High-Resolution Alignment of Sampled Waveforms

KEVIN C. McGILL AND LESLIE J. DORFMAN

Abstract—Waveforms are often sampled faster than the Nyquist rate to obtain desired temporal resolution, even though, theoretically, oversampling adds no information and should not be necessary. This paper shows how high resolution can be achieved efficiently from data sampled at the Nyquist rate by working with coefficients of the Fourier-series expansion of the continuous interpolating waveform. Practical algorithms are presented for aligning and comparing waveforms, locating peaks, resolving superimpositions, and averaging overlapping waveforms. The algorithms prove to be more accurate, and to require fewer computations and less storage than techniques which employ continuous oversampling in many signal-processing applications, particularly template matching.

I. INTRODUCTION

IN order to accurately compare two waveforms with random arrival times, they must first be aligned to the point of maximum correlation [1], [2]. In practice, this is usually accomplished in discrete time by sampling the waveforms, aligning the vectors of samples to maximize their cross correlation, and comparing the aligned vectors [2]–[7]. Since this method only allows discrete time shifts, its temporal resolution is only ±0.5 sampling intervals, and a high sampling rate is needed to achieve sufficient resolution for accurate comparison. Sampling at the Nyquist rate is far from adequate, as illustrated in Fig. 1, and the rule of thumb is to oversample by a factor of 5–7 [6], [7]. Alternatively, high resolution can be attained by interpolating the Nyquist-rate samples as prescribed by the sampling theorem [8], but this is considered computationally burdensome.

This paper presents some practical algorithms for efficiently working with sampled waveforms at resolutions finer than the sampling interval. The approach is to treat the waveforms as continuous functions represented by the coefficients of their expansions in a suitable basis rather than as sets of discrete samples. The problem of aligning out-of-phase waveforms then becomes a continuous-time optimization problem which can be solved with theoretically unlimited resolution. For reasons of computational efficiency, the complex exponentials rather than the sinc functions of the sampling theorem are used as the basis functions. The approach leads to efficient algorithms for aligning and comparing waveforms, locating peaks, resolving superimpositions of interfering waveforms, and averaging overlapping waveforms. Since they make full use of the information in the Nyquist-rate samples, the algorithms are more accurate and require lower sampling rates, less storage, and, in fact, fewer computations than discrete-time techniques in many applications.

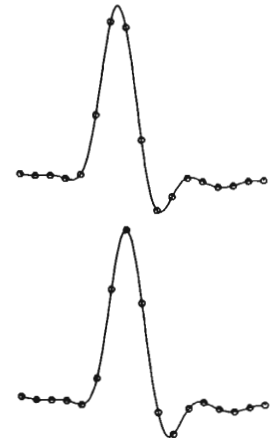

Fig. 1. Two copies of the same waveform, sampled near the Nyquist rate. Sampling at the Nyquist rate only sketches the underlying continuous waveform, and two renditions are apt to appear quite different if the samples are out of phase.

Methods

The waveforms used below to illustrate the algorithms are motor-unit spikes from filtered electromyograms (EMG's). The EMG's were recorded from human biceps muscles during moderate contractions using a concentric needle electrode, sampled at the rate of 10 kHz, and digitally filtered by the simple low-pass differentiator [9]

$$x_n = y_{n+1} - y_{n-1} \tag{1}$$

where y is the raw signal and x is the filtered signal. The purpose of the filter is described in detail elsewhere [10], but basically it transforms motor-unit action potentials into sharp spikes and attenuates low-frequency potentials which are not of interest. (Unlike a true differentiator, it also suppresses high-frequency noise.) The resulting spikes are typically 16–32 sampling intervals in duration, including the initial and final low-amplitude tails, and have high signal-to-noise ratios [(spike energy)$^{1/2}$/rms noise amplitude] ~15–40, depending on the spike—because of the reduced low-frequency interference. The traces in Figs. 2–4 are all 32 sampling intervals long, and the sample points have been interpolated for clarity.

Manuscript received March 28, 1983; revised July 11, 1983.

The authors are with the Department of Neurology, Stanford University School of Medicine, Stanford, CA 94305, and the Rehabilitative Research and Development Center, Palo Alto Veterans Administration Medical Center, Palo Alto, CA 94304.

0018-9294/84/0600-0462$01.00

II. Aligning and Comparing Waveforms

The Discrete Fourier Transform Representation

Any finite-energy waveform $x(t)$ can, for practical purposes, be considered both band-limited to some frequency range $[0, 1/2T]$ and time-limited to some period $[0, NT]$. $x(t)$ can then be (approximately) reconstructed from its N samples $x_n = x(nT), n = 0, 1, \cdots, N-1$, as follows:

$$x(t) \simeq \hat{x}(t) = \frac{X_0}{N} + \frac{1}{N} \sum_{k=1}^{N/2-1} \cdot \left[X_k \exp\left(\frac{j2\pi kt}{NT}\right) + X_k^* \exp\left(\frac{-j2\pi kt}{NT}\right) \right] \quad 0 \leqslant t \leqslant NT \tag{2}$$

where

$$X_k = \sum_{n=0}^{N-1} x_n \exp\left(\frac{-j2\pi nk}{N}\right) \qquad k = 0, 1, \cdots, N-1 \tag{3}$$

is the discrete Fourier transform (DFT) of the sequence $x(nT)$, and $*$ denotes complex conjugation. Notice that (2) makes use of the fact that $X_{N/2+1}, \cdots, X_{N-1}$—which are coefficients of negative-frequency sinusoids despite their positive subscripts—equal the conjugates of the positive-frequency coefficients $X_{N/2-1}, \cdots, X_1$ since x is real. Equation (2) also assumes that $X_{N/2}$ is zero since the sampling theorem allows signal energy only at frequencies strictly less than $1/2T$.

