Micropower Mixed-Signal Acoustic Localizer

Milutin Stanacevic and Gert Cauwenberghs
Department of Electrical and Computer Engineering, Johns Hopkins University
3400 N. Charles St. Baltimore, MD 21218, USA
{miki,gert}@jhu.edu

Abstract

A micropower VLSI system-on-chip is presented for 3-D localization of a broadband acoustic source. The direction of the source is estimated by relating spatial and temporal differentials of acoustic signals acquired over a planar geometry of four microphones. Two stages of mixed-signal least-squares adaptation perform common-mode rejection for increased sensitivity in the analog differentials, and directly produce digital estimates of the direction cosines. The $3\,\mathrm{mm}\times 3\,\mathrm{mm}$ *chip in 0.5* μ*m CMOS technology quantizes signal delays with 2* μ*s resolution at 2 kHz sampling rate and at 32* μ*W power dissipation.*

1. Introduction

Battery-operated integrated sensor systems call for low-power sensing, signal conditioning and signal processing at the sensor level for extraction of a limited number of essential features specific to the task, such as the direction of a source of interest. There have been a number of VLSI systems for acoustic direction finding reported in the literature [1, 2, 3]. Trends towards miniaturization of acoustic sensor arrays pose a challenge in attaining high precision of localization due to shrinking aperture.

Gradient flow [5] is a signal conditioning technique for source localization designed for arrays of very small aperture, *i.e.*, of dimensions significantly smaller than the shortest wavelength in the sources. Consider a traveling acoustic wave impinging on an array of four microphones, in the configuration of Figure 1. The 3-D direction cosines of the traveling wave $\mathbf{u}$ are implied by propagation delays τ_1 and τ_2 in the source along directions p and q in the sensor plane. Direct measurement of these delays is problematic as they require sampling in excess of the bandwidth of the signal, increasing noise floor and power requirements. However, indirect estimates of the delays are obtained, to first order, by relating spatial and temporal derivatives of the acoustic field [5]:

$$\begin{aligned} \xi_{10}(t) &\approx \tau_1 \dot{\xi}_{00}(t) \\ \xi_{01}(t) &\approx \tau_2 \dot{\xi}_{00}(t) \end{aligned} \tag{1}$$

where ξ_{10} and ξ_{01} represent spatial gradients in p and q directions around the origin ($p = q = 0$), ξ_{00} the spatial common mode, and $\dot{\xi}_{00}$ its time derivative. From least-square estimates of the time delays in (1), we directly obtain bearing estimates of azimuth angle θ and elevation angle ϕ of $\mathbf{u}$ relative to p and q as

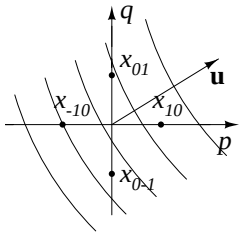

Figure 1. Configuration of sensors for spatial gradient estimation.

$$\begin{aligned} \tau_1 &= \frac{1}{c}|\mathbf{r}_1|\ \cos\theta \sin\phi \\ \tau_2 &= \frac{1}{c}|\mathbf{r}_2|\ \sin\theta \sin\phi, \end{aligned} \tag{2}$$

where $\mathbf{r}_1$ and $\mathbf{r}_2$ are unit orthogonal vectors along p and q directions.

2. System overview

The system block diagram implementing gradient flow for bearing estimation is shown in Figure 2. Spatial gradients are approximated by evaluating finite differences over the four sensors on the planar grid shown in Figure 1. Acurate bearing estimation assumes accurate sensing of the gradients. Differential amplification is performed in the analog domain, as a low-power alternative to high-resolution analog-to-digital conversion and subsequent digital signal processing. Two stages of mixed-signal adaptation compensate for common mode errors in the differential amplification, and produce digital estimates of delays τ_1 and τ_2 from the spatial and temporal differentials.

