DSP beamforming and talker tracking with a linear microphone array

Harvey F. Silverman

Citation: The Journal of the Acoustical Society of America **87**, S2 (1990); doi: 10.1121/1.2028157
View online: https://doi.org/10.1121/1.2028157
View Table of Contents: https://asa.scitation.org/toc/jas/87/S1
Published by the Acoustical Society of America

ARTICLES YOU MAY BE INTERESTED IN

A frequency-domain beamforming algorithm for wideband, coherent signal processing
The Journal of the Acoustical Society of America **76**, 1132 (1984); https://doi.org/10.1121/1.391405

Machine learning in acoustics: Theory and applications
The Journal of the Acoustical Society of America **146**, 3590 (2019); https://doi.org/10.1121/1.5133944

Broadband beamforming for a microphone array
The Journal of the Acoustical Society of America **96**, 845 (1994); https://doi.org/10.1121/1.410323

On the design and implementation of linear differential microphone arrays
The Journal of the Acoustical Society of America **136**, 3097 (2014); https://doi.org/10.1121/1.4898429

Beamforming for a circular microphone array mounted on spherically shaped objects
The Journal of the Acoustical Society of America **109**, 185 (2001); https://doi.org/10.1121/1.1329616

discrete Fourier transform (DFT), discrete cosine transform (DCT), and discrete Hartley transform (DHT) are used for signals with different eigenvalue spread. Real-time implementations of this class of canceler with digital signal processors (DSP) will also be discussed.

8:40

A2. A microphone array for echo cancellation in audioconference rooms. Y. Grenier and M. Xu (Telecom Paris, Dept. Signal, 46 rue Barrault, 75634 Paris Cedex 13, France)

This paper describes an adaptive acoustic antenna (or microphone array) designed for sound recording in audioconferencing rooms. This microphone array achieves three goals simultaneously: the sound from the useful sources (speakers) is recorded without distortion, several jammers emitting known signals can be cancelled (this is the case for the sound received from the other room and emitted by the loudspeakers), and finally, the ambient noise and the room reverberation are reduced. This last effect enforces the "presence effect" in the recordings. In the present approach, echo cancellation is seen as the problem of rejecting a jammer, and it is done by signal combination rather than by subtraction. To achieve these goals, the microphone array uses adaptive beamforming techniques. The beamformer is a compromise between data-independent or fixed beamformers and data-dependent or adaptive ones. This makes it robust with respect to localization errors: the beamformer uses an identified localization of the jammers but weighs this information by some knowledge of the signal-to-noise ratio. The localization of the main jammer (the echo coming from the loudspeaker) requres the identification of a long impulse response (the acoustic channel between the loudspeaker and each microphone). A model to identify this response in the time-frequency domain is developed. This model involves a "double convolution" and its adaptive identification has some analogies with subband adaptive methods. The estimated parameters of this model are used to compute the optimal weights for the beamformer every 8 ms. The paper will present simulations of these techniques based upon a first prototype that consists of an array of 15 microphones on a semicircle with diameter 1 m. The performances of this adaptive microphone array are characterized in terms of echo rejection, convergence rate, and level of dereverberation.

9:05

A3. Acoustic echo cancellation in subbands. Walter Kellermann (Acoustics Res. Dept., AT&T Bell Labs., Murray Hill, NJ 07974)

Acoustic echo cancellation for teleconferencing or hands-free telephony constitutes a challenging task for today's digital signal processing techniques. Compared to the well-established line echo cancellation in telephone networks, the problem takes on a considerably larger size: The impulse response of the echo path to be compensated is longer and may vary more rapidly in time. Moreover, larger bandwidths are often desirable. Therefore, acoustic echo cancellation requires algorithms that adapt far more coeffecients and converge faster than those commonly used in line echo cancelers. On the other hand, direct implementation of sufficiently advanced adaptive filtering algorithms is usually prohibited because of their computational complexity and inherent numerical difficulties. As a way out, the subband approach realizes a "divide and conquer" strategy and promises a reduction of computational complexity and favorable circumstances for fast convergence at the cost of some extra delay. After comparing several possible subband structures, the frequency subband structure is discussed in some detail. For adaptation using LMS-type algorithms, reasons for the improved convergence behavior compared to a fullband implementation are explained. Design examples are presented to illustrate the effectiveness of the method.

9:30

A4. DSP beamforming and talker tracking with a linear microphone array. Harvey F. Silverman (Lab. for Eng. Man/Machine Systems (LEMS), Div. of Eng., Box D, Brown Univ., Providence, RI 02912)

One of the problems for all speech input is the necessity for the talker to be encumbered by a head-mounted, hand-held, or fixed position microphone. An intelligent, electronically aimed undirectional microphone would overcome this problem. Array techniques hold the best promise to bring such a system to practicality, although the acoustic problems are manifold. High-speed digital signal processing (DSP) chips have made it possible to attack the essential problems of forming the beam, aiming the microphone, and tracking a unique talker. A useful beamforming algorithm and a two-step talker-tracking algorithm are introduced. In the latter, step 1 is a rather conventional filtered cross-correlation method; the delay between some pair of microphones is determined to high accuracy using interpolation on the sampled data. Then, using the fact that the delays for a point source should fit a hyperbola, a best hyperbolic fit is obtained using nonlinear optimization. Results indicate that this method works reliably for signal-to-noise ratios of less than 10 dB. [This work principally supported by NSF Grant No. MIP-8809742 and DARPA/NSF Grant No. IRI-8901882.]