A key point for the subsequent development is that $\hat{x}(t)$ is a continuous function of the continuous variable t, and that the techniques of continuous function theory (e.g., differentiation with respect to t) can be brought to bear by operating on the coefficients X_k. The DFT representation is well suited for aligning waveforms because time shifts which are fractions of the sampling interval can be computed by simple rotations, i.e.,

$$X_{k,\phi} = X_k \exp\left(\frac{j2\pi k\phi}{N}\right) \tag{4}$$

where $X_{k,\phi}$ is the DFT of $x((n+\phi)T)$. Another advantage of the DFT representation is that the sums in (3) can be computed quickly using the FFT algorithm [11]. For display purposes, a smooth interpolation of the samples x_n can be computed by using the FFT algorithm to inverse transform a longer version of X_k which has been appropriately padded with zeros [11].

In theory, a waveform cannot be both band-limited and time-limited. Consequently, although $\hat{x}(t)$ interpolates the sample values $x(nT)$ exactly, it only approximates $x(t)$ elsewhere. Also, since $\hat{x}(t)$ is periodic, it approximates $x(t)$ only in the interval $[0, NT]$, and (4) is really a circular delay. Nevertheless, band-limitedness can essentially be achieved by using an anti-aliasing filter, and both the approximation error and the wrap-around error (for small ϕ) can be kept to the level of the background noise by choosing N large enough to include as much of the waveform as rises significantly above the noise level. (Truncation resulting in discontinuity between x_{N-1} and x_0 will give rise to Gibbs oscillations at the beginning and end of the interpolating waveform.) Hereafter, we will no longer distinguish between the original waveform $x(t)$ and the interpolating waveform $\hat{x}(t)$, referring instead to both as the "underlying" waveform.

The Alignment Algorithm

It is possible to align two waveforms to a resolution finer than the sampling interval using their DFT representations. The optimal least-squares alignment between two waveforms $x(t)$ and $s(t)$ is achieved by the offset ϕ which minimizes the alignment error

$$e^2 = \sum_{n=0}^{N-1} [x((n+\phi)T) - s(nT)]^2. \tag{5}$$

Equivalently, ϕ maximizes the cross correlation between $x(t+\phi T)$ and $s(t)$. The error can be expressed in the frequency domain using Parseval's formula

$$e^2 = \frac{1}{N} |X_0 - S_0|^2 + \frac{2}{N} \sum_{k=1}^{N/2-1} |X_{k,\phi} - S_k|^2 \tag{6}$$

where S_k is the DFT of $s(nT)$.

Equation (6) can be minimized using Newton's method [12], which proceeds iteratively as follows:

$$\phi^{(p+1)} = \phi^{(p)} + \bar{u}^{(p)} \tag{7}$$

$$\bar{u}^{(p)} = \begin{cases} u^{(p)} & \text{if } |u^{(p)}| \leqslant \frac{1}{2} \text{ and } \left.\frac{d^2e^2}{d\phi^2}\right|_{\phi^{(p)}} > 0 \\ -\frac{1}{2} \operatorname{sign}\left(\left.\frac{de^2}{d\phi}\right|_{\phi^{(p)}}\right) & \text{otherwise} \end{cases} \tag{8}$$

$$u^{(p)} = -\left.\frac{de^2}{d\phi}\right|_{\phi^{(p)}} \Big/ \left.\frac{d^2e^2}{d\phi^2}\right|_{\phi^{(p)}} \tag{9}$$

where the superscript indicates the iteration number. By ensuring that the step $\bar{u}$ is always in the direction of the negative gradient and limiting the step size to half a sampling interval, (8) guarantees convergence to the nearest local minimum. The required derivatives are obtained by differentiating (6) to get

$$\frac{de^2}{d\phi} = \frac{4}{N} \sum_{k=1}^{N/2-1} \left(\frac{2\pi k}{N}\right) \operatorname{Im}\{X_{k,\phi} S_k^*\} \tag{10}$$

$$\frac{d^2e^2}{d\phi^2} = \frac{4}{N} \sum_{k=1}^{N/2-1} \left(\frac{2\pi k}{N}\right)^2 \operatorname{Re}\{X_{k,\phi} S_k^*\}. \tag{11}$$

The range of initial offsets from which the algorithm will converge to the global minimum is fairly broad for wavelets with sharp peaks, as illustrated in Fig. 2. The achievable resolution depends on the signal-to-noise ratio, and iterations should continue until the step size becomes of this order. It can be shown [1], [13] that for ideal low-pass noise (i.e., noise with a flat spectral density up to the folding frequency), the achievable resolution is $\sigma/\beta\sqrt{E}$ where σ is the rms noise amplitude, E is the wavelet's energy, and β is the wavelet's normalized rms