2.1. Spatial and temporal derivative sensing

Estimates of ξ_{00}, ξ_{10} and ξ_{01} are obtained from the sensor observations $x_{-1,0}$, $x_{1,0}$, $x_{0,-1}$ and $x_{0,1}$ as:

$$\begin{aligned} \xi_{00} &\approx \tfrac{1}{4}\big(x_{-1,0} + x_{1,0} + x_{0,-1} + x_{0,1}\big) \\ \xi_{10} &\approx \tfrac{1}{2}\big(x_{1,0} - x_{-1,0}\big) \\ \xi_{01} &\approx \tfrac{1}{2}\big(x_{0,1} - x_{0,-1}\big) \end{aligned} \tag{3}$$

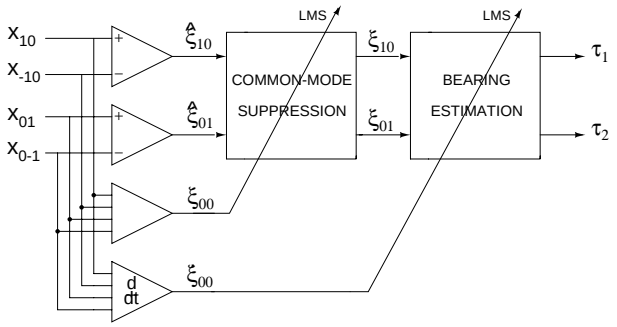

Figure 2. System diagram

Computation of the gradients is implemented using sampled-data switched-capacitor (SC) circuits. The advantage of this realization is application of correlated-double sampling (CDS) to significantly reduce common-mode offsets and $1/f$ noise [4]. A cascoded inverter is used as high-gain amplifier in these and subsequent SC circuits, supporting high density of integration, and high energetic efficiency.

The common-mode component is decomposed in differential form $\xi_{00} = \xi_{00}^{+}[n] - \xi_{00}^{-}[n]$ with

$$\begin{aligned}\xi_{00}^{+}[n] &= \frac{1}{8}(x_{10}[n-\frac{1}{2}] + x_{-10}[n-\frac{1}{2}] \\ &\quad + x_{01}[n-\frac{1}{2}] + x_{0-1}[n-\frac{1}{2}]) \qquad (4)\\ \xi_{00}^{-}[n] &= -\frac{1}{8}(x_{10}[n] + x_{-10}[n] + x_{01}[n] + x_{0-1}[n]).\end{aligned}$$

The computation is illustrated in Figure $3a$), and the corresponding switched-capacitor realization is given in Figure $3b$). The clocks ϕ_1 and ϕ_2 are nonoverlapping, and ϕ_{1e} replicates ϕ_1 with its falling edge slightly preceding the falling edge of ϕ_1. The contribution ξ_{00}^{+} to ξ_{00} represents the estimate of the average signal at time instance $nT - \frac{T}{2}$, while the contribution ξ_{00}^{-} represents the inverted estimate at time instance nT. The difference between both contributions signals hence produces an unbiased estimate of ξ_{00} centered at time $nT - \frac{T}{4}$.

An estimate of the temporal derivative signal $\dot{\xi}_{00}$ centered at same time instance $nT - \frac{T}{4}$ is computed differentially in similar manner, by differencing signal averages (or the reference level) at time instances $nT - \frac{T}{2}$ and nT,

$$\begin{aligned}\dot{\xi}_{00}^{+}[n] &= 0\\ \dot{\xi}_{00}^{-}[n] &= -\frac{1}{8}(x_{10}[n] + x_{-10}[n] + x_{01}[n] + x_{0-1}[n])\\ &\quad + x_{10}[n-\frac{1}{2}] + x_{-10}[n-\frac{1}{2}] \qquad (5)\\ &\quad + x_{01}[n-\frac{1}{2}] + x_{0-1}[n-\frac{1}{2}].\end{aligned}$$

The temporal derivative computation and its switched-capacitor realization are illustrated in Figure 4.

The spatial gradients are computed in fully differential mode, to provide increased clock and supply feedthrough

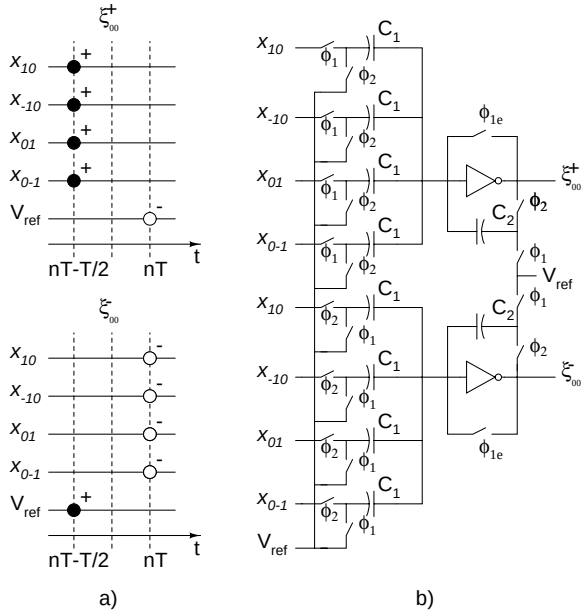

Figure 3. a) Computation of spatial common-mode signal ξ_{00}. b) Switched-capacitor realization.

