we get

$$E\left[\left(\int_{Tm/PN-\Delta}^{Tm/PN+\Delta} \tilde{n}(\tau_1)\,d\tau_1\right)\left(\int_{Ti/PN-\Delta}^{Ti/PN+\Delta} \tilde{n}(\tau_1)\,d\tau_1\right)\right] = \begin{cases} 8W\sigma^2\Delta^2 & m = i \\ 0 & m \neq i. \end{cases} \quad (59)$$

Thus, the autocorrelation of the noise vector in the right-hand side of (55) is

$$R = 8W\sigma^2\Delta^2 I. \quad (60)$$

The estimator of (57) is unbiased, and its error covariance is given by

$$\begin{aligned} E[(\boldsymbol{B} - \hat{\boldsymbol{B}})(\boldsymbol{B} - \hat{\boldsymbol{B}})^*] &= 8W\sigma^2\Delta^2(H^*H)^{-1}H^*[(H^{ast}H)^{-1}H^*]^* \\ &= 8W\sigma^2\Delta^2(H^*H)^{-1}. \end{aligned} \quad (61)$$

The k_1k_2th element of the $N \times N$ matrix $Z = H^*H$ can be found approximately using (33):

$$z_{k_1k_2} = 4\Delta^2 \sum_{m=0}^{PN-1} \exp\left\{-j\pi\left[1 - \cos\left(\frac{\pi m}{PN}\right)\right](k_1 - k_2)\right\} \approx 4\Delta^2 PN(-1)^{(k_1-k_2)}J_0(\pi(k_1 - k_2)). \quad (62)$$

Thus, if the matrix Q is defined by (35), then the error covariance matrix of the estimator of (57) is given by

$$E[(\boldsymbol{B} - \hat{\boldsymbol{B}})(\boldsymbol{B} - \hat{\boldsymbol{B}})^*] = \sigma^2\frac{2W}{PN}Q^{-1}. \quad (63)$$

Thus, to minimize the error covariance of our estimate, we must choose W as small as possible. However, for the above analysis to hold, it must be chosen larger than the bandwidth of the "clean" MRI signal $s(t)$. If B_s denotes the highest frequency component of $s(t)$, to minimize the error covariance, we must choose $W = B_s$. From (63) it is clear that to improve the error covariance, the number of samples must be chosen as large as possible. However, due to assumption of (58), with W fixed to B_s, PN cannot become too large. Specifically, its largest allowed value which does not violate assumption of (58) is

$$(PN)_{\max} = 2WT = 2B_sT. \quad (64)$$

Substituting the above equation into (63) we get

$$E[(\boldsymbol{B} - \hat{\boldsymbol{B}})(\boldsymbol{B} - \hat{\boldsymbol{B}})^*] = \frac{\sigma^2}{T}Q^{-1}. \quad (65)$$

As mentioned earlier, matrix Q^{-1} has been numerically found to be diagonally dominant, and its diagonal elements for $N = 128$ are approximately 1.23. Thus,

$$\text{MSE}\ (i) \approx 1.23\frac{\sigma^2}{T}. \quad (66)$$

Therefore, the error covariance of the linearly sampled estimator of (57) is identical to that of the optimal continuous time filter of (39).