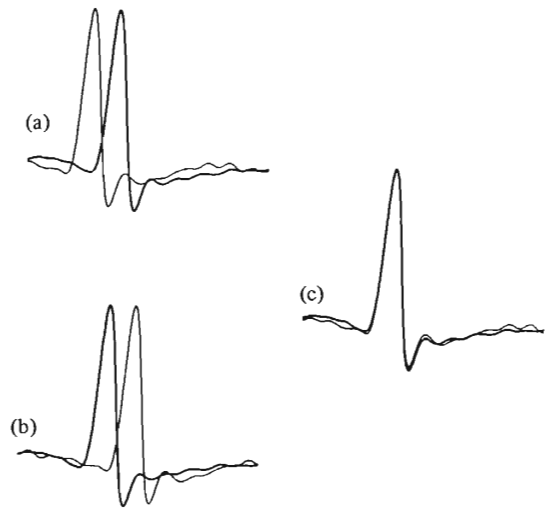

Fig. 2. Waveform alignment. (a) and (b) show the maximum initial offsets between the two wavelets from which the alignment algorithm will converge to the global least squares alignment (c).

bandwidth, which typically lies between 0.4 and 0.9 for sharp-peaked wavelets. A conservative rule is to stop the algorithm after computing $X_{k,\phi(p+1)}$ when $|u^{(p)}| < \sigma/\sqrt{E}$. When the wavelets are initally aligned so that their peak samples coincide, convergence to $|u| < 0.05$ sampling intervals typically takes only one or two iterations.

Resolving Superimpositions

The DFT approach can be extended to determine the best fit of two or more templates to a superimposition of interfering waveforms. The alignment error in this case is

$$e^2 = \frac{1}{N}\sum_{k=0}^{N-1} \left|X_k - S_{1,k,\phi_1} - \cdots - S_{M,k,\phi_M}\right|^2 \tag{12}$$

where

$$S_{i,k,\phi} \triangleq S_{i,k} \exp\left(\frac{j2\pi k\phi}{N}\right) \tag{13}$$

and $S_{i,k}$ and ϕ_i are the DFT and offset of the ith (of M) template. This error can be minimized using the multidimensional form of Newton's method [12]

$$\boldsymbol{\phi}^{(p+1)} = \boldsymbol{\phi}^{(p)} + \boldsymbol{u}^{(p)} \tag{14}$$

$$\boldsymbol{H}^{(p)}\,\boldsymbol{u}^{(p)} = -\boldsymbol{g}^{(p)} \tag{15}$$

where

$\boldsymbol{\phi}$ is the vector of offsets: $(\phi_1\ \phi_2 \cdots \phi_M)^T$

$\boldsymbol{u}$ is the step: $(u_1\ u_2 \cdots u_M)^T$

$\boldsymbol{H}$ is the Hessian matrix: $[\boldsymbol{H}]_{i,j} = \dfrac{\partial^2 e^2}{\partial\phi_i \partial\phi_j}$

$\boldsymbol{g}$ is the gradient: $\left(\dfrac{\partial e^2}{\partial\phi_1}\ \dfrac{\partial e^2}{\partial\phi_2} \cdots \dfrac{\partial e^2}{\partial\phi_M}\right)^T$.

As in the single-dimensional case, two modifications to Newton's method are needed to ensure convergence: indefinite Hessians must be treated with care, e.g., by using the modified Cholesky factorization method [13] to solve (15); and the norm of the step size must be limited to half of a sampling interval. Iterations should proceed until the norm of the step size is sufficiently small. The derivatives are obtained by differentiating (12) to get

$$\frac{\partial e^2}{\partial\phi_i} = \frac{4}{N}\sum_{k=1}^{N/2-1}\left(\frac{2\pi k}{N}\right) \operatorname{Im}\{S_{i,k,\phi_i} D_k^*\} \tag{16}$$

$$\frac{\partial^2 e^2}{\partial\phi_i\partial\phi_j} = \begin{cases} \dfrac{4}{N}\displaystyle\sum_{k=1}^{N/2-1}\left(\frac{2\pi k}{N}\right)^2 \\ \qquad \cdot \left[\operatorname{Re}\{S_{i,k,\phi_i} D_k^*\} + |S_{i,k,\phi_i}|^2\right] & i = j \\ \dfrac{4}{N}\displaystyle\sum_{k=1}^{N/2-1}\left(\frac{2\pi k}{N}\right)^2 \operatorname{Re}\{S_{i,k,\phi_i} S_{j,k,\phi_j}^*\} & i \neq j \end{cases} \tag{17}$$

where

$$D_k \triangleq X_k - S_{1,k,\phi_1} - \cdots - S_{M,k,\phi_M}. \tag{18}$$

Two examples of multidimensional alignment involving wavelets with sharp peaks are shown in Fig. 3. In our experience, initial alignment based on peak samples almost always leads to convergence to the global minimum. Convergence (to $\|u\|_2 < 0.05$) typically takes two to four iterations in cases in which the peaks of the component waveforms are distinguishable [as in Fig. 3(a)], and three to ten iterations in cases in which the peaks merge [as in Fig. 3(b)].

