Experimental Evaluation of a Nested Microphone Array with Adaptive Noise Cancellers

Y. Rosa Zheng
Missouri University of Science and Technology, zhengyr@mst.edu

M. El-Tanany

R. A. Goubran

Recommended Citation

Experimental Evaluation of a Nested Microphone Array With Adaptive Noise Cancellers

Yahong R. Zheng, Rafik A. Goubran, *Member, IEEE*, and Mohamed El-Tanany

Abstract—This paper proposes a near-field broadband adaptive beamforming scheme for intelligent computer telephony and teleconferencing applications, namely the nested microphone array with adaptive noise canceller (NMA-ANC). The NMA-ANC scheme incorporates an harmonically nested array with a nonuniformly subbanded multirate filter bank. Each subband array employs several near-field delay-filter-and-sum beamformers and an adaptive noise canceller (ANC). The proposed NMA-ANC is evaluated via a noise rejection experiment and dereverberation experiment performed in an anechoic chamber and a real conference room, respectively. The experiment data are recorded by a multichannel digital recording system developed using commercial off-the-shelf (COTS) equipments. A perceptual analysis/measurement system (PAMS) test is also carried out using a COTS digital speech level analyzer. The results of the experimental evaluation and PAMS test show that the proposed NMA-ANC scheme is able to improve the sound quality by adaptively rejecting multiple interfering signals and attenuating the reverberant noises and avoiding the desired signal cancellation.

Index Terms—Adaptive noise canceller, broadband beamforming, dereverberation, interference rejection, multichannel digital recording, near-field beamforming, nested microphone array (NMA), subband multirate systems.

I. INTRODUCTION

INTELLIGENT teleconferencing and computer telephony are the emerging communication systems involving intelligent multiple sensors and networking [1]. These systems generally employ multiple video cameras and microphone arrays. The multisensor system is expected to interact with the unknown environment, learn to configure its own model of the surroundings, and operate itself without complete control of a human operator [2]. This means that the system will be able to adaptively monitor the sound sources in the room, track the desired talker, and control the camera to the desired scene [3].

Microphone arrays play important roles in these intelligent communication systems. They provide hands-free auditory data acquisition. Array processing of the received data falls into two basic categories. One is beamforming, which enhances the quality of the desired sound signal and suppresses the interfering signals and environmental noises. Another is source

localization and tracking, which guides the video cameras and the microphone beamformers to pick up the desired signal sources. These two types of microphone array processing may be utilized in several layers of multisensor fusion and integration [2].

This paper focuses on microphone array beamforming for speech quality enhancement. The technical challenges involved in microphone array beamforming are three-fold. First, broadband beamforming is required because speech and audio signals are broadband signals with high degrees of nonstationarity. Broadband beamformers are difficult to design due to the frequency dependent array properties [4]. Second, near-field beamforming is needed because the signal sources are generally located within the near field of the array aperture. This means that the simplified far field assumption is not valid and it can cause severe degradation of the array performance in the near-field scenario [5]. The curvature of the signal propagation has to be taken into account which requires near-field beamforming techniques. The third challenge is the desired signal cancellation phenomena [6] encountered by adaptive beamforming in reverberant environments. Conventional adaptive beamformers would cancel the desired signal at their output due to the high correlation of the desired direct path signal and the reflected multipath signals.

As a compromise solution to the three technical challenges, we have proposed a multirate subband beamformer in previous works [8]–[10]. It incorporates a harmonically nested array with a nonuniformly subbanded multirate filter bank. Each subband employs an octave-band adaptive beamformer designed and implemented by a generalized sidelobe canceller (GSC) [9]. In this paper, we replace the adaptive GSC beamformer of each subband by a switched-beam adaptive noise canceller (ANC). The switched-beam ANC consists of several near-field delay-filter-and-sum (DFS) beamformers followed by a conventional ANC. The resulting system is called the nested microphone array with adaptive noise cancellers (NMA-ANC). The advantages of the NMA-ANC scheme include the better robustness against location errors, higher interference rejection performance, and easier mitigation of the desired signal cancellation in reverberant environments [11]. The robustness against location error is critical for practical implementations because the localization of the sound source is generally accurate only within a couple of degrees [7].

This paper also evaluates the performances of the NMA-ANC using the experimental data recorded in an anechoic chamber and a real conference room. Several near-field signal sources are recorded by a multichannel digital recording system. The experimental results show that the NMA-ANC achieves high noise

Y. R. Zheng is with the Department of Electrical and Computer Engineering, University of Missouri-Columbia, Columbia, MO 65211 USA (e-mail: zhengyr@missouri.edu).

R. A Goubran and M. El-Tanany are with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON K1S 5B6, Canada.
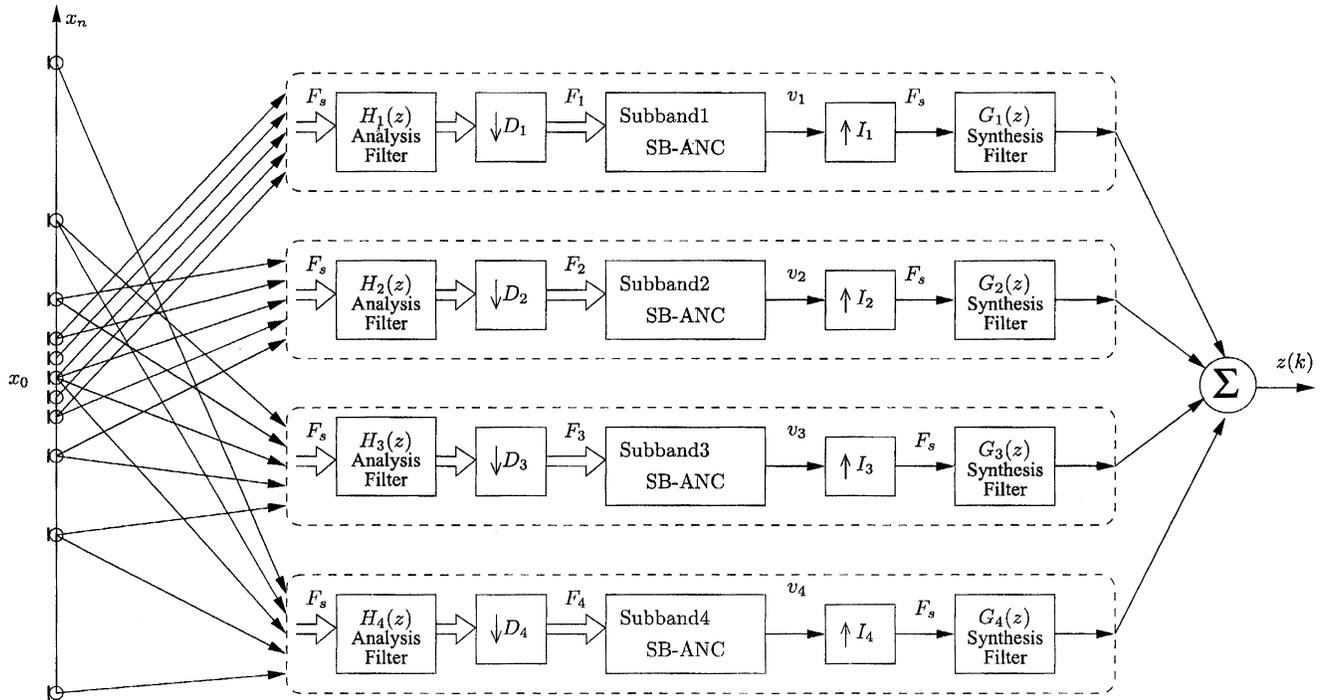
Fig. 1.   Structure of the NMA with ANC (NMA-ANC). The 11-element microphone array is grouped into four subarrays, each having five elements and covering an octave subband. The structure of the switched-beam ANC used in each subarray is detailed in Fig. 5.

