Subspace algorithms for noise reduction in cochlear implants (L)

Philipos C. Loizou,[a] Arthur Lobo, and Yi Hu
Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75083-0688

(Received 28 March 2005; revised 22 August 2005; accepted 22 August 2005)

A single-channel algorithm is proposed for noise reduction in cochlear implants. The proposed algorithm is based on subspace principles and projects the noisy speech vector onto "signal" and "noise" subspaces. An estimate of the clean signal is made by retaining only the components in the signal subspace. The performance of the subspace reduction algorithm is evaluated using 14 subjects wearing the Clarion device. Results indicated that the subspace algorithm produced significant improvements in sentence recognition scores compared to the subjects' daily strategy, at least in stationary noise. Further work is needed to extend the subspace algorithm to nonstationary noise environments. © *2005 Acoustical Society of America.* [DOI: 10.1121/1.2065847]

PACS number(s): 43.66.Ts, 43.71.Ky [AJO] Pages: 2791–2793

I. INTRODUCTION

Several noise-reduction algorithms have been proposed for cochlear implant (CI) users (van Hoesel and Clark, 1995; Hamacher *et al.*, 1997; Wouters and Vanden Berghe, 2001). Most of these algorithms, however, were based on the assumption that two or more microphones were available. van Hoesel and Clark (1995) tested an adaptive beamforming technique with four Nucleus-22 implantees using signals from two microphones—one behind each ear—to reduce noise coming from 90° of the patients. Results indicated that adaptive beamforming with two microphones can bring substantial benefits to CI users in conditions for which reverberation is moderate, and only one source is predominantly interfering with speech. Adding, however, a second microphone contralateral to the implant is ergonomically difficult without requiring the CI users to wear headphones or a neckloop (bilateral implants might provide the means, but their benefit is still being investigated). Alternatively, monaural multimicrophone techniques can be used and such techniques are now becoming commercially available (e.g., BEAM in Nucleus devices).

In general, single-microphone noise reduction algorithms are more desirable and cosmetically more appealing than the algorithms based on multiple-microphone inputs. A few single-microphone noise-reduction strategies (Weiss, 1993; Hochberg *et al.*, 1992; Yang and Fu, 2005) have been proposed for cochlear implants, some of which were implemented on old cochlear implant processors based on feature extraction strategies (F0/F1/F2 and MPEAK strategies) and some of which were implemented on the latest processors. Weiss (1993) demonstrated that preprocessing the signal with a standard noise reduction algorithm could reduce the errors in formant extraction. The latest speech processors, however, are not based on feature extraction strategies but are based on vocoder-type strategies. Recently, Yang and Fu (2005) evaluated a spectral-subtractive algorithm using the latest implant processors. Significant benefits in sentence recognition were observed for all subjects with the spectral-subtractive algorithm, particularly for speech embedded in speech-shaped noise.

In brief, only a few studies (e.g., Yang and Fu, 2005) were conducted to investigate the benefits of preprocessing the noisy speech signal by a noise reduction algorithm and feeding the enhanced signal to implant listeners. In the present study we evaluate the performance of a subspace noise reduction algorithm that is used as a preprocessor for signal enhancement.

II. EXPERIMENT 1: EVALUATION OF SUBSPACE ALGORITHM

In this experiment, we investigate the potential benefits of first preprocessing the noisy signal with a noise reduction algorithm and then feeding the "enhanced" signal to the CI processor. For noise reduction, we use a custom subspace-based algorithm (Hu and Loizou, 2002).

A. Subjects

A total of 14 Clarion implant users participated in this experiment consisting of 9 Clarion CII patients and 5 Clarion S-series patients. The majority of the CII patients were fitted with the CIS strategy, and the S-series patients were fitted with the SAS strategy. All subjects had at least 1 yr of experience with their implant device (see Table I).

B. Subspace algorithm

The signal subspace algorithm was originally developed by Ephraim and Van Trees (1995) for white input noise and was later extended to handle colored noise (e.g., speech-shaped noise) by Hu and Loizou (2002). The underlying principle of the subspace algorithm is based on the projection of the noisy speech vector (consisting of, say, a segment of speech) onto two subspaces: the "signal" subspace and the "noise" subspace. The noise subspace contains only signal

[a]Address correspondence to Philipos C. Loizou, Ph.D., Department of Electrical Engineering, University of Texas at Dallas, P.O. Box 830688, EC 33, Richardson, Texas 75083-0688. Electronic mail: loizou@utdallas.edu; phone: (972) 883-4617; fax: (972) 883-2710.

TABLE I. Subject information.

