A SUBBAND BEAMFORMER ON AN ULTRA LOW-POWER MINIATURE DSP PLATFORM

Edward Chau, Hamid Sheikhzadeh, Robert Brennan, Todd Schneider

Dspfactory Ltd., 80 King Street South, Suite 206, Waterloo Ontario, Canada N2J 1P5
email address:{echau,hsheikh,rbrennan,tschneid}@dspfactory.com

ABSTRACT

This paper presents the design and implementation of a subband cardioid beamformer on an ultra low-power miniature DSP platform, using a 2-microphone endfire array. The subband beamformer extends the classical time-domain, narrow-band algorithm to a frequency-domain, broadband implementation, so it is suitable for general speech and audio applications. An oversampled, weighted overlap-add filterbank is used to allow wide gain and phase adjustments for low power, low group delay requirements. A subband IIR filter is proposed to overcome the non-zero bandwidth of the frequency bands, and to introduce a nearly linear phase adjustment across the bands. The subband implementation allows the flexibility of integrating the beamformer with additional algorithms at different frequency ranges. The beamformer has been implemented in real-time on Dspfactory's Toccata platform, which has been specifically designed for ultra low-power, miniature, head-mounted audio devices. At 1.25 Volts with a 5 MIPS DSP core, the Toccata consumes only about 800 micro Watts without microphones and receivers.

1. INTRODUCTION

Beamforming and directional processing have been popular approaches in the reduction of environmental noises in hearing instruments and other applications [1, 2, 3, 4]. Particularly for hearing aids and mobile headset devices, there are the general constraints of small physical size and limited processing power. These constraints prevent the implementation of more advanced algorithms unless a significant amount of additional resource is provided – such as the separately housed and powered microphone array in [1]. The algorithm designer for these low-resource devices must, therefore, weigh the constraints against the complexity of the algorithm in order to implement a successful design.

This paper presents the design and implementation of a 2-microphone endfire subband cardioid beamformer on an ultra low-power, miniature, programmable DSP platform. The focus of this paper is on implementing a working beamformer for miniature hearing instruments and headsets on highly resource-constrained DSP devices, instead of on designing a beamformer that may be theoretically superior but impractical to implement. Taking advantage of several innovations in the design of the Toccata DSP platform [5] and the subband beamforming algorithm, the beamformer is implemented in real-time, running on a 5 MIPS DSP core at 1.25 V. This paper briefly describes the innovations that have made the implementation practical. First of all, the use of an oversampled, weighted overlap-add (WOLA) filterbank [6] allows a wide range of gain and phase adjustments in the frequency domain – an essential operation in the subband beamformer. The WOLA filterbank on the Toccata is specifically designed for low power and low group delay requirements [5, 6]. Moreover, frequency-selective processing can also be easily implemented with the WOLA filterbank, so additional algorithms can be integrated into the beamformer at different frequency ranges. Finally, the extension of the original time-domain beamforming algorithm to the frequency domain, especially with the use of the proposed subband IIR filter (Section 2), overcomes the narrow-band limitation of the original algorithm by essentially turning a broadband beamforming problem into a number of narrow-band time-frequency problems.

For evaluation, the spatial and frequency responses of the subband beamformer are measured acoustically in a recording studio. Section 3 describes the test results. It is found that, allowing for real-world considerations such as microphone noise, minor reverberations in the recording studio, and microphones mismatch, the subband beamformer compares favorably to the theoretical cardioid beamformer.