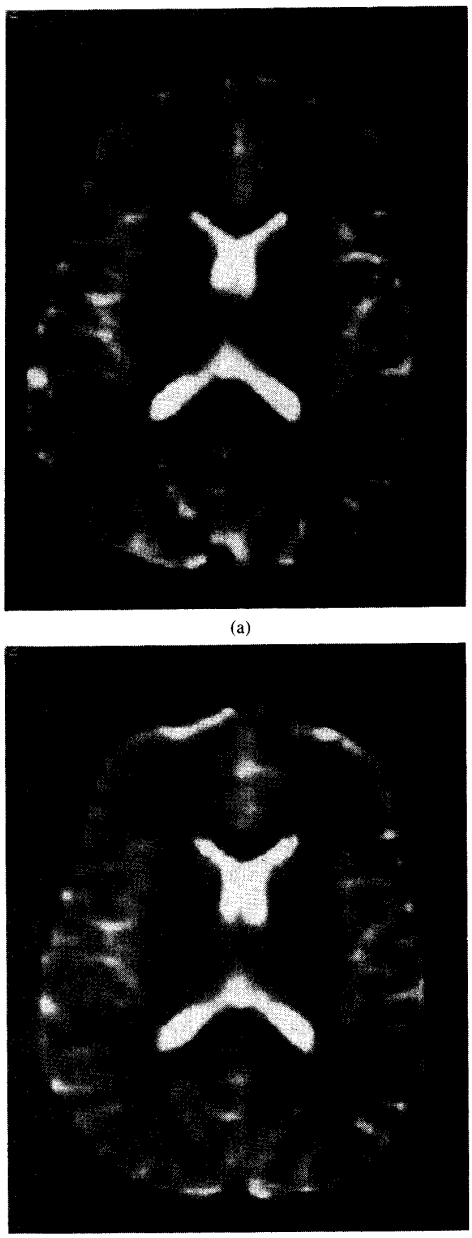
V. Experimental Results

In this section, we present few examples of our proposed estimators for the sinusoidal gradient case. One current application of the sinusoidal processing is for high speed "instant" imaging [19]. By using a sinusoidal frequency encoding gradient, complete magnetic resonance images are regularly produced in as little as 30 ms at rates up to 16 images per second. A comparison of MRI images produced with these sinusoidal gradients is shown in Fig. 3. These images were produced on General Electric Signa 1.5 Tesla scanner retrofitted with Advanced NMR System's Instascan equipment.

Fig. 3(a) shows an image derived from nonlinearly sampled data as described in (45) of Section IV-A1. Fig. 3(b) on the other hand, shows the same head processed from linearly sampled data using the linear least squares estimator of (57) in Section IV-A2. From (50) and (64) it is clear that the number of samples required by linear sampling is approximately $\pi/2$ times that of nonlinear sampling. The two images look similar and artifact free except for the small, slightly different truncation artifact[3] due, most likely, to subtle timing inaccuracies in the sampling.

To quantify the signal-to-noise enhancement (not easily noticeable on the images of Fig. 3) we used a simple phantom, consisting of a 17-cm sphere homogeneously filled with CU_2SO_4 doped water. For this measurement, we acquired 16 images, each a 32 ms acquisition, each image separated by 5 s using the two sampling methods. We determined the relative error covariance indirectly by measuring the average signal-to-noise ratios for the images sets. While the absolute error covariance depends on the exact definition of signal to noise (in particular, on the explicit definition of "noise"), in the large SNR limit,

[3]Gibb's ringing.



(a)

(b)

Fig. 3. Spin echo instant image of human head with (a) nonlinearly (b) linearly sampled data. In both cases, a 128 × 128 image was formed with TE = 73 ms, $TR = \infty$, and 10-mm slide thickness.

the relative covariance is simply the square of the ratio of SNR's for the two methods.

The signal was measured by averaging the magnitude in the same 2 cm^2 uniform region on all the images. The noise, on the other hand, was determined by measuring the standard deviation in a signal free region, over a much larger region of interest.[4] Comparing sets of 16 images using the two methods, we observed an amplitude signal-to-noise increase of 13 ± 1% for the linearly sampled images; the error in this measurement is dominated by the error in the estimate of the noise. Squaring this ratio yields a 27 ± 2% smaller error covariance, in agreement with predictions of (66) and (51).

[4]We measure outside the sphere because the standard deviation within the object itself is dominated by Gibb's ringing.

VI. Conclusions

In this paper, we derived the optimal continuous time filter and its various discrete time implementations for the reconstruction of MRI images from constant and sinusoidal gradients. Our approach was to formulate the problem as a linear parameter estimation one, by writing the observed signal as the sum of the noise-free FID signal and a zero mean, white, Gaussian random process with intensity σ^2/T. The noise-free FID signal was modeled to be a linear combination of the uniform samples of the object distribution. We began by deriving the optimal ML estimate for the constant gradient case. This ML estimate was based on the observed continuous waveform, and our discrete implementation of it was based on its uniform samples. The error variance of the discrete time estimator was shown to be identical to that of the continuous time one, namely proportional to σ^2/T.