To resolve a superimposition for which it is not known which templates or even how many templates are involved, it is necessary to try all possible template combinations and pick the one that gives the best fit. Each combination starts with zero, one, or more than one template aligned with each peak sample of the superimposition. The search can be quite time consuming if there are many templates because of the large number of combinations (although we have developed techniques to speed the search somewhat [13]). De Figueiredo and Gerber [16] have recently published an algorithm for resolving superimpositions which reduces the set of admissible combinations using a cross-correlation method.

III. Locating Waveform Landmarks

The precise locations of waveform landmarks—zero crossings, peaks, points of fastest rise—are needed to accurately measure times of occurrence, latencies, and peak amplitudes. The waveform's DFT representation provides an efficient way of locating these landmarks with high resolution. Suppose that we want to find a zero of the mth derivative of $x(t)$ ($m \geq 0$). The value of the derivative evaluated at $t = \phi T$ can be expressed in the frequency domain as follows:

$$\frac{d^m x}{d\phi^m} = \frac{1}{N}\sum_{k=1}^{N/2-1}\left[\left(\frac{j2\pi k}{N}\right)^m X_{k,\phi} + \left(\frac{-j2\pi k}{N}\right)^m X_{k,\phi}^*\right] \tag{19}$$

and the zero can be found using Newton's method. The resolution achievable in locating a peak in the presence of ideal low-

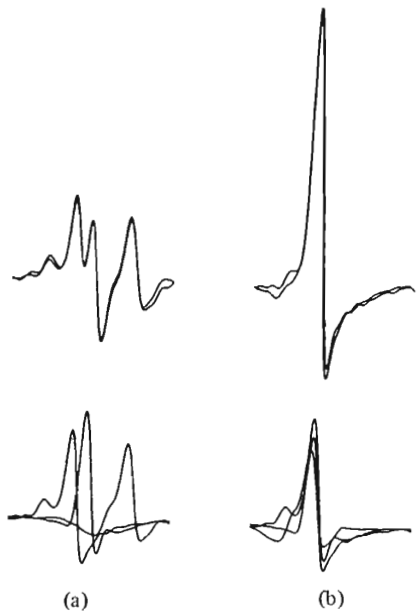

Fig. 3. Resolving superimpositions. (a) The top traces show a segment of EMG arising from the superimposition of three motor-unit spikes and the best fitting summation of the unit's templates as determined by the multidimensional alignment algorithm. (Each template is the average of 50 spikes from the appropriate unit.) The lower traces show the three templates at their optimal alignments. (b) Another superimposition of the same three units and its resolution as determined by the algorithm.

pass noise can be shown [13] to be $\pi\sigma/\sqrt{3}x''(t_p)$ where σ is the rms noise amplitude and t_p is the time at which the peak occurs. A safe rule is to iterate until the step size becomes less than $\sigma/2A$ where A is the amplitude of the peak.

Interestingly, the landmark location problem is equivalent to an alignment problem. Equation (19) can be rewritten

$$\frac{d^m x}{d\phi^m} = -\frac{2}{N}\sum_{k=1}^{N/2-1}\left(\frac{2\pi k}{N}\right)\operatorname{Im}\left\{X_{k,\phi}\left(\frac{j2\pi k}{N}\right)^{m-1}\right\}. \quad (20)$$

The ϕ which sets $d^m x/d\phi^m = 0$ also minimizes the alignment error [sets $de^2/d\phi = 0$ in (10)] when

$$S_k = \left(\frac{-j2\pi k}{N}\right)^{m-1}. \quad (21)$$

This S has a particular significance: it is the DFT of the N-point $(m-1)$st-order differentiator kernel [14]. Thus, for example, locating a wavelet's peak is equivalent to aligning the wavelet with the unit spike

$$s_n = \begin{cases} 1 & \text{if } n = 0 \\ 0 & \text{if } n = 1, 2, \cdots, N-1 \end{cases} \quad (22)$$

$$s(t) \simeq \frac{\sin(\pi t/T)}{\pi t/T} \quad (23)$$

with DFT

$$S_k = 1 \qquad k = 0, 1, \cdots, N-1. \quad (24)$$

We have found it convenient to specify a canonical registration for a wavelet's DFT which is based on a prominent landmark rather than the arbitrary phase of the samples. In particular, we use the peak-finding algorithm to rotate the DFT until the wavelet's peak occurs at time $t = 0$. The location of the peak is a good estimate of a sharp-peaked wavelet's time of occurrence at high signal-to-noise ratios, and in this case, canonical registration aligns wavelets well enough that they can be accurately compared without having to use the alignment algorithm.

IV. Averaging Overlapping Waveforms

Wavelets identified by their resemblance to a template or by a separate trigger are often averaged in order to reduce the effects of noise or to track slow changes in shape. To prevent the wavelets' random sampling phases from distorting the average, the wavelets should be properly aligned by the methods described above. The averaging can then be performed in the frequency domain, component by component, without having to transform back to the time domain.