1.1. Narrow-Band Cardioid Beamformer

The classical narrow-band cardioid beamformer belongs to a well-known class of delay-and-sum beamformers [7], and is characterized by its heart-shaped ("cardioid") beampattern. Let $y(t)$ be the output of the classical narrow-band cardioid beamformer with a two-microphone endfire array. The signal $y(t)$ can be described by

$$y(t) = x_b(t - \tau) - x_f(t) \qquad (1)$$

where $x_b(t)$ and $x_f(t)$ are the signals received from the back and front microphone respectively, and τ is a time-delay applied to $x_b(t)$. The value of τ is found by $\tau = \frac{d}{c}$, where d is the distance between the microphones, and c is the speed of sound. In frequency domain, the delay τ is simply a linear phase shift $\phi = \omega\tau$ and, correspondingly,

$$Y(\omega) = e^{-j\phi} X_b(\omega) - X_f(\omega) \qquad (2)$$

The optimal cardioid beampattern, i.e. maximum beamformer gain of 2 at 0-degree direction-of-arrival (DOA), can be obtained if the value of d is $\frac{1}{4}$ times the wavelength of the signal frequency, f_m. For signals that are at frequencies lower than f_m, the beamformer gain is reduced, even though the beampattern maintains a cardioid shape. For signals that are at frequencies higher than f_m, the shape of the beampattern is distorted due to spatial aliasing. Figure 1 shows the changes in the theoretical beampattern at different signal frequencies (f_m being the frequency for which the narrow-band beamformer is designed).

1.2. Cardioid Beamformer for Broad-Band Signals

A common way of compensating for the reduction in the beamformer gain for frequencies below f_m is to apply a post-filter with
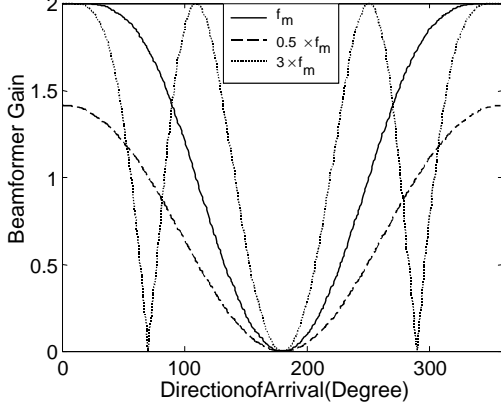
Fig. 1. Beampatterns of a Narrow-Band Cardioid Beamformer

a magnitude response that equals the inverse of the beamformer gain. The post-filter, which is most likely an FIR filter when implemented in the time-domain, will equalize the beamformer gains over the range of frequencies. Since the post-filter cannot compensate for the distortion in the beampatterns due to spatial aliasing for frequencies higher than f_m, the value of f_m should be at least equal to, if not higher than, the highest frequency of interest for a particular application. The major drawback of the post-filter is that it will also amplify the internal microphone noise along with the external signal. Especially for the low-frequency components, where the theoretical post-filter gain is quite high (almost a factor of 10 at 250 Hz), the amplified microphone noise is noticeable in quiet environment. Hence, for the low-frequency components, the actual post-filter gain is generally designed to be lower than theoretically needed. Moreover, the post-filter gain can also be designed so that a maximum post-filtered beamformer gain of 1 is obtained, as opposed to a theoretical maximum gain of 2 (as shown in Figure 1). Attaining a smaller maximum gain requires a lower compensation factor, which in turn reduces the microphone noise amplification.

2. SUBBAND CARDIOID BEAMFORMER

By implementing the cardioid beamformer in the subband frequency domain, a broadband signal is essentially broken down into a number of narrow-band time-frequency signals distributed across the frequency bands. With a few additional signal processing to take into account of the non-zero bandwidth of the frequency bands, the original narrow-band algorithm can be adapted to broadband signals in the subband beamformer. The design of the subband beamformer is described in the following.

As a sampling rate of 16 kHz is used, the highest frequency of interest is 8 kHz. However, due to the non-zero bandwidth of the frequency bands, the center frequency of the highest frequency band is actually used. In a 16-band filterbank implementation, the bandwidth of each frequency band is 500 Hz, with the center frequency of the lowest band at 250 Hz. The center frequency of the highest band is 7750 Hz, which corresponds to a microphone separation distance of $d = \frac{\lambda}{4} = \frac{c}{4 \times 7750} \approx 11$ mm. The amount of spatial aliasing for the small range of frequencies between 7750 and 8000 Hz is considered insignificant. Figure 2 shows the block diagram of the subband cardioid beamformer system using these parameters.