reduction (NR) factor and de-reverberation gain. A perceptual analysis/measurement system (PAMS) test is also included to verify the experimental results.

The rest of this paper is organized as follows. Section II introduces the structure of the NMA-ANC, and its design and implementation for wideband computer telephony. Section III describes the experimental equipment, measurement procedures, and environments. Section IV presents the data processing techniques and the performances of the NMA-ANC in terms of interference rejection and de-reverberation. Section V draws the conclusion.

II. NESTED MICROPHONE ARRAY ANC

The general structure of the NMA-ANC is depicted in Fig. 1. It consists of a nested microphone array grouped into several subarrays, an analysis filter bank with a down sampler for each subband, a switched-beam ANC for each subband, and an up sampler for each subband, and a synthesis filter bank. Signals received by the nested array elements are sampled at a high frequency F_s. The sampled data are grouped into several subarrays. Each subarray is processed by its corresponding analysis filter $H_i(z)(i = 1, 2, \ldots)$, and then decimated by the down sampler D_i. After the decimation, the switched-beam ANC of each subarray operates at a lower sampling rate F_i, where $F_i = F_s/D_i$. The ANCs' outputs are interpolated by the up-samplers I_i and combined via the synthesis filters $G_i(z)$.

The general structure of the NMA-ANC is now applied to the wideband telephony or teleconferencing with specific design and implementations. The following subsections provide the details of the nested array, the analysis and synthesis filters, and the switched-beam ANC.

A. Harmonically-Nested Array

The wideband telephony utilizes the frequency band of [50, 7000] Hz and the sampling frequency of $F_s = 16$ kHz, according to the G.722 standard [12]. The harmonically nested array is designed to cover the frequency range $B = [50, 7200]$ Hz. It is composed of four equispaced linear subarrays, each having $M = 5$ elements, as shown in Fig. 2. Subarray1 is designed for the highest frequency range [3600, 7200] Hz. The inter-element spacing is $d = 2.4$ cm, which is less than half the wavelength of the high frequency edge. Subarray2 is designed for the frequency range [1800, 3600] Hz with inter-element spacing being $2d$. Subarray1 is nested within Subarray2 with $(M + 1)/2$ superimposed elements, assuming M is odd. Subarray3 is designed similarly to cover the subband $B_3 = [900, 1800]$ Hz with inter-element spacing of $4d$. Subarray4 covers the remaining subband $B_4 = [50, 900]$ Hz with inter-element spacing of $8d$. The compound nested array has a total of 11 elements with a total size of 76.8 cm. If larger array size and system complexity are allowed, more subarrays may be added to further enhance the performance of the low subband. In practice, the system of four subarrays provides satisfactory performance with a reasonable complexity.

B. Multirate Filter Bank

Each subarray requires an analysis filter and a synthesis filter to avoid aliasing and imaging. With smaller bandwidth covered by each subarray, temporal multirate sampling is incorporated with the nested array via down-samplers and up-samplers. The analysis filters $H_i(z)$ and the down-samplers D_i can be implemented by a multistage tree structure, as depicted in Fig. 3(a). Each stage of the tree consists of a high-pass filter $HP_i(z)$, a low-pass filter $LP_i(z)$ and down-samplers. The high-pass and
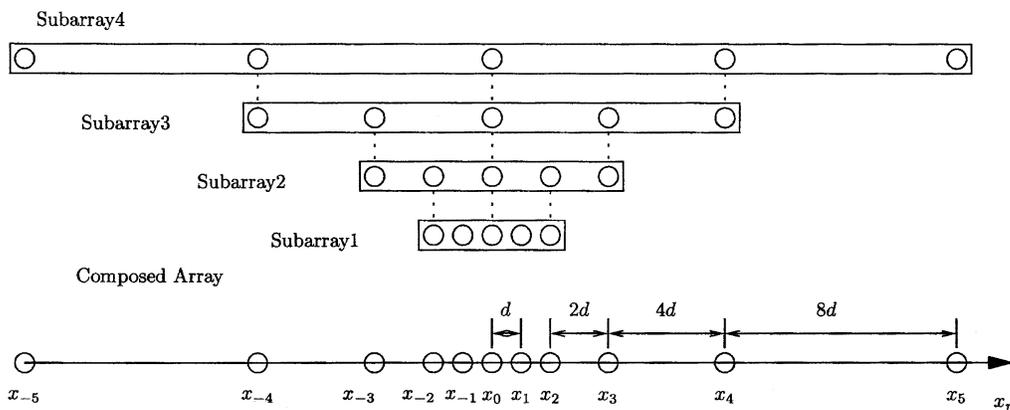
Fig. 2. Super-imposed configuration of the 11-element harmonically nested array.
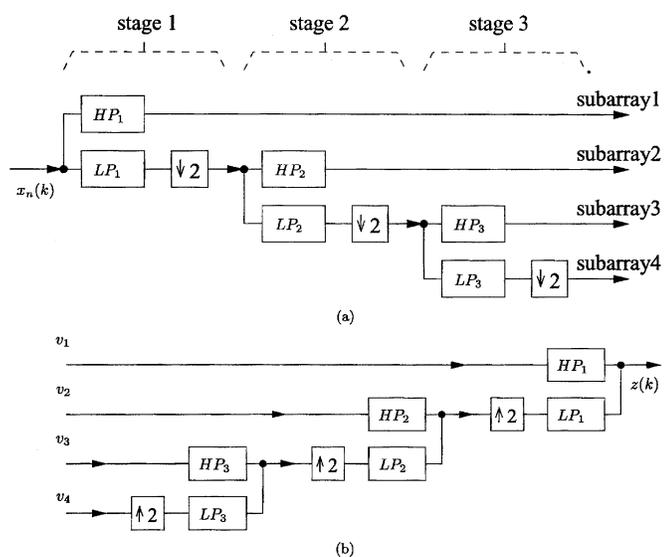