Many signals contain wavelets which arise from several identified units and interfere with one another. Simple averaging is suboptimal in this case because it treats the interference as unavoidable noise. More accuracy can be obtained by selectively averaging only those occurrences which are free of interference, or by blocking or attenuating those portions of each occurrence in which the interference occurs. Even better averages can be obtained by making a second pass through the data, cancelling the interference with the first-pass wavelet estimates.

The frequency domain provides a way of averaging all the wavelets simultaneously, making optimal use of the available information. Let the identified activity be divided into Q epochs, each N-samples long, such that each epoch contains either one distinct wavelet or several overlapping ones. Let $X_{q,k}$ be the DFT of the qth epoch, and let

$$\delta_{q,m} = \begin{cases} 1 & \text{if unit } m\text{'s wavelet occurs in epoch } q \\ 0 & \text{if it does not} \end{cases}$$

$\phi_{q,m}$ = the offset of unit m is epoch q, if it occurs where $m = 1, \cdots, M$, and M is the number of units.

The problem is then to estimate the DFT's of the wavelets $S_{m,k}\, m = 1, \cdots, M$ so as to minimize the total squared error

$$e^2 = \sum_{q=1}^{Q}\sum_{k=1}^{N-1} e_{q,k}^2 \quad (25)$$

where

$$e_{q,k}^2 = \left|X_{q,k} - \delta_{q,1} S_{1,k,\phi_{q,1}} - \cdots - \delta_{q,M} S_{M,k,\phi_{q,M}}\right|^2 \quad (26)$$

is the error associated with the kth frequency component in epoch q. The notation of (13) is used here to denote time-shifted DFT's. Since the frequency components are orthogonal, (25) can be minimized component by component, i.e., by finding

$$\min_{S_{k,1}, \cdots, S_{k,M}} \sum_{q=1}^{Q} e_{q,k}^2 \qquad \text{for } k = 0, 1, \cdots, N/2 - 1. \quad (27)$$

This is a standard linear least-squares problem [15] which can be written

$$\begin{bmatrix} \delta_{1,1} \exp\left(\frac{j2\pi k\phi_{1,1}}{N}\right) & \cdots \delta_{1,M} \exp\left(\frac{j2\pi k\phi_{1,M}}{N}\right) \\ \vdots & \vdots \\ \delta_{Q,1} \exp\left(\frac{j2\pi k\phi_{Q,1}}{N}\right) & \cdots \delta_{Q,M} \exp\left(\frac{j2\pi k\phi_{Q,M}}{N}\right) \end{bmatrix} \cdot \begin{bmatrix} S_{1,k} \\ \vdots \\ S_{M,k} \end{bmatrix} \doteq \begin{bmatrix} X_{1,k} \\ \vdots \\ X_{Q,k} \end{bmatrix} \tag{28}$$

or in matrix notation

$$\boldsymbol{A}_k \boldsymbol{S}_k \doteq \boldsymbol{X}_k. \tag{29}$$

The symbol $\doteq$ means that $\boldsymbol{S}_k$ is to minimize $(\boldsymbol{X}_k - \boldsymbol{A}_k\boldsymbol{S}_k)^* (\boldsymbol{X}_k - \boldsymbol{A}_k\boldsymbol{S}_k)$ where * denotes the complex conjugate transpose. The solution to (28) can be found by solving the normal equations

$$\boldsymbol{A}_k^* \boldsymbol{A}_k \boldsymbol{S}_k = \boldsymbol{A}_k^* \boldsymbol{X}_k. \tag{30}$$

During data acquisition, the normal equations can be stored explicitly and updated each time a new epoch is observed, as follows:

$$(\boldsymbol{A}_k^* \boldsymbol{A}_k)^{(q)} = (\boldsymbol{A}_k^* \boldsymbol{A}_k)^{(q-1)} + \begin{bmatrix} \delta_{q,1} \exp\left(\frac{-j2\pi k\phi_{q,1}}{N}\right) \\ \vdots \\ \delta_{q,M} \exp\left(\frac{-j2\pi k\phi_{q,M}}{N}\right) \end{bmatrix} \cdot \left[\delta_{q,1} \exp\left(\frac{j2\pi k\phi_{q,1}}{N}\right) \cdots \delta_{q,M} \exp\left(\frac{j2\pi k\phi_{q,M}}{N}\right)\right] \tag{31}$$

$$(\boldsymbol{A}_k^* \boldsymbol{X}_k)^{(q)} = (\boldsymbol{A}_k^* \boldsymbol{X}_k)^{(q-1)} + \begin{bmatrix} \delta_{q,1} \exp\left(\frac{-j2\pi k\phi_{q,1}}{N}\right) \\ \vdots \\ \delta_{q,M} \exp\left(\frac{-j2\pi k\phi_{q,M}}{N}\right) \end{bmatrix} X_{q,k}. \tag{32}$$

After all the epochs have been acquired, the normal equations can be solved by one of several computationally sound methods, e.g., the Cholesky factorization method [15].

