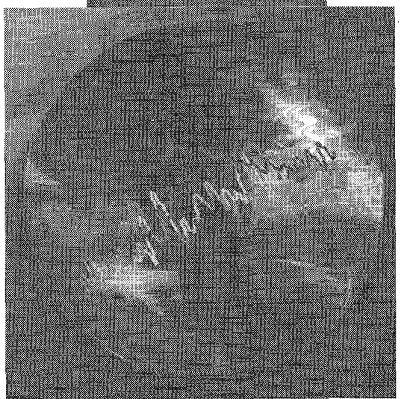


Reducing Correlated Noise in Digital Hearing Aids

A Wavelet-based Method for Extracting Speech from Background Noise

Nathaniel A. Whitmal[1], Janet C. Rutledge[1], and Jonathan Cohen[2]
[1]Dept. of Electrical and Computer Engineering
Northwestern University
[2]Department of Mathematical Sciences
DePaul University

The intelligibility of speech in communication systems is generally reduced by interfering noise. This interference, which can take the form of environmental noise, reverberation, competing speech, or electronic channel noise, reduces intelligibility by masking the signal of interest. The reduction in intelligibility is particularly troublesome for listeners with hearing impairments, who have greater difficulty understanding speech in the presence of noise than do normal-hearing listeners.

Numerous digital signal processing (DSP)-based speech enhancement systems have been proposed to improve intelligibility in the presence of noise. Several of these systems have difficulty distinguishing between noise and consonants, and consequently attenuate both. Other methods, which use imprecise estimates of the noise, create audible artifacts that further mask consonants. The objective of the present study is to develop a new noise-reduction method that can reduce additive noise without impairing intelligibility. The new method could be used to improve intelligibility in a wide variety of applications, with special attention given to digital hearing aids and other portable communication systems (e.g., cellular telephones).

In this article, we present a new wavelet-based method for reducing correlated noise in noisy speech signals. We provide background information on the intelligibility problem and on previous attempts to address it. A theoretical framework is then proposed for reduction of correlated noise, along with some preliminary experimental results.

Background

Intelligibility for Hearing-Impaired Subjects

Studies comparing the speech-perception capabilities of normal-hearing and hearing-impaired listeners in noise have shown that under identical testing conditions, hearing-impaired listeners generally have lower speech-recognition scores, and require higher speech sound pressure levels (SPLs) to achieve the performance of normal-hearing listeners [1]. In one study, Tillman, Carhart, and Olsen [2] reported that listeners with normal hearing could understand 50% of word lists (and 95% of sentences) uttered at a signal-to-noise ratio (SNR) of –5 dB in ambient noise exceeding 60 dB SPL. Hearing-impaired listeners, in comparison, required SNR levels of 9 dB to achieve similar intelligibility.

Many of the differences in perception capability of normal-hearing and hearing-impaired listeners are due to the elevated thresholds of the hearing-impaired listeners. Classic studies of intelligibility [1, 3, 4] liken the effects of these higher thresholds to those of an "internal" masking noise. Other sources of intelligibility loss in hearing-impaired listeners include reduced spectral and temporal distortion, cochlear damage, cognitive processing disorders, and recruitment of loudness (i.e., an acute increase in a subject's perception of loudness with respect to signal level) [5]. The effects of each of these factors (which vary widely from subject to subject) are often worsened by the low fidelity and limited dynamic range of conventional analog hearing aids. In particular, for subjects with recruitment of loudness, volume settings that compensate for elevated thresholds can amplify both the signal and the background noise to uncomfortable listening levels [5].

Previous Noise Reduction Approaches

Many of the noise reduction algorithms considered in studies related to hearing aids

 0739-5175/96/$5.00©1996 September/October 1996

fall into one of two classes: multiple-microphone methods, and single-microphone methods. Multiple-microphone methods exploit the correlation between signals received at spatially separate inputs to identify noise and subtract it from the noisy speech. In applications affording use of spatially separate inputs (e.g., noisy cockpits, automobiles, and some industrial environments), this approach has been shown to improve intelligibility. (Typical examples of improvements are given in [6].) Single-microphone approaches, which rely on statistical models of speech and noise, are more appropriate for compact portable systems (e.g., digital hearing aids, cellular telephones) and other applications for which obtaining spatial information is infeasible.

The results of numerous noise-reduction studies led authors of several literature reviews (e.g., [7, 8]) to conclude that existing single-microphone noise reduction methods were incapable of providing consistent improvements in intelligibility. These results are typified by the recent study of Levitt, et al. [9], who presented results of experiments evaluating the capabilities of four noise-reduction algorithms for use as preprocessors for digital hearing aids. The noise-reduction algorithms included least-mean-square (LMS) filter-based algorithm [6], short-time Wiener filtering, low-frequency spectral subtraction [10], and a sinusoidal-model-based noise-reduction method that reduced noise by discarding the low-amplitude components of a sinusoidal model of the noisy speech [11]. Each of the four methods was evaluated by both normal-hearing and hearing-impaired listeners. The results of the study indicated that:

1. Adaptive noise canceling provided significant intelligibility gains, which decreased in the presence of reverberation and subject head movement.

2. Wiener filtering increased intelligibility for half of the hearing-impaired listeners, and reduced intelligibility for all of the normal-hearing listeners.

3. Spectral subtraction and sinusoidal modeling removed both interfering noise and crucial high-frequency cues, thereby improving SNR (and perceived quality) without improving intelligibility.

The first three noise-reduction approaches evaluated by Levitt, et al. [9], are typical of many approaches in the literature, in that they attempt to preserve features unique to the speech waveform and remove features unique to the noise waveform. An alternative approach to speech enhancement maps segments of the speech waveform into a set of time-varying parameters that may be used to synthesize a modified version of the signal (as with the sinusoidal-model-based approach mentioned above) or input to subsequent processors, such as hearing-loss compensation systems [12].

Wavelet-based parameterization provides a viable alternative to conventional short-time processing methods, since the resulting parameters provide both time-domain and frequency-domain information about the signal. The tree-structured wavelet packet transform of Coifman and Wickerhauser [13] is particularly well-suited for speech processing, as it allows the signal to be decomposed into subbands giving appropriate temporal and spectral resolution for a given application. One potential application of wavelet-based parameterization in digital hearing aids is the TVFD (time-varying, frequency dependent) compensation algorithm developed by Rutledge and Clements [14] for recruitment of loudness compensation. The TVFD algorithm uses a parameterization (derived alternately from a sinusoidal model or from wavelets [15]) to apply time-varying multi-band amplitude compression to speech signals. An important objective of this study is the development of a parametric noise-reduction preprocessing method for wavelet-based TVFD compensation.