In Figure 2, $X_f(\omega_i, k)$ and $X_b(\omega_i, k)$ are the complex-valued filterbank analysis output of the signals received from, respectively, the front and back microphone, where ω_i is the center frequency of band i, and k is the time index. The beamformer delay τ (from Equation 1) is implemented by a complex-valued multiplication

$$U_b(\omega_i, k) = X_b(\omega_i, k) \times G(\omega_i) \qquad (3)$$

where $G(\omega_i)$ is a complex number, $G(\omega_i) = e^{-j\omega_i\tau}$, with

$$|G(\omega_i)| = 1$$
$$\angle G(\omega_i) = -\omega_i\tau$$

for all ω_i, $\omega_i = 250 + (i-1) \times 500$, $i = 1, 2, 3, ..., 16$. Note that, for each band i, $X_b(\omega_i, k)$ is a close approximation to the original time-domain signal $x_b(t)$ around the frequency ω_i, so $U_b(\omega_i, k)$ is the same signal phase shifted by $-\omega_i\tau$. In the ideal case where the bandwidth of each frequency band is infinitesimal, i.e. $\omega_i = \omega \in \Re, 0 < \omega \leq 8000$, the multiplication with $G(\omega)$ will apply a linear phase shift to $X_b(\omega, k)$, and the beamformer output can then be computed according to Equation 2. However, in reality, due to the non-zero bandwidth of the frequency band, and the fact that the center frequency of each band, ω_i, is used to calculate the phase shift $\angle G(\omega_i)$, there will be small errors in the phase shift within each band. Therefore, a subband IIR filter is proposed to compensate for these errors. The subband IIR filter is an IIR filter applied in the subband frequency domain instead of in the time domain. Let BW denotes the bandwidth of the frequency bands. The subband IIR filter, $H(\Omega)$, is an all-pass filter with an approximately linear phase response for $0 < |\Omega| \leq \frac{BW}{2}$, with

$$|H(\Omega)| = 1$$
$$\angle H(0) = 0$$
$$\angle H(\frac{BW}{2}) = -\frac{BW}{2}\tau$$

In other words, the phase response of the filter changes approximately linearly from 0 at $\Omega = 0$ to $-\frac{BW}{2}\tau$ at $\Omega = \frac{BW}{2}$. Then, for each band i, the filter $H(\Omega)$ is applied to the time series $U_b(k)$ (the constant ω_i is dropped for simplicity). Note that the real and imaginary parts of $U_b(k)$ are treated as two separate but related series, so $H(\Omega)$ is a real-valued filter. $U_b(k)$ (with separate real and imaginary series) is basically a phase-shifted approximation of $x_b(t)$ around frequency ω_i, so the spectra of $U_b(k)$ have no signal information for $|\Omega| > \frac{BW}{2}$. Let the real and imaginary series of $U_b(k)$ be $U_{br}(k) = \Re\{U_b(k)\}$ and $U_{bi}(k) = \Im\{U_b(k)\}$, respectively. Then, the spectra of $U_b(k)$ are $U_{br}(\Omega) = \mathcal{F}(U_{br}(k))$ and $U_{bi}(\Omega) = \mathcal{F}(U_{bi}(k))$, where $\mathcal{F}(\cdot)$ denotes the Fourier transform. Note that the relationship between Ω and k is analogous to the relationship between ω and t. It is clear that after applying the filter $H(\Omega)$ to $U_{br}(\Omega)$ and $U_{bi}(\Omega)$, so that

$$V_{br}(\Omega) = H(\Omega)U_{br}(\Omega)$$
$$V_{bi}(\Omega) = H(\Omega)U_{bi}(\Omega),$$