We then derived the optimal ML estimator based on the observed continuous waveform, for the sinusoidal gradient case, and showed its error variance to be proportional to $1.23\,\sigma^2/T$. Two discrete implementations of the continuous time ML estimator were proposed. The first one which involved nonuniform sampling of the FID signal resulted in an error variance proportional to $(\pi/2)(\sigma^2/T)$, and the second one which involved uniform sampling of the FID signal resulted in the same error variance as the continuous time ML estimator, provided the sampling rate was fast enough. This second discrete time implementation, therefore, corresponds to the optimal sampling and reconstruction scheme for FID signals resulting from sinusoidal gradients. Experimental results in support of our theoretical predictions were presented.

Two major conclusions can be drawn from our results. First, uniform sampling of the FID signal in time results in the lowest possible error variance for both constant and time-varying gradients. Second, the optimal ML estimator for sinusoidal gradients has higher error variance than that of constant gradients. The derivations for the sinusoidal gradient case can be generalized to arbitrary time-varying gradients. Future work will be directed towards the tradeoff between smoothness of the encoding gradients and their associated error characteristics.

Appendix
Discrete Time Implementations for the Constant Gradient Case

In this Appendix, we derive various discrete time implementations of the continuous time optimal filter shown in Section III-A. As we will see, the schemes proposed in this section are either suboptimal, or result in structures

which are impractical from an implementations point of view. Nevertheless, their performance characteristics are useful in deriving the discrete time implementation of Section III-B which is shown to achieve the MSE performance of the optimal analog filter of Section III-A.

A. Implementation Based on Samples of the Observed Signal

Although the estimator of (17) is optimal, it is a continuous time filter and therefore we must approximate it with a discrete time filter. The most straightforward approximation is an estimator of the form

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{N}\sum_{k=0}^{N-1} r\left(\frac{kT}{N}\right) e^{j(2\pi ik/N)}. \tag{67}$$

The above estimator is unbiased and its error covariance is given by

$$\begin{aligned}\text{MSE}\ (i) &\equiv E\left[\left|B\left(\frac{iL_y}{N}\right) - \hat{B}\left(\frac{iL_y}{N}\right)\right|^2\right] \\ &= \frac{1}{N^2}\sum_{k_1=0}^{N-1}\sum_{k_2=0}^{N-1} E\left[n\left(\frac{k_1T}{N}\right) n^*\left(\frac{k_2T}{N}\right)\right] \\ &\quad \cdot \exp\left(j\frac{2\pi i(k_1-k_2)}{N}\right) \\ &= \frac{\sigma^2}{N}\delta(0).\end{aligned} \tag{68}$$

Therefore, the error covariance of the discrete estimator of (67) is a delta function which is infinitely large. This has to do with the fact that the variance of a sample of a white process is infinite. In the next section we will propose a way of circumventing this problem.

B. Implementation Based on Small Integrals of the Observed Signal

One way to overcome the infinite error covariance of the estimator of the previous section is to use short integrals of the received signal $r(t)$ instead of its samples. Since most analog-to-digital converters (A/D) have finite aperture time, in practice, obtaining small integrals of $r(t)$ is straightforward. Using this technique, we can modify the discrete estimator of (67) in the following manner:

$$\hat{B}\left(\frac{iL_y}{N}\right) = \sum_{k=0}^{N-1}\left(\int_{Tk/N-\Delta}^{Tk/N+\Delta} r(t)\,dt\right) e^{j(2\pi ik/N)} \tag{69}$$

where the integration of the kth sample of $r(t)$ is carried out from $t = Tk/N - \Delta$ to $t = Tk/N + \Delta$. Strictly speaking, the above estimator is biased since we have

$$\begin{aligned}E\left[\hat{B}\left(\frac{iL_y}{N}\right)\right] &= \sum_{k=0}^{N-1}\sum_{l=0}^{N-1} B(l)e^{j(2\pi ik/N)} \\ &\quad \cdot \left(\int_{Tk/N-\Delta}^{Tk/N+\Delta} e^{-j2\pi lt/T}\,dt\right) \\ &= \frac{NT}{i\pi}\sin\left(\frac{2\pi i\Delta}{T}\right).\end{aligned} \tag{70}$$