Wavelet-based Noise Reduction

The computational efficiency of wavelet transforms and the desirable time-frequency localization properties of their basis functions have motivated intensive research in their use for noise reduction. Many of the investigated approaches (e.g., [16-18]) resemble a form of principal component analysis, where the observed signal is projected onto a specially chosen set of basis functions, and expansion coefficients thought to represent noise are attenuated. The modified coefficients are then inverse-transformed to provide an estimate of the original signal. Differences in the various techniques are mainly distinguished by the methods of choosing the basis functions and the methods of selecting the coefficients to be preserved. The methods adapted for use in this work are described below.

Local Discriminant Bases and Noise Reduction

The method proposed in this article uses the local discrimination bases (LDBs) developed by Saito and Coifman [19] as a feature extractor for pattern recognition and discrimination. Their method generalizes the "best-basis" algorithm of Coifman and Wickerhauser [13], which, for a given signal, produces the minimum-entropy partition of a tree-structured basis system.

In the LDB formulation, the statistical properties of each of C classes of patterns are described by C time-frequency-energy (TFE) maps constructed from sets of training data. For class $\ell \in \{1,2^j,\ldots,C\}$ with N_ℓ training signals $\{\mathbf{x}_\ell(1),\mathbf{x}_\ell(2),\ldots,\mathbf{x}_\ell(N_\ell)\}$, the TFE map $\Gamma_\ell(j,k,m)$ was defined as:

$$\Gamma_\ell(j,k,m) = \frac{\sum_{i=1}^{N_\ell} \left|\phi(j,k,m)^{\mathbf{H}}\mathbf{x}_\ell(i)\right|^2}{\sum_{i=1}^{N_\ell} \left\|\mathbf{x}_\ell(i)\right\|^2} \quad (1)$$

where $\phi(j,k,m)$ is the wavelet-packet basis vector at scale $j \in \{1,2,\ldots,\log_2 N\}$, frequency band $k \in \{0,1,\ldots,2^j-1\}$, and position $m \in \{0,1,\ldots,(2^{-j})N-1\}$in the binary tree, and the superscript "H" denotes Hermitian transposition. For discrimination between two classes, a "best-basis" algorithm was used to select the basis partition maximizing the Kullback-Liebler distances:

$$D(\{p\},\{q\}) = \sum_{i=1}^{N} p_i \log_2 \frac{p_i}{q_i} \quad (2)$$

between the "probability" distributions $\{p_i\}_{i=1}^N$ and $\{q_i\}_{i=1}^N$ corresponding to pairwise combinations of each of the C classes. (The probability distribution $\{p_i\}_{i=1}^N$ is defined as $p_i \equiv \left|\phi_i^H \mathbf{x}\right|^2 / \|\mathbf{x}\|^2$, where ϕ_i was the ith of N basis vectors whose subspaces form a disjoint cover of the binary tree.) The expansion coefficients were then ranked in order of discrimination ability, and a subset of the highest-ranked coefficients was used as input to two conventional pattern-classification algorithms. Saito and Coifman also applied the LDB approach to the denoising problem by selecting bases that maximized discrimination between the classes of "signal+noise" and "noise." For the special case of white noise, the "probability" distribution $\{q_i\}_{i=1}^N$ for the noise class becomes a uniform distribution, and the LDB and best-basis algorithms yield the same basis partition.

Minimum Description Length Criterion

For selection of coefficients, the proposed approach uses the minimum description length (MDL) criterion developed by Rissanen [20] and applied independently by Saito [17] and by Pesquet, et al. [21], to enhance signals in additive white noise. The MDL criterion is an information-theoretic measure that has been used in a number of applications to estimate the order of parametric models.

In the framework of MDL and the context of noise reduction, the noisy observation $\mathbf{x} \in \mathbf{R}^N$ is modeled as an output symbol from a discrete memoryless information source. The description length of $\mathbf{x}$, defined as the length (in bits) of a theoretical binary codeword used to describe $\mathbf{x}$, is given as:

$$L(\mathbf{x},\lambda^{(k)}) = L(\lambda^{(k)}) + L(\mathbf{x}|\lambda^{(k)}) \tag{3}$$

where $L(\lambda^{(k)})$ and $L(\mathbf{x}|\lambda^{(k)})$ are the lengths of codewords, respectively, describing a kth order parametric model $(\lambda^{(k)})$ of $\mathbf{x}$ and the prediction error of that model. $L(\lambda^{(k)})$ and $L(\mathbf{x}|\lambda^{(k)})$ are, respectively, described by a universal prefix coding method proposed by Rissanen and by the Shannon coding method [22]. Among admissible parametric models, the model, which produces the minimum description length, is selected as the model most representative of the signal. A detailed discussion of the MDL criterion may be found in [22].

For the model used in this article, Saito showed that:

$$L(\lambda^{(k)}) = \frac{3k}{2}\log_2 N + C_1 \tag{4}$$

where C_1 was a constant independent of basis and order selection. For $L(\mathbf{x}|\lambda^{(k)})$, which essentially described the noise component of $\mathbf{x}$, the codeword length was given as:

$$L(\mathbf{x}|\lambda^{(k)}) = -\log_2 p(\hat{\mathbf{n}}(\lambda^{(k)});\mathbf{R}_N) \tag{5}$$

where $p(\mathbf{n};\mathbf{R}_N)$ was the probability density of the noise (expressed as a likelihood function of autocorrelation matrix $\mathbf{R}_N$), and $\hat{\mathbf{n}}(\lambda^{(k)})$ was a noise estimate derived from the observation and the kth-order model of the signal. For the case of white Gaussian noise of unknown variance, Saito showed that:

$$L(\mathbf{x}|\lambda^{(k)}) = \frac{N}{2}\log_2\left[\|\mathbf{x}\|^2 - \|\Omega^{(k)}\Phi^{\mathbf{H}}\mathbf{x}\|^2\right] + C_2 \tag{6}$$

where $\Omega^{(k)}$ was a rank-k thresholding matrix preserving the largest k elements of transform coefficient vector $\Phi^{\mathbf{H}}\mathbf{x}$ and C_2 was a constant independent of basis and model order. Detailed derivations of these results, along with examples of MDL-processed signals, may be found in [17].

Preliminary observations suggest that the MDL approach works well for the case of signals corrupted by additive white Gaussian noise. However, the restriction of the present MDL algorithm to reduction of white noise limits its appropriateness in many practical applications.

Reduction of Correlated Noise

In this section, a new approach is proposed for reduction of correlated noise. This approach transforms the basis vectors representing the correlated noise into a second set of basis vectors that is better suited for use with the MDL criterion. The method is in part motivated by the approach of Kay and Nagesha [23] for constructing maximum-likelihood (ML) estimates of the parameters of sinusoidal signals in autoregressive (AR) noise. Their approach used a projection of a pth-order prediction-error filter onto the subspace spanned by the sinusoids to reduce their optimization problem to a more tractable linear least-squares estimation problem.