an approximately linear phase shift is introduced to $U_b(k)$. As a result, within each band i, the overall phase shift applied to $V_b(k)$ (with separate real and imaginary series), relative to $X_b(k)$, is the constant phase shift $-\omega_i\tau$ *plus* an additional approximately linear phase adjustment. Thus, in combination with $G(\omega_i)$, the filter $H(\Omega)$ can be used to introduce an approximately linear phase
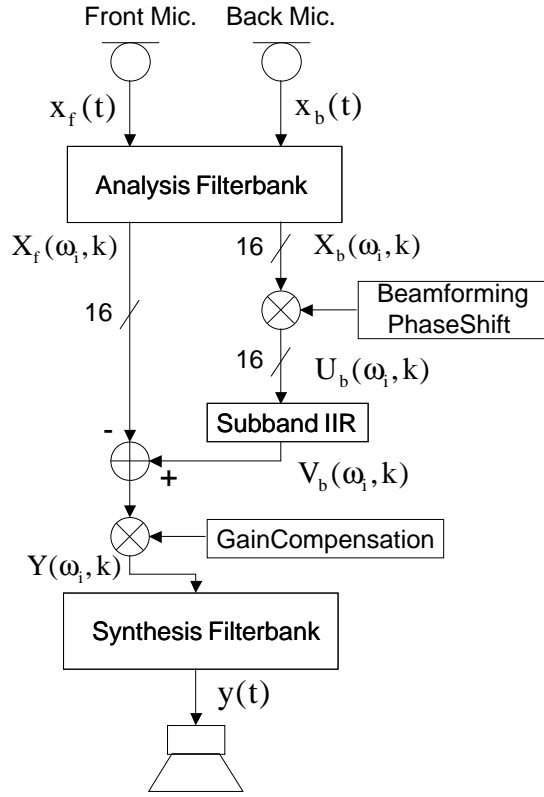
$$\omega_i = 250 + (i-1) \times 500, \quad i = 1,2,3,\cdots,16$$

Fig. 3. Beamforming Phase Shift and Subband IIR Filtering

Fig. 2. Block Diagram of the Subband Beamformer System

real-time on the Toccata DSP platform. The Toccata is based on the ultra low-power DSP design as described in [5]. It consists of a DSP core with a dedicated filterbank coprocessor that allows frequency analysis and synthesis to be done in parallel with any frame-based algorithm processing. The oversampled WOLA filterbank on the Toccata is a vital component that allows the extensive gain and phase adjustments needed for the subband beamformer to be made, whereas a critically-sampled filterbank would not permit these adjustments because of aliasing. Also, the filterbank coprocessor includes a stereo processing mode that allows two input channels to be processed using a single complex FFT-based operation, so the Toccata DSP system is well-suited for implementing a two-microphone subband beamformer.

3. RESULTS

The spatial and frequency responses of the subband beamformer is measured acoustically in a recording studio, with the beamformer running in real-time on the Toccata. A pair of matched hearing aid microphones (Knowles XL-6344-CX) are mounted approximately 11 mm apart in an endfire array. The microphones have a nominal 25 dB "A" weighted noise level at 1 kHz. The mismatch between the two microphones is rated at ± 1 dBV at 200 Hz for sensitivity, and ± 1 degree at 200 Hz for phase. Brief discussions on microphones mismatch and other implementation issues can be found in [8, 9].

To measure the beampatterns, a near-field sound monitor acting as an acoustic source is placed approximately 1 metre away from the microphone array, while the microphone array is rotated azimuthally. A set of 16 band-limited flat-spectrum signals are used as the acoustic test signal, which corresponds to the 16 bands of the filterbank used on the Toccata. The test signals are generated by passing a white-noise signal through appropriate band-pass filters. The output of the beamformer is then recorded onto a PC at 16 kHz sampling rate and 16-bit precision. In order to derive the beampatterns, a reference output is generated from the Toccata with the beamformer turned off. This reference output is simply the output of the front microphone when the array is aimed at zero degree towards the sound source, and can be easily obtained without disturbing the physical configuration of the test equipment set-up. The beampattern at each frequency band is then calculated using the ratio of the beamformer output over the reference output.