Notice that if none of the epochs contains superimpositions, $\boldsymbol{A}_k^* \boldsymbol{A}_k$ is a diagonal matrix whose m, m element counts the epochs in which unit m occurs, and the estimate of the mth unit is the simple average of those epochs after alignment. If epochs containing superimpositions are included, the off-diagonal elements of $\boldsymbol{A}_k^* \boldsymbol{A}_k$ keep track of the interference in $\boldsymbol{A}_k^* \boldsymbol{X}_k$ caused by superimpositions for each pair of units. The major drawback of this method is its large memory requirement: $N(M^2/2 + M)$ words as compared to NM for simple averaging.

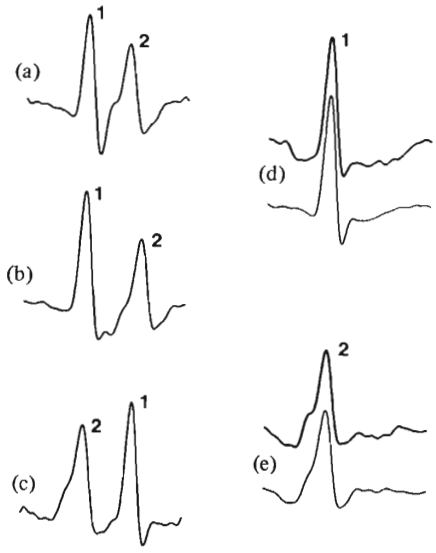

Fig. 4. Averaging overlapping waveforms. (a)–(c) show three EMG segments in which the spikes of two motor units occur side by side and interfere with one another. (d) and (e) show the least-squares averages of the spikes determined from these three segments by the method described in the text (bold) and, for comparison, an interference-free occurrence of each spike (light).

As an example, Fig. 4 shows three segments of EMG in which the spikes of two units interfere with each other, as well as the least-squares averages of the spikes computed from these segments by the above method.

V. Discussion

In order to attain high temporal resolution from sampled data, the data must either be oversampled or the samples must be interpolated. Oversampling is the practical choice in many applications. It bestows a sort of continuity on the samples—neighboring samples do not differ by very much—which makes them suitable for displaying the averaging without interpolation. In addition, it allows many continuous-time signal-processing techniques, such as filtering and correlating, to be borrowed by simply changing integrations to summations, and to be implemented straightforwardly using shift registers and tapped delay lines. (It should be pointed out that the proper anti-alias filter setting when oversampling is not half the sampling rate, but rather half the Nyquist rate of the signal; otherwise, unnecessary high-frequency noise is admitted.) The disadvantages of oversampling include the need for fast A/D converters and long waveform buffers, and the fact that the resolution is limited by the sampling rate. Also, discrete-time optimization techniques are cumbersome, so that aligning waveforms [2]–[7] and resolving superimpositions [4] require time-consuming exhaustive searches of all possible offsets.

Interpolating the sample points allows greater resolution at a lower sampling rate. This is shown in Table I, which compares the average squared alignment errors e^2 when simulated noisy wavelets of random sampling phase were aligned with a noise-free template using the DFT-based algorithm of Section II and using the discrete-time cross-correlation method with oversampling factors (α) of 2, 4, and 8. For the DFT-based algorithm, the average squared error is the sum of the noise variance (which is listed in the right-hand column of the table), a negative bias of about 10 percent due to signal/noise interaction, and a small

TABLE I
AVERAGE ALIGNMENT ERROR e^2 AS A FUNCTION OF SIGNAL-TO-NOISE RATIO (SNR) FOR DISCRETE-TIME ALIGNMENT (WITH OVERSAMPLING FACTORS $\alpha = 2, 4, 8$) AND FOR THE DFT-BASED ALIGNMENT ALGORITHM. THE WAVELET OF FIG. 1 WAS USED AS THE TEMPLATE, WITH $N = 16$, AND 600 NOISY, RANDOMLY ARRIVING WAVELETS WERE SIMULATED AND ALIGNED TO COMPUTE EACH ENTRY. THE LAST COLUMN SHOWS THE TOTAL NOISE VARIANCE FOR N SAMPLES FOR EACH SNR; AND THE LAST ROW SHOWS THE EXPECTED TIME-QUANTIZATION ERROR (e_{tq}^2) FOR EACH BASED ON THE TEMPLATE'S CALCULATED NORMALIZED rms BANDWIDTH OF 0.85. ALL ENTRIES ARE IN UNITS OF (PERCENT ERROR)^2, WHERE 100 PERCENT ERROR CORRESPONDS TO AN rms ERROR EQUAL TO $E^{1/2}$ AND E IS THE TEMPLATE'S ENERGY