However, the bias can be easily removed by modifying it in the following way:

$$\begin{aligned}\hat{B}\left(\frac{iL_y}{N}\right) &= \frac{i\pi}{NT\sin\left(\frac{2\pi i\Delta}{T}\right)}\sum_{k=0}^{N-1} \\ &\quad \cdot \left(\int_{Tk/N-\Delta}^{Tk/N+\Delta} r(t)\,dt\right) e^{j(2\pi ik/N)}.\end{aligned} \tag{71}$$

The above estimator is unbiased and its error covariance is given by

$$\begin{aligned}\text{MSE}\ (i) &= \left(\frac{i\pi}{NT\sin\left(\frac{2\pi i\Delta}{T}\right)}\right)^2 \sum_{k_1=0}^{N-1}\sum_{k_2=0}^{N-1} \\ &\quad \cdot \exp\left(j\frac{2\pi i(k_1-k_2)}{N}\right) \\ &\quad \cdot \left(\int_{Tk_1/N-\Delta}^{Tk_1/N+\Delta}\int_{Tk_2/N-\Delta}^{Tk_2/N+\Delta} E[n(\tau_1)n^*(\tau_2)]\,d\tau_1\,d\tau_2\right) \\ &= \frac{\sigma^2}{2\Delta N}\left(\frac{\frac{2\pi i\Delta}{T}}{\sin\left(\frac{2\pi i\Delta}{T}\right)}\right)^2.\end{aligned} \tag{72}$$

Since in most practical situations $\Delta << T$, we can approximate $\sin(2\pi i\Delta/T)$ with $2\pi i\Delta/T$ for all values of i. In doing so, the estimator of (71) can be rewritten as

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{2\Delta N}\sum_{k=0}^{N-1}\left(\int_{Tk/N-\Delta}^{Tk/N+\Delta} r(t)\,dt\right) e^{j(2\pi ik/N)} \tag{73}$$

and its error covariance is given by

$$\text{MSE}\ (i) \approx \frac{\sigma^2}{2\Delta N}. \tag{74}$$

C. Increasing the Number of Samples of the Observed Signal

We can improve the performance of the estimator of the previous section by increasing the number of samples of the observed signal $r(t)$. Specifically, the error covariance of the estimator with PN samples of $r(t)$ given by

$$\hat{B}\left(\frac{iL_y}{N}\right) = \frac{1}{2\Delta PN}\sum_{k=0}^{PN-1}\left(\int_{Tk/PN-\Delta}^{Tk/PN+\Delta} r(t)\,dt\right) e^{j(2\pi ik/PN)} \tag{75}$$

is

$$\text{MSE}\ (i) \approx \frac{\sigma^2}{2\Delta PN}. \tag{76}$$

It is important to note that the error covariance of the estimator of (75) cannot become arbitrarily small by increasing P: When P becomes large enough so that

$$\frac{T}{PN} < 2\Delta \tag{77}$$

the integration interval of the neighboring samples of $r(t)$ overlap. Under these conditions, the noise samples $\int_{Tk/PN-\Delta}^{Tk/PN+\Delta} n(t)\,dt$ are not independent of each other and the error covariance is no longer given by (76). Thus, (76) only holds when

$$\frac{T}{PN} \geq 2\Delta. \tag{78}$$

Combining the above inequality with (76), the error covariance of the estimator of (75) can be lower bounded in the following way:

$$\text{MSE}\ (i) \geq \frac{\sigma^2}{T}. \tag{79}$$

Thus, we have managed to find a discrete time implementation of the optimal continuous time estimator of (18). Specifically, if $r(t)$ is sampled at $T/2\Delta$ points, where the kth sample is defined to be $\int_{2\Delta k-\Delta}^{2\Delta k+\Delta} r(t)\,dt$, then, the estimator of (75) will have the lowest possible error covariance, i.e., that of the continuous time optimal filter of (18).

In most practical situations, sampling at the rate $T/2\Delta$ is extremely demanding. In fact, most analog-to-digital converters cannot possibly sample as fast as their aperture time. Fortunately, as we will see in Section III-B, we can take advantage of the band limitedness of the "clean" MRI signal $s(t)$ in order to reduce the required sampling rate.