Notation and Assumptions

In the present case, we consider a deterministic speech signal $s \in \mathbf{R}^N$, comprised of a linear combination of $\kappa(<N)$ vectors chosen from a set of orthonormal basis vectors $\{\phi_i\}_{i=1}^N$ spanning $\mathbf{R}^N$. As above, the index set $i \in \{1,2,\ldots,N\}$ is assumed to represent a set of space-frequency-position triplets $\{(j_i,k_i,m_i)\}_{i=1}^N$ representing a disjoint cover of the binary tree. The projection of the signal onto the κ basis elements is described by parameter vector $\lambda_s = [\lambda_{s,1}\ \lambda_{s,2} \ldots \lambda_{s,N}]^{\mathbf{H}}$, such that:

$$\mathbf{s} = \Phi\lambda_s = \sum_{i=1}^{N}\phi_{z_i}\lambda_{s,z_i} \tag{7}$$

where $\Phi^{\mathbf{H}}$ is the $N \times N$ orthogonal transform matrix $[\phi_1\ \phi_2 \ldots \phi_N]^{\mathbf{H}}$ and $Z_\kappa \equiv \{z_i\}_{i=1}^{\kappa}$ is the set of indices of nonzero expansion coefficients in λ_s. The observed signal $\mathbf{x}$ equals $\mathbf{s}+\mathbf{n}$, where $\mathbf{n}$, the additive noise, is a sample realization of a multivariate Gaussian process $N(\mathbf{0}_{(N\times1)},\mathbf{R}_N)$, which is independent of the signal. The projection of the observation onto the complete set of N basis elements is described by parameter vector $\lambda_x = \Phi^{\mathbf{H}}\mathbf{x}$. The N elements of $\mathbf{n}$ are generated by an AR random process $\{\mathbf{n}_\ell\}$:

$$n_\ell = -\sum_{m=1}^{p}\alpha_m n_{\ell-m} + u_\ell \tag{8}$$

where $\{u_k\}$ is a white-noise sequence with variance σ_u^2. The noise parameters σ_u^2, p (<< N), and $\{\alpha_i\}_{i=1}^{p}$ are assumed to be known in advance. (For speech signals, this is a reasonable and practical assumption, since the noise parameters may be estimated during silence intervals.)

We will now use the MDL criterion and the model presented above to determine $\hat{Z}_\kappa$ and $\hat{\kappa}$, the index set and number of nonzero expansion coefficients of the signal estimate.

MDL Model of Correlated Noise

For the model of Eq. (8), the probability density for the vector $\mathbf{n} = [n_0\ n_1 \ldots n_{N-1}]^{\mathbf{H}}$ is [24]:

$$\begin{aligned} p_N(\mathbf{n};\mathbf{R}_N) &= p_{N|\tilde{N}}(\mathbf{n}|\tilde{\mathbf{n}};\mathbf{R}_{\tilde{N}})p_{\tilde{N}}(\tilde{\mathbf{n}};\mathbf{R}_{\tilde{N}}) \\ &= ((2\pi\sigma_u^2)^N \det\mathbf{R}_{\tilde{N}})^{-\frac{1}{2}}\exp\left[-\frac{Q(\mathbf{n},\mathbf{R}_N)}{2\sigma_u^2}\right] \end{aligned}$$

where:

$$Q(\mathbf{n},\mathbf{R}_{\tilde{N}}) = \tilde{\mathbf{n}}^{\mathbf{H}}\mathbf{R}_{\tilde{N}}^{-1}\tilde{\mathbf{n}} + \sum_{\ell=p}^{N-1}\left(\sum_{m=0}^{N-1}\alpha_m n_{\ell-m}\right)^2$$

with $\alpha_0 = 1, \tilde{\mathbf{n}} = [n_0\ n_1 \cdots n_{p-1}]^{\mathbf{H}}$, and $\mathbf{R}_{\tilde{N}} = E\{\tilde{\mathbf{n}}\tilde{\mathbf{n}}^{\mathbf{H}}\}$. The codeword length for the prediction error of the kth order estimate of the signal is then given by:

$$\begin{aligned} L(\mathbf{x}|\lambda^{(k)}) &= -\log_2 p(\hat{\mathbf{n}}(\lambda^{(k)});\mathbf{R}_N) \\ &= \frac{1}{2\ln 2}\left[N\ln(2\pi\sigma_u^2) + \ln\det\mathbf{R}_{\tilde{N}} + \frac{1}{\sigma_u^2}Q(\hat{\mathbf{n}}(\lambda^{(k)}),\mathbf{R}_{\tilde{N}})\right] \end{aligned} \tag{9}$$

where $\lambda^{(k)}$ is unknown. It may be shown [24] that, for N >> p, the likelihood function $p_N(\mathbf{n};\mathbf{R_N})$ is dominated by the conditional distribution $p_{\tilde{N}}(\tilde{\mathbf{n}};\mathbf{R}_{\tilde{N}})$. Hence, the contributions of terms from $p_{\tilde{N}}(\tilde{\mathbf{n}};\mathbf{R}_{\tilde{N}})$ become negligible, and:

$$L(\mathbf{x}|\lambda^{(k)}) \approx \frac{1}{2\sigma_u^2 \ln 2}\sum_{m=0}^{N-1}\left(\sum_{m=0}^{p}\alpha_m \hat{n}_{l-m}\right)^2 + C_3 \qquad (10)$$

where C_3 is a constant independent of basis and model order. Note that the description length $L(\mathbf{x}|\lambda^{(k)})$ depends directly on the value of the unknown parameter vector $\lambda^{(k)}$. In minimizing description length, we will implicitly assume use of the ML estimate of $\lambda^{(k)}$.

The noise estimate corresponding to the kth-order estimate of the signal is:

$$\begin{aligned}\hat{\mathbf{n}}(\lambda^{(k)}) &= \mathbf{x} - \Phi\lambda^{(k)} = \Phi(\lambda_x - \lambda^{(k)}) \\ &= \sum_{i=1}^{k}\left(\lambda_{x,\hat{z}_i} - \lambda_{\hat{z}_i}^{(k)}\right)\phi_{\hat{z}_i} + \sum_{j=k+1}\lambda_{\hat{z}_j}^{(k)}\phi_{\hat{z}_j}\end{aligned} \qquad (11)$$