rejection. The first-order spatial gradient ξ_{10}, likewise centered at time $nT - \frac{T}{4}$, is computed by differencing estimates of ξ_{10} at time instances $nT - \frac{T}{2}$ and nT

$$\begin{aligned}\xi_{10}^{+}[n] &= \frac{1}{4}(x_{10}[n-\frac{1}{2}] - x_{-10}[n])\\ \xi_{10}^{-}[n] &= \frac{1}{4}(x_{-10}[n-\frac{1}{2}] - x_{10}[n]) \qquad (6)\end{aligned}$$

as illustrated in Figure $5a$). The switched-capacitor realization shown in Figure $5b$) includes provisions for common-mode suppression, described next. The first-order spatial gradient in the q direction, ξ_{01}, is computed in identical fashion.

2.2. Common-mode suppression

Prior to bearing estimation, common mode offset correction is performed on the estimated spatial gradients. Common mode errors arise from gain mismatch in the sensors and in the circuits computing the gradients ξ_{10} and ξ_{01}. For ξ_{10} these can be represented to first order as

$$\hat{\xi}_{10} \approx \xi_{10} + \varepsilon_1 \xi_{00}. \qquad (7)$$

To estimate ε_1, a digital sign-sign LMS (SS-LMS) adaptation rule is used. ε_1 is stored as digital value in a 12-bit counter and it is represented in two's complement. The update is performed by incrementing or decrementing the counter based on sign of spatial gradient and average signal [5]

$$\begin{aligned}\varepsilon_1^{+}[n+1] &= \varepsilon_1^{+}[n]\\ &\quad + \mathrm{sgn}(\xi_{10}^{+}[n] - \xi_{10}^{-}[n])\mathrm{sgn}(\xi_{00}^{+}[n] - \xi_{00}^{-}[n])\\ \varepsilon_1^{-}[n+1] &= 2^{12} - 1 - \varepsilon_1^{+}[n+1]. \qquad (8)\end{aligned}$$

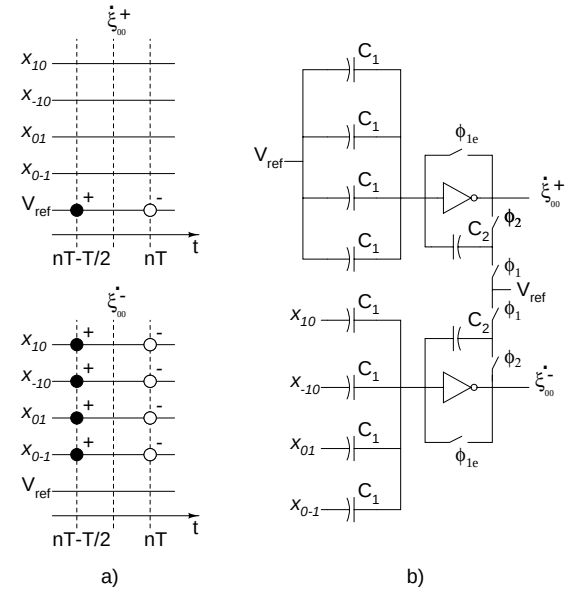

Figure 4. a) Computation of temporal derivative common-mode signal $\dot{\xi}_{00}$. b) Switched-capacitor realization.

The 8 most significant bits are presented to a multiplying D/A capacitor array to construct the LMS error signal. In the case of ξ_{10}, depicted in Figure 5,

$$\begin{aligned} \xi_{10}^{+}[n] &= \hat{\xi}_{10}^{+}[n] - (\varepsilon_1^{+}\xi_{00}^{+}[n] + \varepsilon_1^{-}\xi_{00}^{-}[n]) \\ \xi_{10}^{-}[n] &= \hat{\xi}_{10}^{-}[n] - (\varepsilon_1^{-}\xi_{00}^{+}[n] + \varepsilon_1^{+}\xi_{00}^{-}[n]). \end{aligned} \quad (9)$$

The capacitor array is implemented as an array of 15 columns each containing 16 unit capacitors, controlled by the 4 most significant bits presented in thermometer code, and a column of 15 unit capacitors controlled by 4 least significant bits in thermometer code. The additional 4 bits of stored ε_1 value enable flexibility in the choice of learning rate, that is regulated by the frequency of updating.