	Average of $\left(\frac{100^2\ e^2}{N\ E}\right)$				
SNR	α=2	α=4	α=8	DFT	Noise Variance
10	101.3	93.8	92.1	90.7	100.0
20	31.7	24.9	23.2	22.6	25.0
30	19.4	12.5	10.7	10.3	11.1
40	15.0	8.03	6.28	5.98	6.25
∞	9.84	2.34	0.61	0.30	0.00
e_{tq}^2	9.48	2.37	0.59	0.00	

end-effect error of 0.30 (corresponding to 0.048 percent of the template's energy) due to violation of the time limitedness assumption. For the discrete-time method, the average squared error is the sum of the noise variance, the interaction bias, and an error due to time quantization. Wheeler and Heetderks [5] have derived an expression for the expected time quantization error as follows. Since the alignment offset because of time quantization is uniformly distributed in the interval $[-T/2\alpha, T/2\alpha]$, the variance of a sample point due to this misalignment is $T^2\ |x'(t)|^2/12\alpha^2$. The total expected alignment error is therefore $\beta^2 E/12\alpha^2$ where E is the wavelet's energy and β is its normalized rms bandwidth

$$\beta^2 = \int_{-\infty}^{\infty} [x'(t)]^2\, dt/E \tag{33}$$

$$= 2 \sum_{k=1}^{N/2-1} \left(\frac{2\pi k}{N}\right)^2 |X_k|^2 \Big/ \sum_{k=0}^{N-1} |X_k|^2. \tag{34}$$

The bottom line of Table I shows the expected time quantization error calculated using the above formula.

At the high signal-to-noise ratios shown in Table I, the time quantization error is an appreciable fraction of the noise variance for $\alpha = 2$ or 4, and oversampling by a factor of 8 is necessary to reduce it to less than 5 percent. The DFT-based algorithm, on the other hand, does away with time-quantization errors completely, without oversampling.

The DFT approach is also computationally more efficient than the discrete-time approach at high signal-to-noise ratios which demand large oversampling factors. It requires less storage per wavelet (N versus αN words) and fewer operations per full-wavelet maximum-correlation alignment (approximately $10N$ versus $\alpha^2 N$ real multiplications, where the former figure assumes 1.5 iterations and does not count the FFT, and the latter assumes that α shifts must be searched to find the best alignment). The operation-count savings are even greater when resolving superimpositions.

It should be pointed out that other techniques besides full-wavelet maximum correlation can be used to align wavelets. For example, the correlation can be computed over only the major spike component, or the alignment can be based solely on the location of a landmark, such as the peak (as in canonical registration) or a threshold crossing. These techniques are often preferable to full-wavelet maximum correlation because they require fewer computations, and, at high signal-to-noise ratios, they are nearly as accurate. For these techniques, the frequency domain offers no computational advantages. However, since a time-domain implementation requires oversampling (unless the registering landmark can be detected by analog means and the sampling time-locked to it), a frequency-domain implementation is advantageous if sampling speed is at a premium.

The algorithms in this paper achieve high temporal resolution by approximating the underlying waveforms by continuous functions. Although other bases could be used, for example, the $\sin(\pi(t/T - k))/\pi(t/T - k)$ functions prescribed by the sampling theorem, the complex exponentials have two computation advantages—namely, the speed of the FFT algorithm and their unique time-shift invariance: $\exp(j2\pi k_1 (t + \tau)/NT)$ retains its orthogonality to $\exp(j2\pi k_2 t/NT)$ for all τ for all $k_2 \neq k_1$.

The computational efficiency of the complex exponentials can be seen by considering the use of a different basis. The vector of coefficients of a waveform's expansion in an arbitrary basis $\boldsymbol{X}$ is obtained from the vector of sample values $\boldsymbol{x}$ by a linear transformation $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{x}$. The time-shift operator in the transform domain is also a linear transformation: $\boldsymbol{X}_\phi = \boldsymbol{D}(\phi)\,\boldsymbol{X}$. Maximizing the correlation $\boldsymbol{S}^*\boldsymbol{X}$, where $\boldsymbol{S}$ is the transformed template, requires calculation of the gradient $\boldsymbol{S}^*\boldsymbol{D}'(0)\,\boldsymbol{X}$ and perhaps the second derivative $\boldsymbol{S}^*\boldsymbol{D}''(0)\,\boldsymbol{X}$. For an arbitrary basis, each of these calculations involves a full matrix-vector product which costs N^2 operations (real multiplications), not counting those needed to form $\boldsymbol{D}(\phi)$. However, for certain bases, considerable savings are obtained. For example, if the $\sin(\pi(t/T - k))/\pi(t/T - k)$ functions are used $\boldsymbol{T} = \boldsymbol{I}$ then no operations are needed to form $\boldsymbol{X}$; and $\boldsymbol{D}$, $\boldsymbol{D}'$, and $\boldsymbol{D}''$ are Toeplitz, so multiplications by these matrices can be formulated as convolutions and calculated in less than N^2 operations using fast convolution algorithms. (In fact, if the convolutions are computed using the FFT algorithm, the alignment procedure becomes essentially identical to the algorithm based on the complex exponentials.) When the complex exponentials are used as the basis, $\boldsymbol{X}$ can be calculated using the FFT algorithm which requires only $2N\log_2 N$ operations. Also, because of the time-shift invariance, $\boldsymbol{D}$, $\boldsymbol{D}'$, and $\boldsymbol{D}''$ are diagonal, so that forming $\boldsymbol{D}(\phi)$ requires only $2N$ operations (plus one sine and one cosine evaluation), and forming $\boldsymbol{X}_\phi$, the gradient, and the second derivative requires only $2N$, $1.5N$, and $1.5N$ operations, respectively.