Substituting Eq. (11) into Eq. (10) yields:

$$\begin{aligned}L(\mathbf{x}|\lambda^{(k)}) &= \sum_{l=p}^{N-1}\frac{\sum_{m=0}^{p}\left[\alpha_m\left(\sum_{i=1}^{k}\left(\lambda_{x,\hat{z}_i} - \lambda_{\hat{z}_i}^{(k)}\right)\phi_{\hat{z}_i,l-m} + \sum_{j=k+1}^{N}\lambda_{\hat{z}_j}^{(k)}\phi_{\hat{z}_j,l-m}\right)\right]^2}{2\sigma_u^2\ln 2} + C_3 \\ &= \sum_{l=p}^{N-1}\frac{\sum_{i=1}^{k}\left[\left(\lambda_{x,\hat{z}_i} - \lambda_{\hat{z}_i}^{(k)}\right)\sum_{m=0}^{p}\alpha_m\phi_{\hat{z}_i,l-m} + \sum_{j=k+1}^{N}\lambda_{\hat{z}_j}^{(k)}\sum_{m=0}^{p}\alpha_m\phi_{\hat{z}_j,l-m}\right]^2}{2\sigma_u^2\ln 2} + C_3\end{aligned} \qquad (12)$$

where $\phi_{i,j}$ is the jth element of basis vector ϕ_i. Equation (12), which describes the energy in a signal comprised of a linear combination of filtered basis elements, may be conveniently written in matrix-vector format as:

$$L(\mathbf{x}|\lambda^{(k)}) = \frac{1}{2\sigma_u^2\ln 2}\left\|\mathbf{A^H}(\Phi_s(k)\lambda_s(k) + \Phi_\eta(k)\lambda_\eta(k))\right\|^2 + C_3 \qquad (13)$$

where:

$$\mathbf{A^H} = \begin{bmatrix}\alpha_p & \alpha_{p-1} & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \alpha_p & \alpha_{p-1} & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_{p-1} & \alpha_p & \cdots & 1\end{bmatrix}$$

$$\lambda_s(k) = \begin{bmatrix}\lambda_{x,\hat{z}_1} - \lambda_1^{(k)} \\ \lambda_{x,\hat{z}_2} - \lambda_2^{(k)} \\ \lambda_{x,\hat{z}_k} - \lambda_k^{(k)}\end{bmatrix} \lambda_\eta(k) = \begin{bmatrix}\lambda_{x,\hat{z}_{k+1}} \\ \lambda_{x,\hat{z}_{k+2}} \\ \vdots \\ \lambda_{x,\hat{z}_N}\end{bmatrix}$$

$$\Phi_s(k) = \begin{bmatrix}\phi_{\hat{z}_1,0} & \phi_{\hat{z}_2,0} & \cdots & \phi_{\hat{z}_k,0} \\ \phi_{\hat{z}_1,1} & \phi_{\hat{z}_2,1} & \cdots & \phi_{\hat{z}_k,0} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{\hat{z}_1,N-1} & \phi_{\hat{z}2_1,N-1} & \cdots & \phi_{\hat{z}_k,0}\end{bmatrix}$$

$$\Phi_\eta(k) = \begin{bmatrix}\phi_{\hat{z}_{k+1},0} & \phi_{\hat{z}_{k+2},0} & \cdots & \phi_{\hat{z}_n,0} \\ \phi_{\hat{z}_{k+1},1} & \phi_{\hat{z}_{k+2},1} & \cdots & \phi_{\hat{z}_n,1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{\hat{z}_{k+1},N-1} & \phi_{\hat{z}_{k+2},N-1} & \cdots & \phi_{\hat{z}_n,N-1}\end{bmatrix}$$

To find the ML estimate of $\lambda_S^{(k)}$ (and, therefore, that of $\lambda^{(k)}$), we set the gradient of Eq. (13) with respect to $\lambda_S^{(k)}$ equal to zero, obtaining:

$$\lambda_{\mathrm{s,ML}}(k) = \left(\Phi_s(k)^{\mathbf{H}}\mathbf{AA^H}\Phi_s(k)\right)^{-1}\left(\Phi_s(k)^{\mathbf{H}}\mathbf{AA^H}\Phi_\eta(k)\right)\lambda_\eta(k) \qquad (14)$$

The columns in the matrices $\Phi_s(k)^{\mathbf{H}}\mathbf{A}$ and $\Phi_\eta(k)^{\mathbf{H}}\mathbf{A}$ are seen to be the projections of N-p shifted versions of a time-reversed prediction-error filter onto the designated "signal" and "noise" subspaces. Calculation of these $(N$-$p)$ x N transform coefficients may be done quickly and efficiently by means of the shift-invariant wavelet-packet transform. Discussions of the shift-invariant wavelet and wavelet-packet transforms and their use in denoising are provided in [21, 25].

Substituting Eq. (14) into Eq. (13) reveals that $L(\mathbf{x}|\lambda^{(k)})$ is simply a measure of the energy in the subspace orthogonal to the column space of $\mathbf{A^H}\Phi_s(k)$. Hence, minimizing $L(\mathbf{x}|\lambda^{(k)})$ is equivalent to maximizing the expression:

$$\Lambda(\mathbf{x}|\lambda^{(k)}) \equiv \left\|\mathbf{A^H}\Phi_s(k)\left(\Phi_s(k)^{\mathbf{H}}\mathbf{AA^H}\Phi_s(k)\right)^{-1}\Phi_s(k)^{\mathbf{H}}\mathbf{AA^H}\mathbf{x}\right\|^2 \qquad (15)$$

Setting $\lambda_{\mathbf{x_k}} = \Phi_s(k)^{\mathbf{H}}\mathbf{x}$, we see that the above equation is simply the squared norm of the projection:

$$\begin{aligned}\Lambda(\mathbf{x}|\lambda^{(k)}) &= \left\|\mathbf{A^H}\Phi_s(k)\lambda_{\mathbf{x_k}}\right\|^2 \\ &= \lambda_{\mathbf{x_k}}{}^{\mathbf{H}}\Phi_s(k)^{\mathbf{H}}\mathbf{AA^H}\Phi_s(k)\lambda_{\mathbf{x_k}}\end{aligned} \qquad (16)$$

which, as shown below, is nearly equal to $\sum_{i=1}^{k}\lambda_{x,\hat{z}_i}^2\left\|\mathbf{A^H}\phi_{\hat{z}_i}\right\|^2$. We will now show that this bound is tight enough to be useful in developing the MDL algorithm.

The filtering matrix $\mathbf{A^H}$, which consists of N-p circular shifts of the prediction-error filter, equals $\mathbf{A^H} = \mathbf{SFQ_AF^H}$, where $[\mathbf{F}]_{mn} = e^{j\frac{2\pi}{N}mn}/\sqrt{N}$, $\mathbf{S} = \left[\mathbf{I}_{(\mathrm{N-p})\times(\mathrm{N-p})} \quad \mathbf{0}_{(N-p)\times p}\right]$, and $\mathbf{Q_A}$ is a diagonal matrix containing the values of the FFT of the first row of $\mathbf{A^H}$ divided by $\sqrt{N}$. We may then write:

$$\Phi^{\mathbf{H}}\mathbf{AA^H}\Phi = (\mathbf{F^H}\Phi)^{\mathbf{H}}\mathbf{Q_A^HQ_A}(\mathbf{F^H}\Phi) - (\mathbf{F^H}\Phi)^{\mathbf{H}}\Theta(\mathbf{F^H}\Phi) \qquad (17)$$

where:

$$\theta_{mn} = [\Theta]_{mn} = q_m^* q_n \frac{\sin\frac{p\pi}{N}(m-n)}{N\sin\frac{\pi}{N}(m-n)}e^{j\frac{\pi}{N}(p+1)(m-n)} \qquad (18)$$

with q_n being the nth diagonal element of $\mathbf{Q_A}$. It is straightforward to show, then, that:

$$|\theta_{mn}|^2 \le \frac{\mathrm{p}^2}{\mathrm{N}^4}\left(1 + \sum_{\ell=1}^{p}\alpha_\ell^2\right)^2 \qquad (19)$$