Acknowledgment

The authors would like to acknowledge the simulation, interpretation, and implementation skills of T. Green. The reviewers' helpful suggestions are also deeply appreciated.

References

[1] R. Damadian, *NMR in Medicine*. New York: Springer, 1981.
[2] G. Holland, R. Hawkes, and W. Moore, "NMR tomography of the brain," *J. Comput. Assist. Tomog.*, vol. 4, p. 429, 1980.
[3] F. W. Smith, J. R. Mallard, A. Reid, and J. M. Hutchison, "NMR tomographic imaging in liver disease," *Lancet*, vol. 1, p. 963, May 1981.
[4] S. X. Sallescunha, R. E. Halback, J. H. Battocletti, and A. Sances, "The NMR blood flowmeter-applications," *Med. Phys.*, vol. 8, no. 4, p. 452, July-Aug. 1982.
[5] D. I. Hoult, *Magnetic Resonance in Biology*. New York: Wiley, 1980.
[6] P. Mansfield and P. G. Morris, *NMR Imaging in Biomedicine*. New York: Academic, 1982.
[7] Z. H. Cho and H. S. Kim, "Fourier transform NMR tomographic imaging," *Proc. IEEE*, vol. 70, pp. 1152-1173, 1982.
[8] H. Scudder, "Introduction to computer aided tomography," *Proc. IEEE*, vol. 66, no. 6, June 1978.
[9] W. S. Hinshaw and A. H. Lent, "An introduction to NMR imaging: From Bloch equation to the imaging equation," *Proc. IEEE*, vol. 71, no. 3, Mar. 1983.
[10] Donald B. Twieg, "The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods," *Med. Phys.*, vol. 10, no. 5, pp. 610-621, Sept./Oct. 1983.
[11] I. Shenberg and A. Macovski, "Resolution and noise consideration in MRI systems with time-varying gradients," *IEEE Trans. Med. Imaging*, vol. 4, no. 3, pp. 144-152, Sept. 1985.
[12] A. Macovski, "Volumetric NMR imaging with time-varying gradients," *J. Mag. Res. Med.*, vol. 2, pp. 29-39, 1985.
[13] A. Maeda, K. Sano, and T. Yokoyama, "Reconstruction by weighted correlation for MRI with time-varying gradients," *IEEE Trans. Med. Imaging*, vol. 7, no. 1, pp. 26-31, Mar. 1988.
[14] A. Zakhor and R. R. Rzedzian, "Method for reconstructing MRI signals resulting from time-varying gradients," U.S. Patent 379 995, July 14, 1989.
[15] Harry L. Van Trees, *Detection, Estimation, and Modulation Theory*, vol. 1. New York: Wiley, 1968.
[16] A. Zakhor and A. V. Oppenheim, "Sampling and reconstruction schemes for multidimensional signals," *Proc. IEEE*, to be published.
[17] A. Zakhor and A. V. Oppenheim, "Sampling schemes for reconstruction of multidimensional signals from multiple level threshold crossings," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Apr. 1988, pp. 721-724.
[18] A. Zakhor, "Reconstruction of multidimensional signals from multiple level threshold crossings," Ph.D. dissertation, Massachusetts Institute of Technology, 1987.
[19] R. R. Rzedzian and I. L. Pykett, *Amer. J. Roentgenol.*, vol. 149, p. 245, 1987.

Avideh Zakhor (S'87-M'87) received the B.S. degree from the California Institute of Technology, Pasadena, and the S.M. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, all in electrical engineering, in 1983, 1985, and 1987, respectively.

She has been an Assistant Professor in the Department of Electrical Engineering, University of California, Berkeley, since 1988. Her research interests are in signal processing and its various applications. She holds two patents in NMR signal processing and a pending patent on over-sampled A/D converters.

Dr. Zakhor was a General Motors scholar from 1982 to 1983, received the Henry Ford Engineering Award and Caltech Prize in 1983, and the Presidential Young Investigators Award and IBM Junior Faculty Development Award in 1990.

Robert Weisskoff, photograph and biography not available at the time of publication.

Richard Rzedzian, photograph and biography not available at the time of publication.