Subject	Age	Implant	CI use (yr)	HINT score (quiet)
S1	41	Clarion CII	2	57
S2	26	Clarion CII	2	55
S3	39	Clarion CII	1	90
S4	41	Clarion CII	2	52
S5	70	Clarion CII	2	60
S6	55	Clarion CII	1	86
S7	58	Clarion CII	2	88
S8	66	Clarion CII	3	95
S9	38	Clarion CII	4	35
SS1	56	Clarion S series	1	60
SS2	45	Clarion S series	1	94
SS3	40	Clarion S series	1	55
SS4	52	Clarion S series	1	79
SS5	43	Clarion S series	1	80

components due to the noise, and the signal subspace contains primarily the clean signal. Therefore, an estimate of the clean signal can be made by removing the components of the signal in the noise subspace and retaining only the components of the signal in the signal subspace.

Let $\mathbf{y}$ be the noisy vector, and let $\hat{\mathbf{x}}=\mathbf{Hy}$ be an estimate of the clean signal vector, where $\mathbf{H}$ is a transformation matrix. The noise reduction problem can be formulated as that of finding a transformation matrix $\mathbf{H}$, which, when applied to the noisy vector, would yield the clean signal. After applying such a transformation to the noisy signal, we can express the error between the estimated signal $\hat{\mathbf{x}}$ and the true clean signal $\mathbf{x}$ as $\epsilon=\hat{\mathbf{x}}-\mathbf{x}=(\mathbf{H}-\mathbf{I})\mathbf{x}+\mathbf{H}\ \mathbf{n}$, where $\mathbf{n}$ is the noise vector. Since the transformation matrix will not be perfect, it will introduce some speech distortion, which is quantified by the first term of the error term, i.e., by $(\mathbf{H}\text{-}\mathbf{I})\mathbf{x}$. The second term $(\mathbf{H}\ \mathbf{n})$ quantifies the amount of noise distortion introduced by the transformation matrix. As the speech and noise distortion (as defined above) are decoupled, one can find the optimal transformation matrix $\mathbf{H}$ that would minimize the speech distortion subject to the noise distortion falling below a preset threshold. The solution to this constrained minimization problem for colored noise is given by (Hu and Loizou, 2002):

$$\mathbf{H}_{\mathbf{opt}}=\mathbf{V}^{-\mathbf{T}}\mathbf{\Lambda}(\mathbf{\Lambda}+\mu\mathbf{I})^{-1}\mathbf{V}^{\mathbf{T}}, \tag{1}$$

where μ is a parameter (typical values for $\mu=1-20$), $\mathbf{V}$ is an eigenvector matrix, and $\mathbf{\Lambda}$ is a diagonal eigenvalue matrix obtained from the noisy speech vector (more details can be found in Hu and Loizou, 2002, 2003). In our implementation, we used a variable μ that took values in the range of 1 to 20 depending on the estimated short-term signal-to-noise ratio (see Hu and Loizou, 2003).

The above equation has the following interesting interpretation. The matrix $\mathbf{V}^{\mathbf{T}}$ acts like a data-dependent transform and projects the noisy speech vector into the noise and signal subspaces. The diagonal matrix $\mathbf{\Lambda}(\mathbf{\Lambda}+\mu\mathbf{I})^{-1}$ multiplies the components of the signal in the signal subspace by a gain while zeroing out the components of the signal in the noise subspace. Finally, the matrix $\mathbf{V}^{-\mathbf{T}}$ transforms back the projected signal, i.e., it acts like an inverse transform.

The implementation of the above signal subspace algorithm can be summarized into two steps. Step (1): For each frame of noisy speech $(\mathbf{y})$, use the above transformation given in Eq. (1) to obtain an estimate of the clean signal vector $\hat{\mathbf{x}}$, i.e., $\hat{\mathbf{x}}=\mathbf{H}_{\mathbf{opt}}\mathbf{y}$. Step (2): Use the estimated signal $\hat{\mathbf{x}}$ as input to the CI processor.

The above estimator was applied to 4 ms duration frames of the noisy signal, which overlapped each other by 50%. The enhanced speech vectors were Hamming windowed and combined using the overlap and add approach. No voice activity detection algorithm was used in our approach to update the noise covariance matrix needed to compute the matrix $\mathbf{V}$. The noise covariance matrix was estimated using speech vectors from the initial silent frames of the sentences. Although this procedure for estimating the noise covariance matrix is adequate for stationary noise (such as the one used in this study), it is not adequate for nonstationary environments in which the background spectra (and consequently the noise covariance matrices) constantly change. In nonstationary environments (e.g., restaurant noise), the noise covariance matrix could be estimated and updated whenever a speech-absent segment is detected based on a voice activity detector or a noise-estimation algorithm.

C. Procedure

HINT sentences (Nilsson *et al.*, 1994) corrupted in +5 dB S/N speech-shaped noise (taken from the HINT database) were used for evaluation. Six lists (60 sentences) were processed offline in MATLAB by the subspace noise reduction algorithm. The processed sentences were presented directly to the subjects via the auxiliary input jack of their CI processor at a comfortable listening level. Subjects were fitted with their daily strategy. For comparative purposes, subjects were also presented with six different lists (60 sentences) of HINT sentences corrupted in +5 dB speech-shaped noise, i.e., unprocessed sentences. The presentation order of preprocessed and unprocessed sentences was randomized between subjects.