Sinha and Tewfik [26] have shown that for wavelets with large numbers of vanishing moments, the first matrix in Eq. (17) is essentially diagonal. Invoking the bound of Eq. (19), and noting that the second matrix is a quadratic function of an orthonormal basis, we may then use the Cauchy-Schwarz inequality to show that each term in the second matrix of Eq. (17) is at least a factor of p/N less than the terms on the diagonal of the first matrix. Hence, the columns of $\mathbf{A^H}\Phi_1(k)$ are nearly orthogonal, such that:

$$\Lambda(\mathbf{x}|\lambda^{(\mathbf{k})}) \approx \sum_{i=1}^{k}\lambda_{x,\hat{z}_i}{}^2\left\|\mathbf{A^H}\phi_{\hat{z}_i}\right\|^2 \qquad (20)$$

and

$$L(\mathbf{x}|\lambda^{(k)}) \approx \frac{1}{2\sigma_u^2 \ln 2}\left[\left\|\mathbf{A}^{\mathbf{H}}\mathbf{x}\right\|^2 - \sum_{i=1}^{k}\lambda_{x,\hat{z}_i}^2\left\|\mathbf{A}^{\mathbf{H}}\phi_{\hat{z}_i}\right\|^2\right] + C_3 \quad (21)$$

Equation (21) is analogous to the MDL equation derived by Saito for the case of white noise, with the transformed basis system $\mathbf{A^H}\Phi$ used in place of the conventional basis system. Hence, a fast algorithm similar to the one employed by Saito may be used to select those basis vectors that best represent the signal.

Calculation of TFE Maps

As noted above, application of the LDB method to the denoising problem requires knowledge of the statistics of the "signal+noise" and "noise" classes (denoted as class 1 and 2, respectively). Acquiring training data for both classes is straightforward, since Class 1 is typified by the input signals to the denoising algorithm, and the AR model of the noise may be used to provide a TFE map for Class 2 data. Referring to Eq. (1), we see that as N_2, the cardinality of the set of training signals for the noise class, approaches infinity:

$$\Gamma_2(j,k,m) \rightarrow \frac{E\left\{\left|\phi(j,k,m)^{\mathrm{T}}\mathbf{n}\right|^2\right\}}{R_n(0)} \quad (22)$$

The numerator of Eq. (22) may be calculated by generating the impulse response of the all-pole filter and then computing its stationary wavelet-packet transform for the binary-tree partition of maximal decimation. At each level $j \in \{1,2,...,\log_2 N\}$ of the transform, $E\{|\phi(j,k,m)^{\mathbf{T}}\mathbf{n}|^2\}$ may be calculated at each value of k and m by summing the energy in the coefficients of the corresponding subband. The denominator is then calculated from the impulse response.

Reduction of Residual Noise

In the previous two sections, we described a wavelet-based denoising method that extends the approaches of [16, 17, 21] to the case of autoregressive Gaussian noise. In all of these approaches, truncation of the N-element series expansion describing the noisy signal can impose audible artifacts on the enhanced speech. (The effect is analogous to the "Gibbs effect" observed in Fourier series expansions.) One approach for reducing the audible artifacts is the "cycle-spinning" method discussed in [25]. There, the stationary wavelet transform was used to compute the average of N time-aligned denoised versions of waveforms created from the N admissible circular shifts of the noisy N-sample sequence.

We have adapted the stationary wavelet-packet transform for use in this work by using the MDL criterion to find a universal threshold for coefficient selection. As in [27], we accomplish this by first finding the finite difference of with respect to k:

$$\Delta L(\mathbf{x},\lambda^{(k)}) \equiv L(\mathbf{x},\lambda^{(k)}) - L(\mathbf{x},\lambda^{(k-1)}) = \frac{3}{2}\log_2 N - \frac{\lambda_{x,\hat{z}_k}{}^2\left\|\mathbf{A}^{\mathbf{H}}\phi_{\hat{z}_k}\right\|^2}{2\sigma_u^2 \ln 2} \quad (23)$$

The minimum value of $L(\mathbf{x},\lambda^{(k)})$ is seen to occur for the value k for which $\Delta L(\mathbf{x},\lambda^{(k)})$ goes from negative to positive; i.e., the largest value of k for which:

$$\left|\lambda_{x,\hat{z}_k}\right| \geq \frac{\sigma_u}{\left\|\mathbf{A}^{\mathbf{H}}\phi_{\hat{z}_k}\right\|}\sqrt{3\ln N} \quad (24)$$

Note that each shift-invariant transform coefficient at a given scale j and frequency k corresponds to a shift-variant transform coefficient in more than one circularly shifted version of the signal. Therefore, averaging N denoised, circularly shifted versions of the signal requires that each coefficient be subjected to multiple threshold rules. A more practical and efficient method of combining thresholded transform coefficients over all N circulant shifts is to use an average threshold that selects the coefficient at (j, k, m) by the rule:

$$\left|\lambda_x(j,k,m)\right| \geq \frac{\sigma_u\sqrt{3\ln N}}{\frac{1}{2^{-j}N}\sum_{m'=0}^{2^{-j}N-1}\left\|\mathbf{A}^{\mathbf{H}}\phi(j,k,m)\right\|} \quad (25)$$

Hence, evaluation of $L(\mathbf{x},\lambda^{(k)})$ for varying values of k may be replaced by a simple thresholding operation, similar to those of [18] and [21]. In contrast to those approaches, however, the method presented here provides scale-dependent, frequency-dependent thresholding that is more appropriately tailored to the characteristics of the noise.

Summary of Algorithm

We may now summarize the sequence of operations in our algorithm, which shares many of the elements of [17, 19, 21]:

1. Construct an average "signal+noise" TFE map from energy averages of data for circularly shifted versions of the noisy signal.
2. For the "signal+noise" and "noise" maps, find the LDB and take a stationary wavelet-packet transform, using the partition returned by the LDB algorithm.
3. Use either the MDL or thresholding approaches presented in this article to denoise the circularly shifted versions of the noisy signal.
4. Use the inverse stationary wavelet-packet transform to resynthesize the signal.