Fig. 3. Tree-structured analysis and synthesis filters for noncritical sampling multirate subband systems. (a) Analysis filter. (b) Synthesis filter.
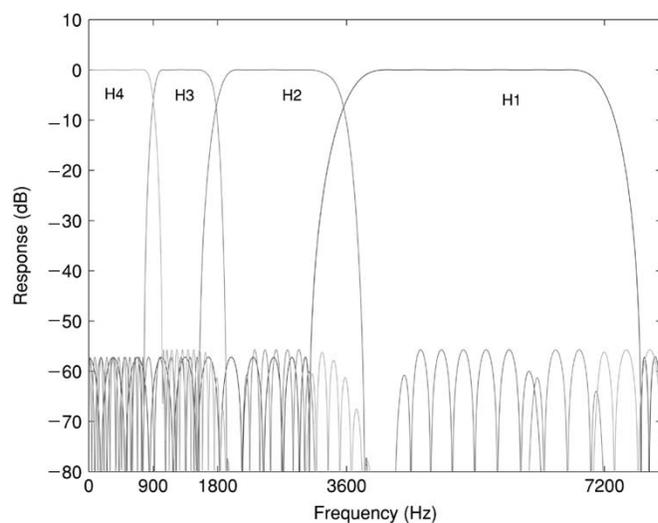

Fig. 4. Frequency response of the analysis filter bank.

low-pass filters are related to the parallel filters $H_i(z)$ in Fig. 1 as

$$
\begin{aligned}
H_1(z) &= HP_1(z) \\
H_2(z) &= LP_1(z) * HP_2(z^2) \\
H_3(z) &= LP_1(z) * LP_2(z^2) * HP_3(z^4) \\
H_4(z) &= LP_1(z) * LP_2(z^2) * LP_3(z^4).
\end{aligned}
\tag{1}
$$

The synthesis filters $G_i(z)$ are the mirror images of the analysis filters and can also be implemented by a tree structure, as shown in Fig. 3(b).

In each stage of the tree, a 49-tap high-pass filter and a 49-tap low-pass FIR filter are designed by Remez method. The equivalent parallel filters have a stop band attenuation of 60 dB and the normalized transition band of 0.0625. The frequency responses of the analysis filters are shown in Fig. 4.

C. Switched-Beam ANC

The switched-beam ANC used in each subband array is illustrated in Fig. 5. It consists of three functional blocks: the array

beamformers, the switches, and the ANC. The signals received at the M-element array are fed into several predesigned beamformers. Each beamformer focuses at a separate spatial location without adapting to the signal environment. The switches select the desired beam as the primary channel and other beams as the auxiliary channels for the ANC. The ANC is a standard adaptive filter which adaptively cancels the noise components in the primary channel and tries to achieve a higher signal-to-interference-and-noise ratio (SINR) at the output.

As an example, three near-field DFS beamformers are predesigned to focus at $\mathbf{x}_{S1} = (0.6 \text{ m}, 90^\circ), \mathbf{x}_{S2} = (1.0 \text{ m}, 50^\circ), \mathbf{x}_{S3} = (1.0 \text{ m}, 120^\circ)$, respectively, where the spherical coordinates are defined such that the origin is at the array phase center and the array axis is at 0°. Each DFS beamformer had 16 tap per element. The array responses are plotted in Fig. 6 at the four in-band frequencies 500, 1800, 3500, and 6800 Hz.

The beamformers are followed by the ANC with two auxiliary channels. Each auxiliary channel of the ANC uses 32 adaptive taps. The delay in the primary channel is set to $Q = 16$.
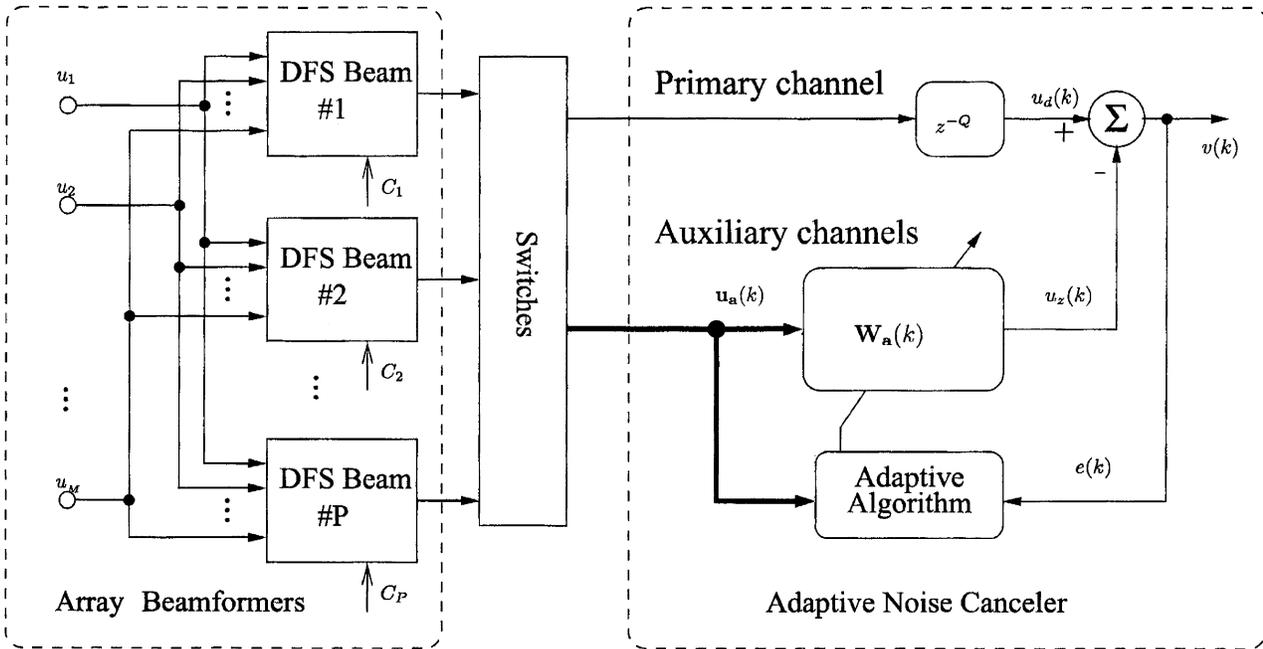
Fig. 5.   Structure of the switched-beam adaptive noise canceller (SB-ANC).

III. DESCRIPTION OF THE EXPERIMENTS

A. Measurement Apparatus

A multichannel audio recording system, as shown in Fig. 7, was used for microphone array recordings. The system consisted of two parts: the generator and the recorder. The generator used a compact disk (CD) player to produce the sound sources. The recorder consisted of high quality microphones, multichannel preamplifiers, multichannel A/D converters, and a personal computer. The details of the apparatus are listed in Table I. All apparatus are commercial off-the-shelf (COTS) products.