D. Results and discussion

The sentences were scored in terms of the percent of words identified correctly (all words were scored). Figure 1 shows the percent correct scores for all subjects. The mean score obtained with sentences preprocessed by the subspace algorithm was 44% correct, and the mean score obtained with unprocessed sentences was 19% correct. ANOVA (repeated measures) tests indicated that the sentence scores obtained with the subspace algorithm were significantly higher $[F(1,13)=33.1,\ p<0.0005]$ than the scores obtained with the unprocessed sentences. As can be seen from Fig. 1, most subjects benefited from the noise reduction algorithm. Subject's SS4 score, for instance, improved from 0% correct to 40% correct. Similarly, subjects' SS1 and SS2 scores improved from roughly 0% to 50% correct.

The above results indicate that the subspace algorithm can provide significant benefits to CI users in regard to the

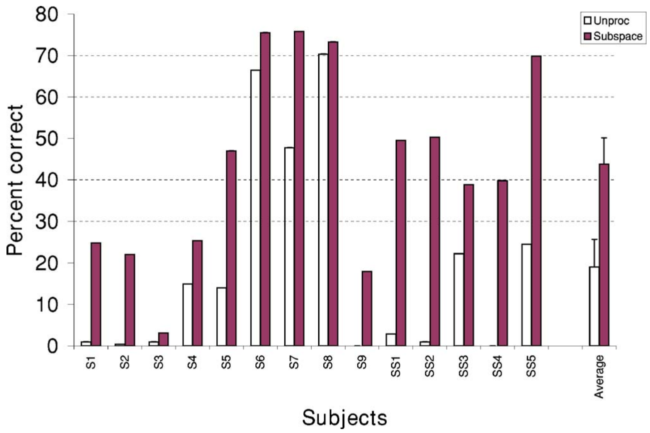

FIG. 1. (Color online) The subjects' performance on the identification of words in sentences embedded in +5 dB S/N speech-shaped noise and pre-processed (dark bars) by the subspace algorithm or left unprocessed (white bars). Subjects S1–S9 were Clarion CII patients and subjects SS1–SS5 were Clarion S-series patients. Error bars indicate standard errors of the mean.

recognition of sentences corrupted by stationary noise. It should be noted that the above signal subspace algorithm was only tested in stationary noise, and it is not clear whether such an intelligibility benefit would be maintained if the algorithm was tested in nonstationary environments (e.g., restaurant, multitalker babble). Further work is needed to extend the subspace algorithm to nonstationary noise environments, particularly with regard to updating the noise covariance matrix based on perhaps a voice activity detector or a noise-estimation algorithm.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01-DC03421 from NIDCD/NIH. The authors would like to thank the anonymous reviewers and Dr. Andrew Oxenham for their valuable comments.

Ephraim, Y., and Van Trees H. L. (**1995**). "A signal subspace approach for speech enhancement," IEEE Trans. Speech Audio Process. **3**, 251–266.

Hamacher, V., Doering, W., Mauer, G., Fleischmann, H., and Hennecke, J. (**1997**). "Evaluation of noise reduction systems for cochlear implant users in different acoustic environments," Am. J. Otol. 18, S46–S49.

Hochberg, I., Boorthroyd, A., Weiss, M., and Hellman, S. (**1992**). "Effects of noise and noise suppression on speech perception by cochlear implant users," Ear Hear. **13**, 263–271.

Hu, Y., and Loizou P. (**2002**). "A subspace approach for enhancing speech corrupted with colored noise," IEEE Signal Process. Lett. **9**, 204–206.

Hu, Y., and Loizou, P. (**2003**). "A generalized subspace approach for enhancing speech corrupted with colored noise," IEEE Trans. Speech Audio Process. **11**, 334–341.

Nilsson, M., Soli, S., and Sullivan, J. (**1994**). "Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise," J. Acoust. Soc. Am. **95**, 1085–1099.

van Hoesel, R., and Clark, G. (**1995**). "Evaluation of a portable two-microphone adaptive beamforming speech processor with cochlear implant patients," J. Acoust. Soc. Am. **97**, 2498–2503.

Weiss, M. (**1993**). "Effects of noise and noise reduction processing on the operation of the Nucleus-22 cochlear implant processor," J. Rehabil. Res. Dev. **30**, 117–128.

Wouters, J., and Vanden Berghe, J. (**2001**). "Speech recognition in noise for cochlear implantees with a two-microphone monaural adaptive noise reduction system," Ear Hear. **22**, 420–430.

Yang, L., and Fu, Q. (**2005**). "Spectral subtraction-based speech enhancement for cochlear implant patients in background noise," J. Acoust. Soc. Am. **117**, 1001–1004.