Experimental Results

A preliminary comparison of the capabilities of conventional and proposed noise-reduction approaches was conducted. A recording of the sentence, "That hose can wash her feet" (followed by 880 ms of silence) was sampled at 8 kHz and combined with each of 5 correlated Gaussian noise sequences to produce waveforms with overall SNRs of 0, 5, 10, 15, and 20 dB. The correlated noise sequences were produced by a fourth-order all-pole filter of the form:

$$n_k = 1.3733n_{k-1} - 0.7858n_{k-2} + 0.1138n_{k-3} + 0.0119n_{k-4} + u_k$$

where $\{u_k\}$ was a sequence of white Gaussian noise. (The filter parameters were derived from digital recordings of ambient noise in an automobile traveling at highway speeds.) The speech signals were processed in 50% overlapped 256-sample segments by each of 6 algorithms that used the Daubechies-20 wavelet:

1. Conventional MDL denoising (i.e., using the white noise assumption) averaged over five circular shifts.
2. LDB-based MDL denoising averaged over five circular shifts, using the true AR model.
3. LDB-based MDL denoising averaged over five circular shifts, using an estimate of the AR model.
4. Conventional shift-invariant thresholding.
5. LDB-based shift-invariant thresholding, using the true AR model.
6. LDB-based shift-invariant thresholding, using an estimate of the AR model.

For algorithms 3 and 6, Levinson's recursion [24] was used

Table 1. Segmental SNR improvements for conventional and proposed MDL and shift-invariant algorithms (speech region only)					
Sentence SNR	0.00	5.00	10.00	15.00	20.00
Mean segmental SNR	-10.27	-5.27	-0.27	4.73	9.73
Output of MDL algorithms	SNR Improvements (dB)				
1. Conventional	6.03	4.79	3.68	2.60	1.04
2. Proposed (true AR model)	11.56	8.72	5.83	2.84	-0.44
3. Proposed (estimated AR model)	11.59	8.74	5.79	2.31	-1.94
Output of shift-invariant algorithms	SNR Improvements (dB)				
4. Conventional	6.75	5.49	4.18	2.88	1.76
5. Proposed (true AR model)	11.75	9.21	7.12	5.33	3.61
6. Proposed (estimated AR model)	11.68	9.48	7.29	5.45	3.59

Table 2. Background noise reduction capabilities of conventional and proposed MDL and shift-invariant algorithms (silence region only)					
Sentence SNR	0.00	5.00	10.00	15.00	20.00
Output of MDL algorithms	Reduction in noise level (dB)				
1. Conventional	9.52	9.85	10.50	12.39	18.34
2. Proposed (true AR model)	24.82	24.82	24.82	24.82	24.82
3. Proposed (estimated AR model)	22.94	26.19	26.90	28.49	28.86
Output of shift-invariant algorithms	Reduction in noise level (dB)				
4. Conventional	10.98	10.98	11.03	11.28	12.20
5. Proposed (true AR model)	29.21	29.23	29.26	29.30	29.42
6. Proposed (estimated AR model)	23.77	23.93	24.52	25.50	30.07

to fit a fourth-order AR model to a 256-sample segment containing only additive noise. (Coefficient estimates for the resulting models were –1.34, 0.74, –0.14, and –0.04.) TFE maps for the "signal+noise" class were derived by averaging coefficient energies for each frame over five circular shifts; TFE maps for the "noise" class were derived from AR models as described above.

Segmental SNR measurements were taken over nonoverlapping 256-sample segments and averaged over both speech and silence regions to produce an objective measure of quality. Data for both speech and silence regions are respectively shown in Tables 1 and 2. Both sets of data indicate that the proposed approaches are better suited for enhancing signals in correlated noise than the conventional approaches. The minor differences between outputs for actual and estimated AR models suggest that the proposed algorithms can be used effectively in practical applications. The data in Tables 1 and 2 also indicate that the proposed algorithms have a more modest advantage in the speech regions than in the silence regions. This is caused, in part, by the implicit prediction-error filtering of the waveforms, which can partially decorrelate the speech and make some speech components subject to attenuation.

The quality of the output signals varied as a function of SNR. At the highest SNR level of 20 dB, the output of the proposed algorithms was difficult to distinguish from the original signal. In contrast, at 0 dB SNR, both the conventional and proposed algorithms produced signals containing substantial residual noise. The quality of the output speech tended to sound less natural and more "synthetic," largely because the number of retained coefficients had decreased. Overall, the shift-invariant algorithms enjoyed a slight advantage in measured segmental SNR, which decreased as the level of the noise increased. Spectrograms of the original sentence, the noisy sentence at 5 dB SNR, and sentences processed by the conventional and proposed shift-invariant thresholding algorithms are shown in Figs. 1-4. Inspection of Figs. 3 and 4 shows that the conventional algorithms tended to remove high-frequency features that contributed to intelligibility, while the proposed algorithms tended to retain them. Figures 3 and 4 also indicate that the proposed algorithms attenuate low-frequency noise components that are retained by the conventional algorithms and become particularly noticeable in regions between adjacent words. Recent results [28] indicate that the proposed algorithms' advantage in nonspeech regions may be useful in enhancing the perceived quality of the output speech, as is borne out by informal listening.

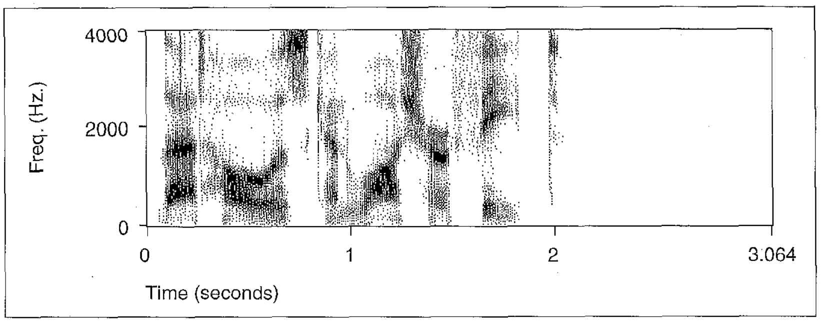
Discussion

In interpreting the results of these experiments, we note that both the conventional and the newly proposed algorithms project the noisy signal onto a set of basis vectors, and then estimate the original signal from a subset of the basis vectors with the largest energy content. The conventional algorithm uses a standard wavelet-packet basis (where all basis vectors contain the same amount of energy), and selects basis vectors solely on the basis of coefficient size. Since both the original speech and the simulated car noise have most of their energies at low frequencies, the conventional algorithm is most likely to preserve basis vectors with substantial energies at low frequencies. In contrast, the filtering imposed on the the new algorithm's basis vectors increases the energies in vectors whose spectrum is most severely compromised by the noise, thereby offseting the low SNRs in these subspaces. This feature allows the new algorithm to retain low-level high-frequency components, which may improve intelligibility.