The "clean" sound source was recorded on a CD recordable disk. The signal source was 16-bit PCM waveform sampled at 44.1 kHz. The length of the signal source was approximately 2 min. It consisted of segments of white noise, a cue frame, male speech, female speech and music clippings. Each segment of the signal was separated by 4 s of silence so it could be easily extracted from the recorded data. The white noise segment at the beginning of the signal source was sufficiently long to allow manual operation of the play and record buttons of the CD player. The cue frame was designed for synchronization of nonsimultaneously recorded data. It consisted of three single frequency tones of 300 Hz, 3.0 kHz, and 7.0 kHz. Each single tone had a length of 0.2 s. The speech and music segments were broadband signals originally sampled at 16 kHz or higher. They were resampled to 44.1 kHz for recording onto the CD disk.

The microphones were omni-directional condenser microphones possessing a flat frequency response from 50 Hz to 15 kHz. They were mounted on a plywood board with the nested array geometry as shown in Fig. 2. There were a total of 11 microphones nested into four subarrays each having five elements. The inter-element spacings of the four subarrays were 2.4, 4.8, 9.6, and 19.2 cm, respectively. The total size of the nested microphone array was 76.8 cm.

Other apparatus were mounted on a metal equipment rack. The multichannel preamplifier had eight mono inputs for microphones with 70-dB gain range. Each input channel had a filter with the low cutoff frequency at 100 Hz. The MIDIMAN digital recording system [13] was configured as a PCI host card plus an external rack-mount unit, which housed the A/D (and D/A) converters. It had eight data channels with word length and sampling rates up to 24-bit/96 kHz. A/D converters had a high dynamic range (A-weighted measured) of 109 dB, and low distortion (measured THD@0 dBFS) of less than 0.001. The PCI host card of the MIDIMAN digital recording system was installed in the personal computer. A multichannel recording software was also installed and configured in the personal computer for eight-channel digital recording.

After the setup of the equipment, channel calibration was performed for each microphone channel by adjusting the gain of the preamplifier. The calibration ensured that no amplitude clipping occurred and the gains of the received signals were accurate within 0.25 dB.

B. Measurement Procedures and Environments

The experiments were performed in an anechoic chamber and a small conference room, respectively. The source CD was played back at several locations and the microphone recordings were done separately for each location. The data were recorded in 16 bit PCM format with a rate of 48 000 samples per second.

Multiple runs of recording were required for each location because our recording equipment had only eight channels available and the 11-element array could not be recorded simultaneously. Each recording run used seven channels to simultaneously record the seven elements of two adjacent subarrays—the first recording run for the seven elements of Subarray1 and Subarray2, the second run for Subarray2 and Subarray3, and the
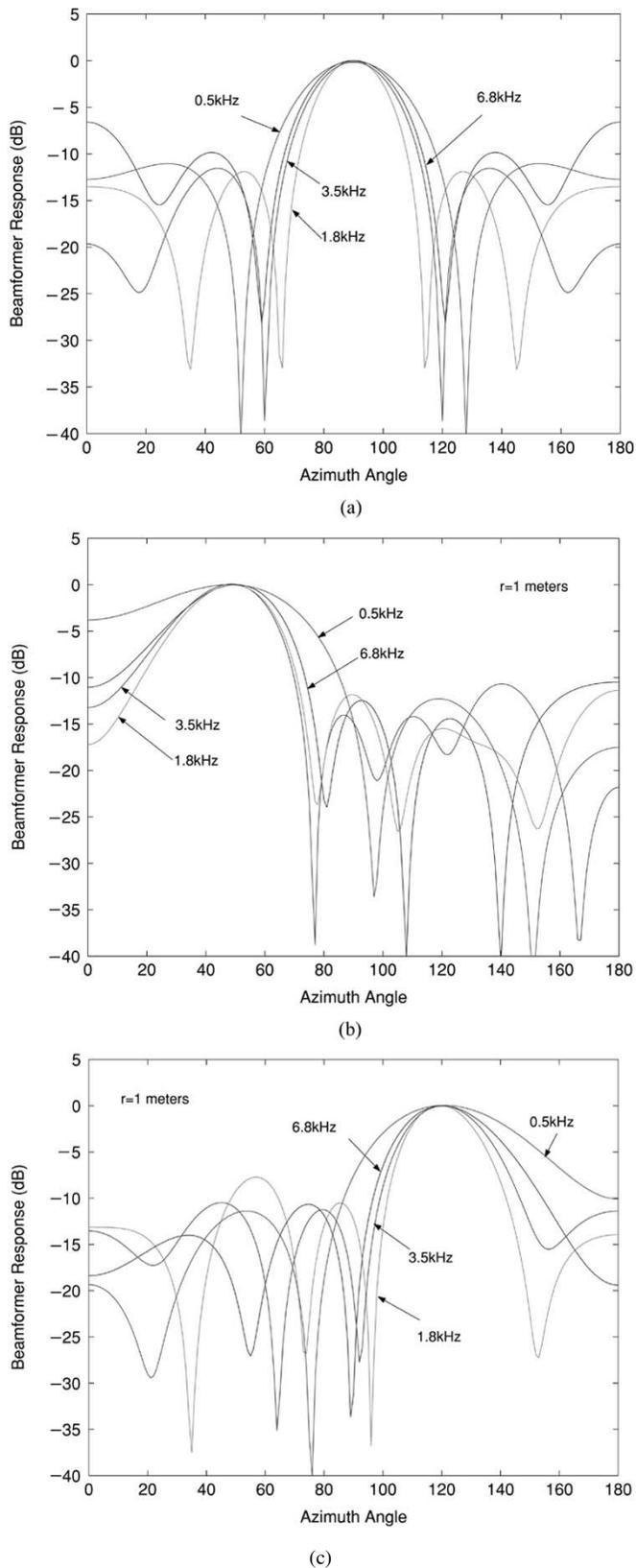
Fig. 6.   Three predesigned DFS beams of the NMA-ANC plotted at the four in-band frequencies 500, 1800, 3500, and 6800 Hz. (a) Beamformer1 focusing at $\mathbf{x}_{S1} = (0.6\,\mathrm{m}, 90°)$, (b) Beamformer2 focusing at $\mathbf{x}_{S2} = (10\,\mathrm{m}, 50°)$, and (c) Beamformer3 focusing at $\mathbf{x}_{S3} = (10\,\mathrm{m}, 120°)$.

third run for Subarray3 and Subarray4. The second run was redundant but it turned out to be very helpful in case there were damaged data in the other two runs. It also helped with the cuing or synchronization of the multiple runs.

The multiple recording runs were not synchronized due to the nonsynchronized manual operation of the play and record buttons. The synchronization was carried out after the recording by identifying and aligning the cue frames of the recorded data of the superimposed microphone elements. After the synchronization, the recorded data were down sampled to 16 kHz.