Figure 5 shows the measured beampatterns. The solid line indicates the averaged beampattern for frequency bands 2 to 16, while the beampattern for the first band is shown separately with the dashed line. The first frequency band is shown separately because, in order to reduce the microphone noise amplification (Section 1.2), the gain compensation for that band is much lower than theoretically required. As seen in the figure, for the first fre-

shift across all frequency bands. Figure 3 illustrates the phase shift operations for a single frequency band. The dashed lines in the figure indicate that the real and imaginary parts of $U_b(\omega_i, k)$ are treated as two separate real-valued series. Hence, in the figure, $X_b(\omega_i, k)$, $G(\omega_i)$, $U_b(\omega_i, k)$, and $V_b(\omega_i, k)$ are complex-valued, while $U_{br}(\omega_i, k)$, $U_{bi}(\omega_i, k)$, $V_{br}(\omega_i, k)$, $V_{bi}(\omega_i, k)$ and the filter $H(\Omega)$ are real-valued.

For comparison, Figure 4 shows the ideal linear phase shift, the uncompensated and compensated subband phase shifts for a 16-band filterbank implementation. The phase shifts for the first three bands are shown. In Figure 4, the "stair-case" pattern of the uncompensated subband phase shift clearly shows the bandwidth effect of the frequency band. After compensation with a first-order subband IIR filter, the phase error is found to be reduced by at least two orders of magnitude. As the compensated subband phase shift is much closer to the ideal linear phase shift, it is difficult to clearly distinguish between the two in the figure.

After combining $V_b(\omega_i, k)$ with $X_f(\omega_i, k)$ (see Figure 2), the beamformer gain compensation is applied by a scaling factor for each frequency band. The scaling factors are designed so that a theoretical maximum gain of 1 is obtained at 0-degree DOA. While implementing the gain compensation by subband scaling factors has the advantage of providing an appropriate magnitude response without introducing an unnecessary phase response like a time-domain post-filter, there will be some error in the compensation due to the non-zero bandwidth of the bands. However, given a narrow enough bandwidth, the error is negligible.

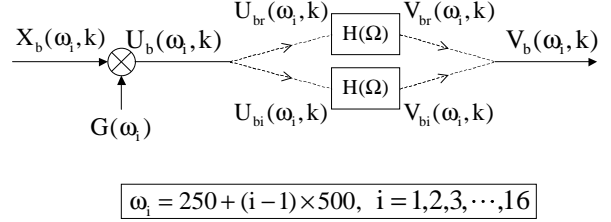Finally, the subband cardioid beamformer is implemented in

quency band (center frequency at 250 Hz), directivity is sacrificed in return for lower microphone noise. For bands 2 to 16, however, an average front-to-back gain of 13 dB is achieved. A maximum front-to-back gain of 20.5 dB is found at band 14 (center frequency at 6750 Hz), while a minimum front-to-back gain of 9 dB is found at band 2 (center frequency at 750 Hz). Taking into account real-world factors such as microphone noise, minor reverberation and microphones mismatch, the measured result does not appear to deviate far from theoretical expectations [8, 9].

4. CONCLUSION

A 2-microphone endfire subband cardioid beamformer has been designed and implemented for miniature head-mounted audio applications. The beamformer is a broadband and frequency-domain extension of the classical time-domain cardioid beamformer. With a few innovations in the DSP platform and algorithm design, the beamformer has been implemented in real-time running on a 5 MIPS DSP core at 1.25 V, consuming about 800 μW without the microphones and receivers. Even with very limited computing resource, it is found that the subband beamformer is comparable to the theoretical cardioid beamformer. For further work, the low-frequency performance of the beamformer can be improved by designing a post-filter (such as a subband Wiener filter) to reduce both microphone and external noises at the lower frequency bands. More sophisticated enhancement to the current algorithm may also be investigated.