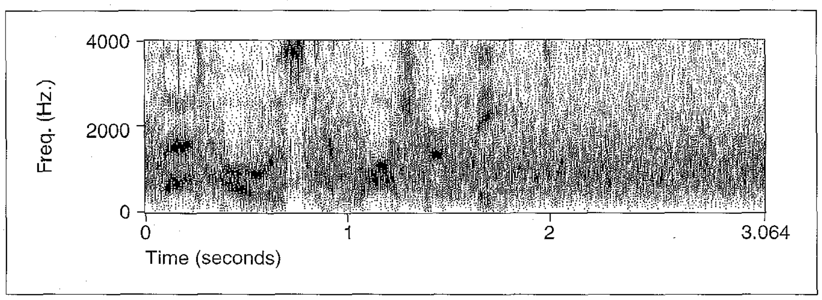
Several comments are appropriate at this point. First, we note that the combination of coefficient attenuation and averaging can produce benign distortion artifacts, which result in objective measures with overly pessimistic values. (This is particularly true at the 20 dB SNR level.) Also, the combined use of the LDB approach and shift-invariant transforms can result in transform structures with impractical computation requirements. Our present work is focused on development of an intelligibility-based LDB algorithm that addresses this concern. Finally, it should be noted that, at least on a conceptual level, one could choose either to filter the basis vectors (as is done in this article), or simply to filter the signal and to apply the conventional algorithm. In the application of interest, where parameters from a denoising algorithm may be used as input to other algorithms, the first approach is preferred. Furthermore, the approach taken in this paper may be useful in extending theoretical results for wavelet-based reduction in white noise to other applications where an AR noise model is appropriate.

Conclusion

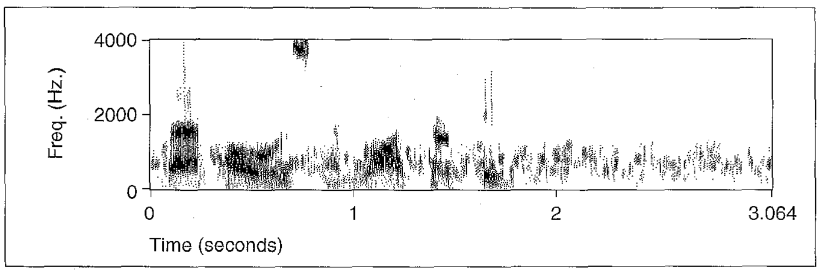
We have investigated a new wavelet-based method for extracting speech from background noise. In particular, we used the local discriminant bases of Saito and Coifman [19] to obtain a basis partition that maximizes discrimination between two classes: speech in noise, and noise alone. In previous work [29], we used the best-basis criterion of Coifman and Wickerhauser [13] and the MDL criterion [17] to extract signals from a background of white Gaussian noise. The LDB-based approaches presented in this article work better than conventional approaches when applied to speech in correlated noise. In particular, improvements in both speech-to-noise ratio and retention of consonants were observed in our experiments. While these results are encouraging, much work remains. We are currently developing a version of the MDL algorithm that uses a prior distribution appropriate to transform coefficients of speech signals. We are also investigating methods for incorporating versions of the Articulation Index [3, 30] into the coefficient selection process, and for using discarded coefficients to update estimates of masking noise parameters. The resulting algorithm will be evaluated as a front-end preprocessor for hearing-loss compensation in digital hearing aids. While the technique is appropriate for a



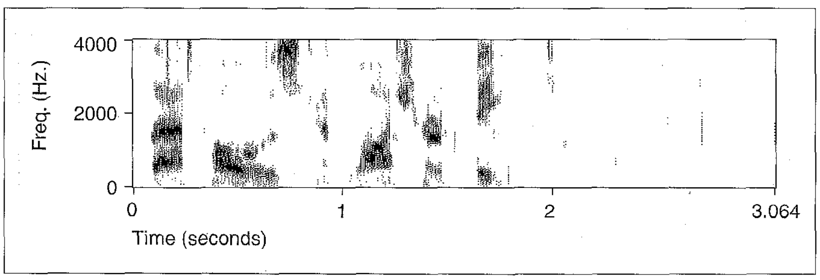
1. "That hose can wash her feet."



2. "That hose can wash her feet" in simulated car noise at 5 dB SNR.



3. Sentence of Figure 2, processed by Algorithm 4 (conventional shift-invariant denoising).



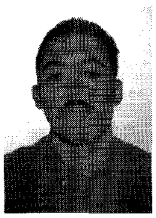
4. Sentence of Figure 1, processed by Algorithm 5 (proposed shift-invariant denoising using true AR model).

variety of compensation processing schemes, it is particularly well suited for the wavelet-based TVFD algorithm described above [12, 15]. This technique will also be evaluated for applicability as a front-end processor to general-purpose speech-enhancement systems and automatic speech-recognition systems.

Acknowledgments

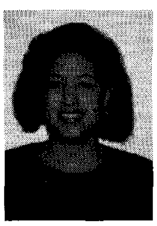
The authors wish to thank Prof. R. Coifman (Yale University) and Dr. M. Randolph (Motorola Corporation) for their encouragement and useful suggestions. A portion of this work was performed while N. Whitmal was a summer intern at the Motorola Corporate Research Laboratories, Schaumburg, Ill.

The work of N. Whitmal was supported in part by a National Science Foundation Graduate Fellowship, by the National Science Foundation under grant BCS-9110247, the Buehler Center on Aging, the Motorola Corporation, and a Monsanto Foundation Faculty Internship. The work of J. Rutledge was supported in part by the National Science Foundation under grant BCS-9110247 and the Buehler Center on Aging. The work of J. Cohen was supported in part by a DePaul University Summer Research Grant.



Nathaniel A. Whitmal, III, was born in Chicago, IL, in 1964. He received the B.S. degree in electrical engineering from the Massachusetts Institute of Technology, Cambridge, MA, in 1986; the M.S. degree in engineering management from the Gordon Institute of Tufts University, Medford, MA, in 1990; and the M.S. degree in electrical engineering from Northwestern University, Evanston, IL, in 1993.

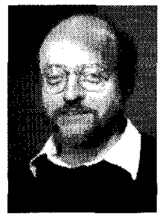
From 1986 to 1992 he was a member of the Technical Staff of Bose Corporation, Framingham, MA, where he was involved in design and development of audio products for international markets. He is currently a Ph.D candidate in the Department of Electrical and Computer Engineering at Northwestern University, where he also holds the position of Monsanto Faculty Intern. His current research interests are in the areas of signal processing and acoustics, with special emphasis on speech enhancement.



Janet C. Rutledge received the B.S. degree in electrical engineering from Rensselaer Polytechnic Institute, Troy, NY, in 1983. She received the M.S. and Ph.D. in electrical engineering from the Georgia Institute of Technology, Atlanta, GA, in 1984 and 1990, respectively. She had a scholarship and fellowship from Bell Laboratories for the B.S. and M.S., and a fellowship from Bell Communications Research for the Ph.D.