Two experiments were carried out at Carleton University. One was the noise rejection experiment in an anechoic chamber and other was the de-reverberation experiment in a real conference room. When the experiment was performed in the anechoic chamber, the equipment rack was placed outside of the anechoic chamber while the microphone array and the loudspeaker were inside the anechoic chamber. The connection cables run through a small hole on a wall of the chamber and it was covered by sound absorbing forms. The sound source CD was played back at several locations in front of the array, as shown in Fig. 8. A Cartesian coordinate system was defined such that the array center was at the origin and the elements laid along the x axis. The sound sources were on the x-y plane with their Cartesian coordinates being (0, 0.60, 0), (0.55, 0.89, 0) and (−0.59, 0.80, 0), respectively. Their corresponding spherical coordinates were (0.6, 90°, 90°), (1.05, 58.4°, 90°), and (0.99, 126.4°, 90°), respectively. All distances were in meters.

Second, the recordings were performed in a small conference room with strong reverberation. The size of the room was $5.0 \times 3.8 \times 3.5$ m. The room was constructed with double plaster board walls, cement floor with linoleum tiles, acoustic tile drop ceiling below a corrugated steel roof, and a double wooden door. There were a square table and six padded chairs in the middle of the room, as shown in Fig. 9. The equipment rack stood in a corner of the room beside the door. The microphone array was placed on a desk in another corner of the room. The phase center of the array was located at 1 m away from the floor and the two walls. The angle between the array axis and the walls was $\beta = 45°$. The sound source was located 0.6 m away from the array center on the y axis. The background noise level in the room was low compared to a typical office environment. Consequently, these recordings were suitable for examination of the beamformer's de-reverberation performance.

IV. DATA ANALYSIS AND RESULTS

A. Noise Reduction Performance

The recordings made in the anechoic chamber were used for evaluation of the interference rejection performance. The desired signal $S1$ was a female speech located at (0, 0.60, 0) in the Cartesian coordinate system. Two interfering signals $S2$ and $S3$ were, respectively, a mix of male speech segments and female speech segments at (0.55, 0.89, 0), and a music signal located at (−0.59, 0.80, 0). The interfering female speech and the desired female speech were generated from different talkers. Each signal was 18 s in length and the sampling rate was 16 kHz.

The speech signals were some English sentences as follows.

- S1: "Welcome to the Code Composer Studio multimedia tutorial. This tutorial has been created to show developers
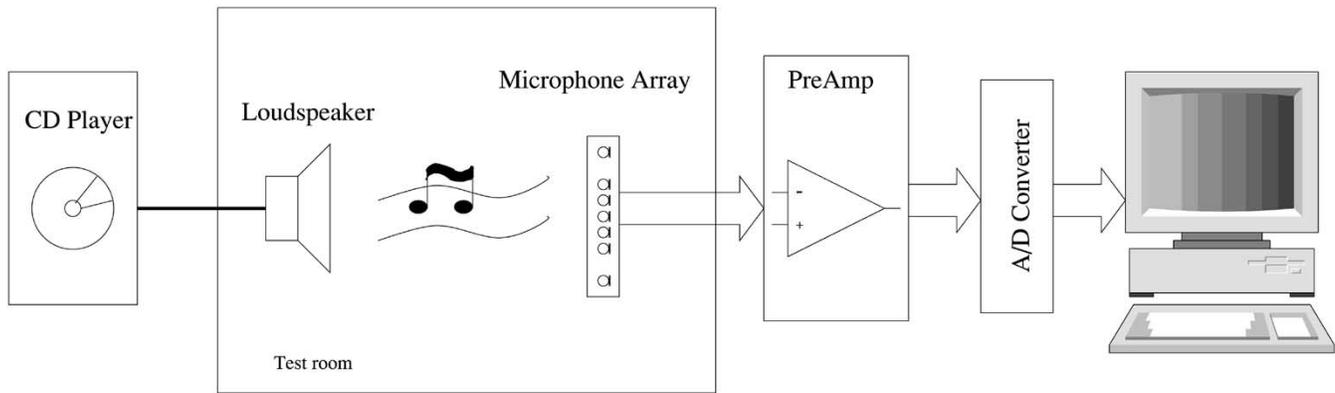
Fig. 7. Multichannel digital recording system for microphone array experiments. A MIDIMAN digital recording card is installed in the personal computer.

TABLE I
EXPERIMENTAL APPARATUS

|	Manufacturer	Model	Description
CD Player	TASCAM	CD-150	Compact Disk Player
Loudspeaker	TANNOY	REVEAL	Omni-directional
Microphone	AUDIO-TECHNICA	AT803b or AT831	Miniature Condenser
Pre-Amplifier	ALLEN & HEATH	MixWizard WZ12:2DX	12-ch Pre-Amp
A/D Converter	MIDIMAN	Delta 1010	8-ch Digital Recording System
Personal Computer	Dell	Pentium 300MHz	Microsoft WIN98, 128MB RAM
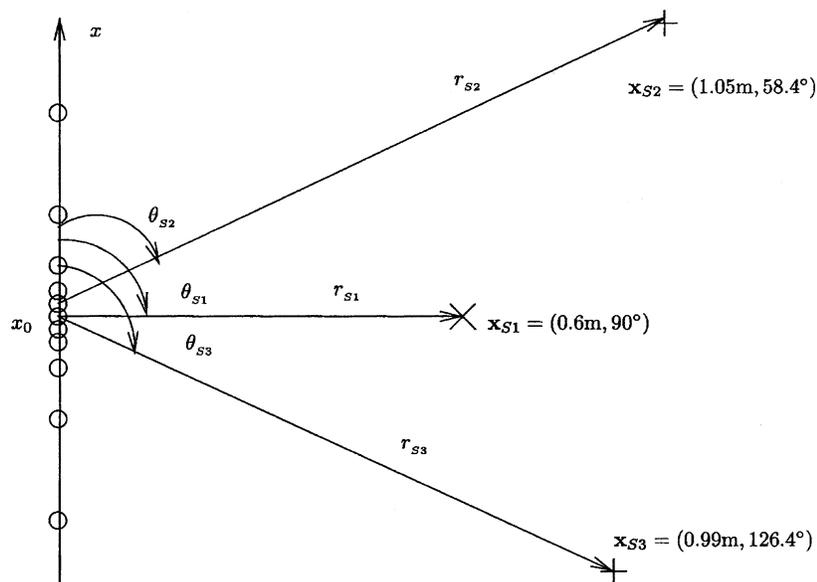
Fig. 8. Signal source locations of the noise rejection experiment in the anechoic chamber.

how to utilize a few of the Code Composer Studio's key features. It is complementary to the tutorial found both in the on-line help and as a pdf file located on the program CD-ROM (female 1)."

- S2: "Incoming file transfer (male)"; "Incoming chat request (male)"; "This is the speaker and sound card test for the Intel configuration wizard. As you listen to this

recording, adjust the volume to a comfortable level using the configuration wizard slider bar (female 2)."