5. REFERENCES

[1] B Widrow, "A microphone array for hearing aids," *Proc. IEEE Symp. Adaptive Systems for Signal Processing, Communications and Control*, 2000.

[2] E D McKinney and V E DeBrunner, "A two-microphone adaptive broadband array for hearing aids," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 1996.

[3] G H Saunders and J M Kates, "Speech intelligibility enhancement using hearing-aid array processing," *J. Acoust. Soc. Am.*, vol. 102, no. 3, pp. 1827–1837, 1997.

[4] I A McCowan, C Marro, and L Mauuary, "Robust speech recognition using near-field superdirective beamforming with post-filtering," *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, 2000.

[5] T Schneider, R Brennan, P Balsiger, and A Heubi, "An ultra low-power programmable DSP system for hearing aids and other audio applications," *Proc. Int. Conf. Signal Processing Applications & Technology*, 1999.

[6] R Brennan and T Schneider, "A flexible filterbank structure for extensive signal manipulations in digital hearing aids," *Proc. IEEE Int. Symp. Circuits and Systems*, pp. 569–572, 1998.

[7] D H Johnson and D E Dudgeon, *Array Signal Processing: Concepts and Techniques*, Simon & Schuster Trade, 1992.

[8] B Csermak, "A primer on a dual microphone directional system," *The Hearing Review*, vol. 7, no. 1, pp. 56–60, 2000.

[9] S C Thompson, "Directional patterns obtained from two or three microphones," *Technical Report, Knowles Electronics*, 2000.
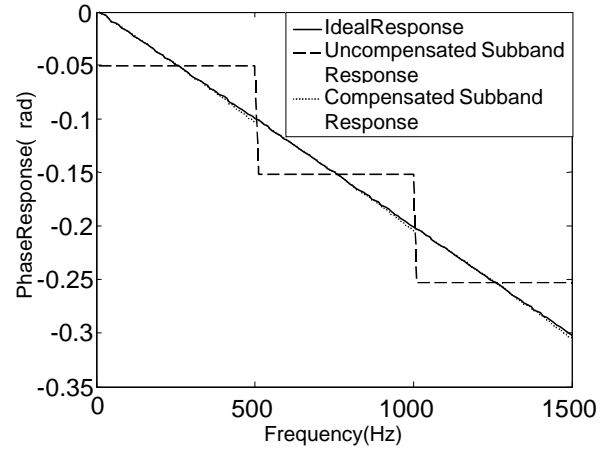
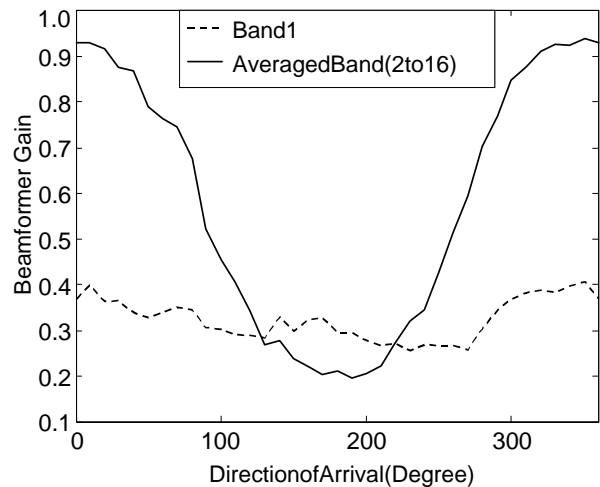**Fig. 4**. Comparison of the Ideal and Subband (both Compensated and Uncompensated) Phase Responses

Fig. 5. Acoustically Measured Beampatterns