One application for which the algorithms presented in this paper are well suited is the classification of brief wavelets such as nerve and muscle action potentials. A program which uses the algorithms to automatically classify motor-unit action potentials in the clinical EMG is described elsewhere [10]. Briefly, the raw EMG is digitally filtered to produce spikes, and each spike is detected by a threshold crossing, transformed into the frequency domain, canonically registered, compared to each template, and averaged into the template it matches most

closely. The motor-unit action potentials are then averaged from the raw EMG using the identified spikes as triggers. The templates are stored as canonically registered DFT's, and all comparing and averaging are done in the frequency domain.

Note added in proof: While this paper was in press, we became aware of Hansen's work [17] which presents an averaging algorithm very similar to the one presented here.

Acknowledgment

We would like to thank K. L. Cummins for many helpful discussions during the course of this work. We are also grateful to T. Kailath and F. E. Zajac for reviewing drafts of this paper.

References

[1] C. W. Helstrom, *Statistical Theory of Signal Detection.* New York: McGraw-Hill, 1965.
[2] M. Abeles and M. H. Goldstein, "Multispike train analysis," *Proc. IEEE*, vol. 65, pp. 762–773, 1975.
[3] J. Bergmans, "Computer-assisted measurement of the parameters of single motor unit potentials in human electromyography," in *New Developments in Electromyography and Clinical Neurophysiology*, vol. 2, J. E. Desmedt, Ed. Basel, Switzerland: Kerger, 1973, pp. 482–488.
[4] V. J. Prochazka and H. H. Kornhuber, "On-line multi-unit sorting with resolution of superposition potentials," *Electroenceph. Clin. Neurophysiol.*, vol. 34, pp. 91–93, 1973.
[5] B. C. Wheeler and W. J. Heetderks, "The effects of sampling rate on the signal-to-noise ratios of biological signals," in *Proc. 8th New England Bioeng. Conf.*, 1980, pp. 446–450.
[6] ——, "A comparison of techniques for classification of multiple neural signals," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 752–759, 1982.
[7] R. S. LeFever and C. J. De Luca, "A procedure for decomposing the myoelectric signal into its constituent action potentials–Part I: Technique, theory, and implementation," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 144–157, 1982.
[8] A. Papoulis, *Signal Analysis.* New York: McGraw-Hill, 1977.
[9] S. Usui and I. Amidror, "Digital low-pass differentiation for biological signal processing," *IEEE Trans. Biomed. Eng.*, vol. BME-29, pp. 686–693, 1982.
[10] K. C. McGill, K. L. Cummins, and L. J. Dorfman, "Automatic decomposition of the clinical electromyogram," in preparation.
[11] A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing.* Englewood Cliffs, NJ: Prentice-Hall, 1975.
[12] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization.* New York: Academic, 1981.
[13] K. C. McGill, "A method for quantitating the clinical electromyogram," Ph.D. dissertation, Stanford Univ., Stanford, CA, 1984.
[14] R. J. Marks, II, and M. W. Hall, "Differintegral interpolation from a band-limited signal's samples," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 872–877, 1981.
[15] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems.* Englewood Cliffs, NJ: Prentice-Hall, 1974.
[16] R.J.P. De Figueiredo and A. Gerber, "Separation of superimposed signals by a cross-correlation method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1084–1089, 1983.
[17] J. C. Hansen, "Separation of overlapped waveforms having known temporal distributions," *J. Neurosci. Meth.*, vol. 9, pp. 127–139, 1983.

Kevin C. McGill was born in Indianapolis, IN, on June 19, 1952. He received the A.B., B.S.E.E., and M.S.E.E. degrees from the University of Notre Dame, Notre Dame, IN, in 1974, 1975, and 1979, respectively.

From 1976 to 1978 he was with the Naval Avionics Center, Indianapolis. He is currently a Biomedical Engineer with the Rehabilitative Research and Development Center, Palo Alto Veterans Administration Medical Center, Palo Alto, CA, and a Ph.D. candidate in electrical engineering at Stanford University, Stanford, CA. His research interests include biomedical signal processing and the quantitative analysis of electromyograms.

Leslie J. Dorfman was born in Montreal, P.Q., Canada, on September 11, 1943. He received the B.Sc. degree in experimental psychology from McGill University, Montreal, in 1964, and the M.D. degree from Albert Einstein College of Medicine, Bronx, NY, in 1968.

He took residency training in internal medicine at Greenwich Hospital, Greenwich, CT, and in neurology at Stanford University Medical Center, Stanford, CA, and fellowship training in electromyography at the National Hospital for Neurological Diseases, Queen Square, London, England. Since 1974 he has been on the faculty of Stanford University School of Medicine, where he is currently an Associate Professor of Neurology and Director of the Laboratories of Electromyography and Sensory Evoked Potentials. In 1982 he became Program Medical Director for the Rehabilitative Research and Development Center, Palo Alto Veterans Administration Medical Center, Palo Alto, CA. His research interest is the application of signal processing techniques to clinical neurophysiological data for diagnosis of neurological disorders.