The three signals were received by the array separately. They were scaled to have the same power at the array's phase center then mixed together. Uncorrelated white noises were also added to each element with -20 dB power with respect to the signals. The power spectrum densities (PSD) of the input signals are
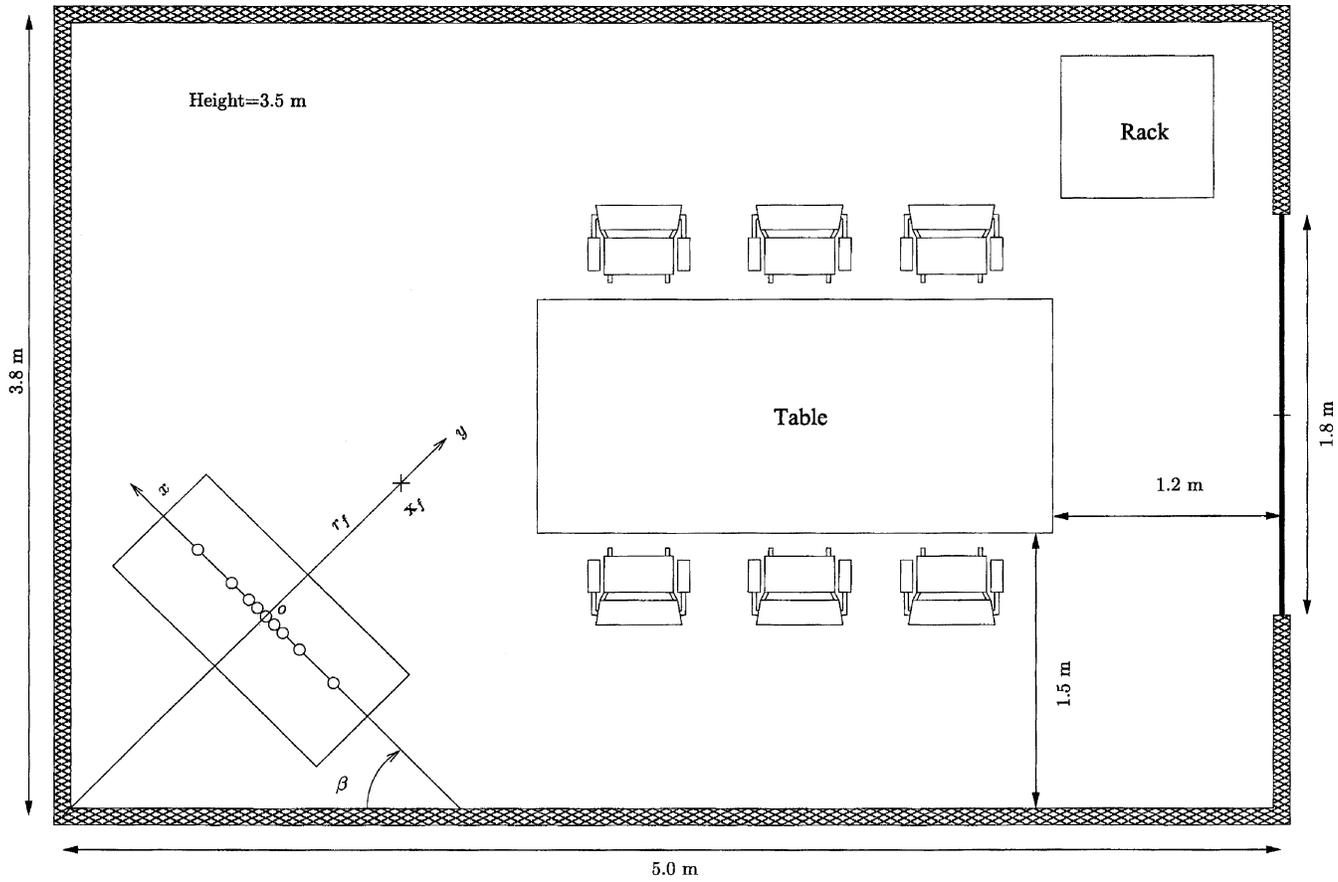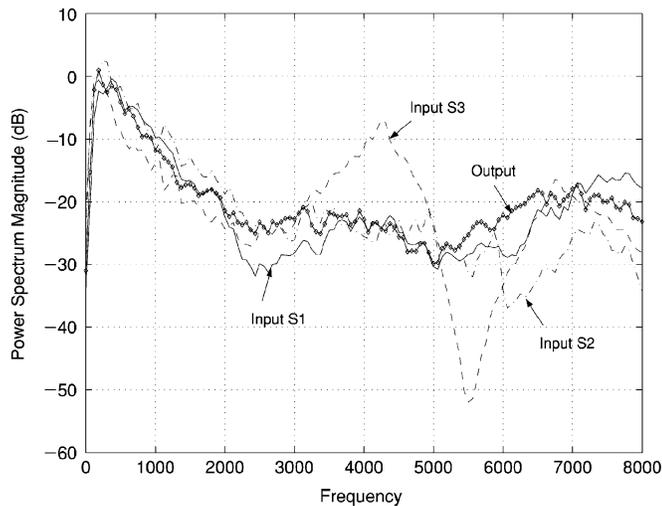
Fig. 9.   Experiment setups of the conference room.

Fig. 10.   Power spectral density of the three audio inputs and the array output, where $S1$ is the desired signal and $S2$ and $S3$ are the interfering signals.

shown in Fig. 10. The two speech signals ($S1$ and $S2$) had energies concentrated in the low frequency band, while the music signal ($S3$) had high energies spread in the lowest subband and the highest subband. The input SINR's of the subbanded signals were not the same in every subband.

The three input signal sources were fed into the adaptive NMA-ANC simultaneously. The adaptation of each subband ANC was controlled by a simple [voice activity detector (VAD)]. The VAD estimated the power of the desired signal $S1$ in every frame of several hundred samples and selected a threshold according to the subband signal energy. If the estimated power was above the threshold, then the VAD was on and the adaptation of the ANC was stopped. If the VAD was off, then the ANC would adapt to the signal inputs. The Normalized Least Mean Square (NLMS) algorithm was used for the ANC with the step size $\mu = 0.02$. The adaptation was converged within 10 s of the signal input. This duration corresponded to 16×10^4 samples in Subarray1 (the highest frequency band), or 8×10^4 samples in Subarray2, or 4×10^4 samples in Subarray3, or 2×10^4 samples in Subarray4 (the lowest frequency band). The output power and SINR were computed using the data after the convergence. The results were listed in Table II. The compound NMA-ANC achieved a SINR of 23.9 dB at the output, and a NR of about 26.4 dB.