Since August 1990 she has been an assistant professor in the Electrical and Computer Engineering Department at Northwestern University, Evanston, IL, with an adjunct appointment in the Audiology and Hearing Sciences Dept. Her research interests include modeling and compensating for the effects of sensorineural hearing loss, speech recognition, speech enhancement, and audio signal processing.

Professor Rutledge is a member of the speech technical committee of the Acoustical Society of America and a Fellow of the Hugh Knowles Center for Clinical and Basic Science in Hearing and its Disorders at Northwestern University. She serves on the NIH NIDCD Ad Hoc Committee on Hearing Aid Research.



Jonathan Cohen is a professor and chairman of the Mathematics Department at DePaul University. Professor Cohen received his Ph.D. in harmonic analysis from the Department of Mathematics, Washington University, St. Louis, MO, in 1977. His work includes papers on singular integrals, partial differential equations, potential theory, real variable theory and wavelet analysis in the processing of speech signals.

Dr. Cohen is a member of the American Mathematics Society and the Society for Industrial and Applied Mathematics.

Address for Correspondence: *Nathaniel A. Whitmal* and *Janet C. Rutledge* - Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208; *Jonathan Cohen* Department of Mathematical Sciences, DePaul University, Chicago, IL 60614.

References

1. **Plomp R:** Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J Acoust Soc Amer* 63:533-549, 1978.
2. **Tillman TW, Carhart R, Olsen, WO:** Hearing aid efficiency in a competing speech situation. *J Speech Hear Res* 13:789-811, 1970.
3. **French NR, Steinberg JC:** Factors governing the intelligibility of speech sounds. *J Acoust Soc Amer* 19:91-119, 1947.
4. **Zurek PM, Delhorne LA:** Consonant reception in noise by listeners with mild and moderate sensorineural hearing impairment. *J Acoust Soc Amer* 82:1548-1559, 1987.
5. **Killion, M:** The K-Amp hearing aid: an attempt to present high fidelity for persons with impaired hearing. *Amer J Audiology* 2:52-73, 1993.
6. **Weiss M:** Use of an adaptive noise canceler as an input preprocessor for a hearing aid. *J Rehab Res Dev* 24:93-102, 1987.
7. **Makhoul J, McAulay R:** Removal of Noise from Noise-Degraded Speech Signals, National Academy of Sciences, National Academy Press, Washington DC, 1989.
8. **Boll SF:** Speech enhancement in the 1980's: noise suppression with pattern matching. In: Furui S, Sondhi MM: Advances in Speech Signal Processing. Marcel Dekker, New York NY, 1991.
9. **Levitt H, Bakke M, Kates J, Neuman A, Schwander T, Weiss M:** Signal processing for hearing impairment. *Scand Audiol* S38:7-19, 1993.
10. **Boll SF:** Suppression of acoustic noise in speech using spectral subtraction. *IEEE Trans Acoust Speech Sig Proc* 27:113-120, 1979.
11. **Kates J:** Speech enhancement based on a sinusoidal model. *J Speech Hear Res* 37:449-64, 1994.
12. **Rutledge JC:** Speech enhancement for hearing aids. In: Akay M (ed): Time-Frequency and Wavelets in Biomedical Engineering. IEEE Press: work in press.
13. **Coifman RR, Wickerhauser MV:** Entropy-based algorithms for best basis selection. *IEEE Trans Inf Theory* 38:713-738, 1992.
14. **Rutledge JC, Clements M:** Compensation for recruitment of loudness in sensorineural hearing impairments using a sinusoidal model of speech. *Proc ICASSP-91* 5:3641-3644, 1991.
15. **Drake LA, Rutledge JC, Cohen J:** Wavelet analysis in recruitment of loudness compensation. *IEEE Trans Sig Proc* 12:3306-3312, 1993.
16. **Berger J, Coifman RR, Goldberg M:** Removing noise from music using local trigonometric bases and wavelet packets. *J Audio Engng Soc*: 808-818, 1994.
17. **Saito N:** Simultaneous noise suppression and signal compression using a library of orthonormal bases and the minimum description length criterion. In: Foufoula-Georgiou E, Kumar P: Wavelets in Geophysics. Academic Press, 1994.
18. **Donoho D:** De-noising by soft thresholding. *IEEE Trans Inf Theory* 41: 613-627, 1995.
19. **Saito N, Coifman RR:** Local discriminant bases and their applications. *J Math Imag Vision*: 5:337-358, 1995.

20. **Rissanen, J:** Modeling by shortest data description. *Automatica* 14: 465-471, 1978.

21. **Pesquet J-C, Krim H, Carfantan H, Proakis JG:** Estimation of noisy signals using time-invariant wavelet packets. *27th Asilomar Conf Sig Syst Comp*: 635-660, 1995.

22. **Rissanen, J:** *Stochastic Complexity in Statistical Inquiry*. World Scientific Publishing: Singapore, 1989.

23. **Kay S, Nagesha V:** Maximum likelihood estimation of signals in autoregressive noise. *IEEE Trans Sig Proc* 42: 88-101, 1994.

24. **Box GEP, Jenkins GM:** *Time Series Analysis — Forecasting and Control*. Holden Day: San Francisco, 1970.

25. **Coifman RR, Donoho DL:** Translation-invariant de-noising. In: Antoniadis A (ed): *Wavelets and Statistics*. Springer-Verlag: 1995.

26. **Sinha D, Tewfik AH:** Low bit rate transparent audio compression using adapted wavelets. *IEEE Trans Sig Proc* 41: 3463-3479, 1993.

27. **Krim H, Pesquet J-C:** On the statistics of best bases criteria. In: Antoniadis A (ed): *Wavelets and Statistics*. Springer-Verlag: 1995.

28. **Tyler R, Johnson TJ:** Attenuating noise during pauses in speech. NIH Interdisciplinary Forum on Hearing Aid Research and Development, September, 1995.

29. **Whitmal NA, Rutledge JC, Cohen J:** Wavelet-based noise reduction. *Proc ICASSP-95*: 3003-3006, 1995.

30. **Fletcher H, Galt RH:** Perception of speech and its relation to telephony. *J Acoust Soc Amer* 22:89-151, 1950.

News Notes

1996 Biomedical Engineering Rankings

According to *U.S. News and World Report*, the following schools represent the top 20 biomedical engineering programs in the United States:

1. Johns Hopkins University
2. Duke University
3. Massachusetts Institute of Technology
4. University of California, San Diego
5. University of Washington
   Case Western Reserve University
7. University of Pennsylvania
8. University of Utah
9. University of Michigan
10. Georgia Institute of Technology
11. University of California, Berkeley
12. Northwestern University
13. Carnegie Mellon University
    Rice University
    Stanford University
16. Columbia University
    University of Virginia
18. Boston University
    Rensselaer Polytechnic Institute
    University of Texas, Austin

Adapted from The AIMBE News, *Spring 1996*