The PSD of the NMA-ANC output is shown in Fig. 10. The output PSD was very close to that of the desired input $S1$. The difference between the output PSD and the desired signal PSD was the contribution of the interference power. It was very small indicating that high noise reduction factor was achieved. The experimental results verified that the design of the robust NMA-ANC beamformer was successful.

B. Dereverberation Performance

The recordings made in the conference room were used for the de-reverberation performance evaluation. The signal source

TABLE II
SINR OF THE NMA-ANC AND ITS SUBBANDS EVALUATED
BY THE NOISE REJECTION EXPERIMENTAL DATA

	BF Input	ANC Input	ANC Output
Subband $B_1 = [3.6, 7.2]$ kHz	−2.4 dB	10.1 dB	28.3 dB
Subband $B_2 = [1.8, 3.6]$ kHz	−7.6 dB	5.9 dB	25.8 dB
Subband $B_3 = [0.9, 1.8]$ kHz	−6.1 dB	9.3 dB	24.6 dB
Subband $B_4 = [0.1, 0.9]$ kHz	−0.4 dB	9.2 dB	24.4 dB
Compound NMA-ANC	−1.8 dB	9.1 dB	23.9 dB

was the same female speech used in the noise rejection case. The recorded reverberant signals were processed by the NMA-ANC with the ANC switched off at all times by the VAD.

To evaluate the SINR at the beamformers' output, the input signal had to be decomposed into the direct path and the reflected paths. This can be easily performed in computer simulations. However, separating the direct path signal from its reflected paths was difficult for real room recordings. We took an alternative of the clean signal source recorded in the anechoic chamber as the direct path signal. This signal was filtered by the NMA-ANC and the output was obtained as the desired signal output. The signal recorded in the conference room was calibrated with respect to the clean signal, then processed by the NMA-ANC, and the output was the total output containing the desired signal and the suppressed reverberant signals. The output interference power was estimated as the difference between the total output power and the desired output power. The SINR was measured based on these signal powers. As the result, the input SINR was 6.1 dB. The output SINR was 9.3 dB. The NMA-ANC obtained a de-reverberation gain of 3.2 dB.

Fig. 11 plots the waveforms of the direct path signal, the recorded reverberant signal, and the output signal processed by the NMA-ANC beamformer. Only the first 1-s waveforms of the signals were plotted to show the details of the reverberation. The clean signal recorded in the anechoic chamber had very low noises in the nonspeech frames, as shown in Fig. 11(a). The reverberant signal recorded in the conference room was shown in Fig. 11(b). The nonspeech frames were covered by the reflected speech signals, except for the beginning of the signal. The waveforms of the speech frames were also different from the clean speech. Fig. 11(c) showed the output signal processed by the NMA-ANC beamformer. The reverberation was partially suppressed and the speech waveform was restored close to the clean signal. The benefit of the 3.2-dB de-reverberation gain was evident.

C. PAMS Test

The PAMS is an objective test of listening effort (LE) and listening quality (LQ) specified by ITU-T Recommendation P.800. The mean opinion score (MOS) calculated by PAMS is typically within one half of a MOS of that determined by a well controlled subjective test in a laboratory ([14], pp. 17–23). The
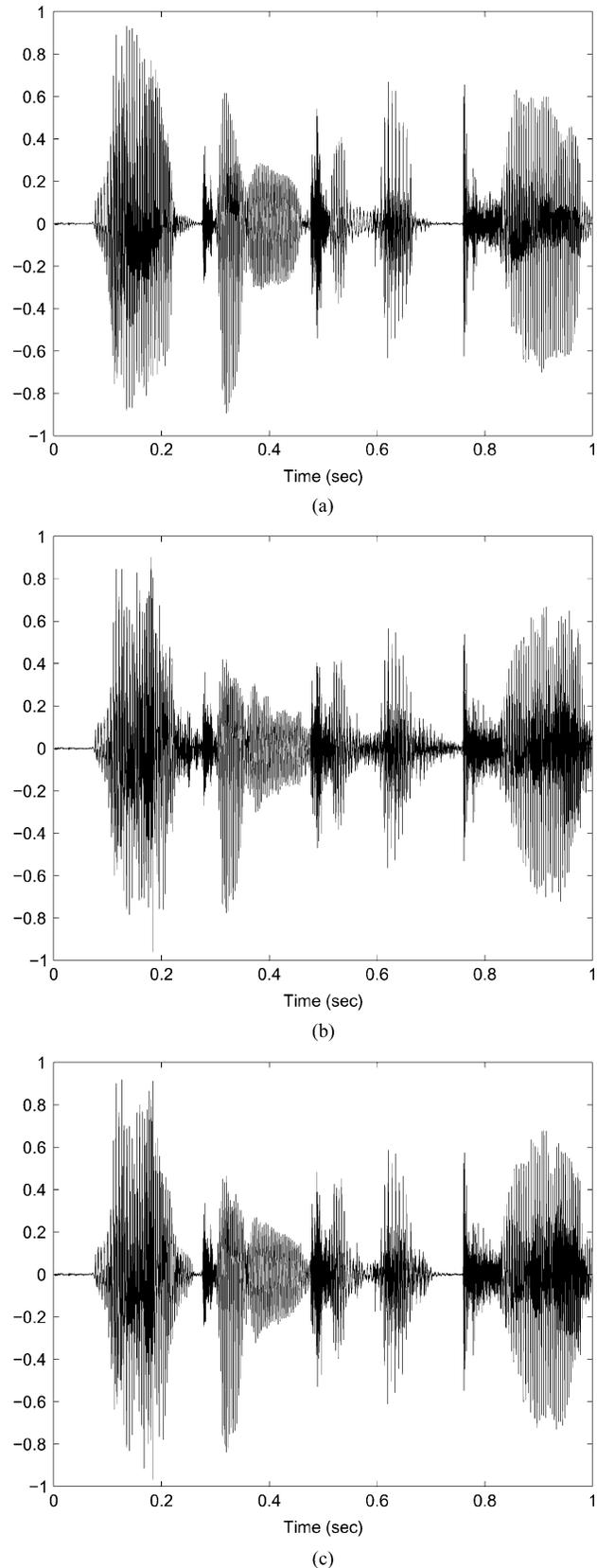

Fig. 11. Speech waveforms of the de-reverberation experiment. (a) Clean signal source recorded in an anechoic chamber, (b) reverberant signal recorded in a conference room, and (c) output signal after de-reverberation processing.

standard MOS gives a measure of perceptual quality, as listed in Table III.

TABLE III
STANDARD MOS SCORES

	Listening Quality	Listening Effort
5	Excellent	Complete relaxation possible; no effort required
4	Good	Attention necessary; no appreciable effort required
3	Fair	Moderate effort required
2	Poor	Considerable effort required
1	Bad	No meaning understood with any feasible effort

TABLE IV
LISTENING EFFORT (LE) AND LISTENING QUALITY (LQ)
SCORES OBTAINED BY THE PAMS TEST.

	Noise Rejection		De-reverberation	
	LE	LQ	LE	LQ
Array Input	1.8	1.0	2.9	2.4
NMA-ANC Output	4.2	3.5	3.5	3.1

A digital speech level analyzer (DSLA), a COTS product made by Malden Electronics, Ltd., was used to perform the PAMS test for the NMA-ANC. The clean speech signal source was used as the reference input to the DSLA. The input signal at the array's phase center and the output signals of the beam-formers were the test signals fed separately to the DSLA. The resulting LE and LQ scores are listed in Table IV.

The results of the noise rejection experiments shows that the noisy input at the array had a LQ score of only 1.0. The NMA-ANC scheme improved the LQ to 3.5, which corresponded to the NR factor of 26.4 dB. The recorded reverberant input at the array had a $LQ = 2.4$. The output of the NMA-ANC achieved an LQ of 3.1, which corresponded to the de-reverberation gain of 3.2 dB. The LE scores were higher than the corresponding LQ scores. This was common to all experiments, according to [14]. All the results were within expectations. They verified the experimental results of the NMA-ANC.

V. CONCLUSION

This paper has proposed a near-field broadband adaptive beamforming scheme for intelligent computer telephony and teleconferencing applications, namely the NMA-ANC. An 11-element microphone array was designed for the [50, 7000] Hz band using 16-kHz sampling rate. The array consisted of four nested subarrays each having five elements to cover an octave subband. Each subband array employed several near-field DFS beamformers and an ANC. With the assistance of VAD, the proposed NMA-ANC is able to adaptively reject multiple interfering signals, attenuate the reverberant noises and avoid the desired signal cancellation.

Experiments of noise reject and de-reverberation have been performed in an anechoic chamber and a small conference room, respectively. A multichannel digital recording system has been developed using COTS products. The results of the experiments have shown that the NMA-ANC achieves a noise reduction factor of 26.4 dB and a de-reverberation gain of 3.2 dB. The PAMS test has also been carried out using a COTS digital speech level analyzer. The results have shown that the NMA-ANC can improve the listening quality from 1.0 to 3.5 in the interference rejection scenario and from 2.4 to 3.1 in the reverberant environment.

REFERENCES

[1] E. M. Petriu, N. D. Georganas, D. C. Petriu, D. Makrakis, and V. Z. Groza, "Sensor-based information appliances," *IEEE Instrum. Meas. Mag.*, vol. 3, pp. 31–35, Dec. 2000.
[2] R. C. Luo, C.-C. Yih, and K. L. Su, "Multisensor fusion and integration: Approaches, applications, and future research directions," *IEEE Sensors J.*, vol. 2, pp. 107–119, Apr. 2002.
[3] D. Zotkin, R. Duraiswami, V. Philomin, and L. S. Davis, "Smart video-conferencing," in *IEEE Int. Conf. Multimedia Expo.*, vol. 3, 2000, pp. 1597–1600.
[4] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," in *IEEE ASSP Mag.*, Apr. 1988, pp. 4–24.
[5] J. G. Ryan, "Criterion for the minimum source distance at which plane-wave beamforming can be applied," *J. Acoust. Soc. Amer.*, vol. 104, no. 1, pp. 595–598, July 1998.
[6] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and cures," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 469–478, May 1982.
[7] J. Huang, N. Ohnishi, and N. Sugie, "Sound localization in reverberant environment based on the model of the precedence effect," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 842–846, Aug. 1997.
[8] Y. R. Zheng, R. A. Goubran, and M. El-Tanany, "A broadband adaptive beamformer using nested arrays and multirate techniques," in *Proc. IEEE DSP Workshop*, Hill County, TX, Oct. 2000.
[9] ——, "On constraint design and implementation for broadband adaptive array beamforming," in *IEEE ICASSP*, Orlando, FL, May 2002.
[10] ——, "A nested sensor array focusing on near field targets," in *Proc. IEEE Sensors Conf.*, Toronto, ON, Canada, Oct. 2003.
[11] Y. R. Zheng, "Spatial-temporal subband beamforming for near field adaptive array processing," Ph.D dissertation, Carleton Univ., Ottawa, ON, Canada, 2002.
[12] P. Mermelstein, "G.722, a new CCITT coding standard for digital transmission of wideband audio signals," *IEEE Commun. Mag.*, vol. 26, no. 1, pp. 8–15, Jan. 1988.
[13] Product Information and User Guide [Online]. Available: http://www.midiman.com
[14] *Digital Speech Level Analyzer User Guide*, Malden Electronics, Ltd., Revision 3.3, 1997–2000.

Yahong R. Zheng received the B.S. degree in electrical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 1987, the M.S. degree in electrical engineering from Tsinghua University, Beijing, China, in 1989, the Ph.D. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 2002.

From 1989 to 1994, she was a Senior Member of the Scientific Staff with the Peony Electronic Group, Beijing, China. From 1994 to 1997, she held positions with GPS Solutions at Sagem Australasia Pty., Ltd., Sydney, Australia, and Polytronics Pty., Ltd., Toronto, ON, Canada. Currently, she is an NSERC Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Missouri, Columbia. Her research interests include digital signal processing algorithms, array signal processing, and channel estimation and modeling for wireless communications. She has published a number of papers in international journals and conferences and holds two patents in China.

Rafik A. Goubran (M'87) received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University, Cairo, Egypt, in 1978 and 1981, respectively. He received the Ph.D. degree in electrical engineering from Carleton University, Ottawa, ON, Canada, in 1986.

In January 1987, he joined the Department of Systems and Computer Engineering, Carleton University, where he is now Professor and Chair. He was involved in several research projects with industry and government organizations, including Nortel Networks, Mitel Networks, Bell Canada, NEC Corporation, the Department of National Defense (DND), and the National Research Council of Canada (NRC). His research interests include digital signal processing and its applications in audio and biomedical engineering, voice transmission over IP networks (VoIP), noise and echo cancellation, and beamforming using microphone arrays.

He is a member of the association of Professional Engineers of Ontario.

Mohammed El-Tanany received the B.Sc. and M.Sc. degrees in electrical engineering from Cairo University, Giza, Egypt, in 1974 and 1978, respectively, and the Ph.D. degree in electrical engineering from Carleton University, Ottawa, ON, Canada in 1983.

He was with the Advanced Systems division of Miller Communications, Kanata, ON, from 1982 to 1985, with principal involvement in the research and development of digital transmission equipment for mobile satellite types of applications, as well as for VHF airborne high-speed down links. He joined Carleton University in 1985. He is currently a Professor with the Department of Systems and Computer Engineering. His research activities are mainly in the areas of digital wireless communications with emphasis on the transmission subsystems from a hardware design and algorithm design points of view, digital TV and digital audio broadcasting systems, experimental characterization and empirical modeling of wireless communications channels, and digital TV channels in various environments and at various frequency bands.