The range of common-mode error correction is designed to compensate for at most 5 percent of the average signal in the spatial gradients. Since the minimum size of capacitor array is constrained by the size of unit capacitor, a T-cell is used to attenuate the output swing of multiplying D/A capacitor array. While the transfer characteristic of the T-cell is not linear, it is monotonically increasing. Stray sensitivity in the T-cell does not influence the differential linearity.

2.3. Bearing estimation

For implementation of bearing time-delay estimation, digital SS-LMS differential on-line adaptation is used similar to common-mode error correction. Bearing estimates are represented as 12-bit values in two's complement [5]

$$\begin{aligned} \tau_1^{+}[n+1] &= \tau_1^{+}[n] \\ &\quad +\mathrm{sgn}(e_{10}^{+}[n] - e_{10}^{-}[n])\mathrm{sgn}(\xi_{10}^{+}[n] - \xi_{10}^{-}[n]) \\ \tau_1^{-}[n+1] &= 2^{12} - 1 - \tau_1^{+}[n+1], \end{aligned} \quad (10)$$

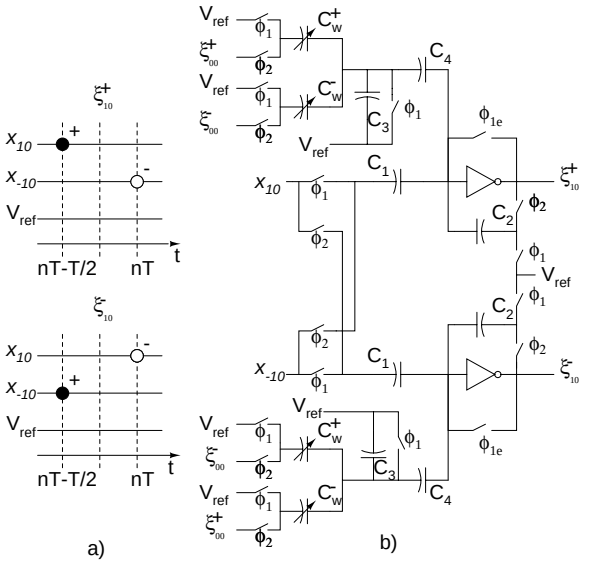

Figure 5. a) Evaluation of first-order spatial gradient along p direction, ξ_{10}. b) Switched-capacitor differential realization.

with the 8 most significant bits used for computation of LMS error signal

$$\begin{aligned} e_{10}^{+}[n] &= \xi_{10}^{+}[n] - (\tau_1^{+}\dot{\xi}_{00}^{+}[n] + \tau_1^{-}\dot{\xi}_{00}^{-}[n]) \\ e_{10}^{-}[n] &= \xi_{10}^{-}[n] - (\tau_1^{-}\dot{\xi}_{00}^{+}[n] + \tau_1^{+}\dot{\xi}_{00}^{-}[n]). \end{aligned} \quad (11)$$

The switched-capacitor realization is shown in Figure 6. Clock ϕ_{2e} represents a replica of ϕ_2 with its falling edge slightly preceding the falling edge of ϕ_2.

The chip outputs bit-serial digital estimates of τ_1 and τ_2, obtained directly from the bearing registers (10) at convergence. The number of update iterations (10) to convergnce depends on the signal and by virtue of the SS-LMS rule does not exceed 2^{12}, determined by the register length.

3. Experimental Results

The VLSI acoustic localizer measures 3×3 mm^2 in 0.5 μm CMOS technology. Figure 7 depicts the micrograph and system floorplan of the chip. The measured transfer characteristic of one multiplying D/A converter is shown in Figure 8. System-level accuracy of bearing estimation was demonstrated using a 200 Hz synthetic sine wave input signal, and a system sampling frequency of 2 kHz, characteristic of acoustic signatures of ground vehicles. The signal was presented to x_{10} and x_{01}, and under variable delays to x_{-10} and x_{0-1}. Digital bearing estimates were observed over observation windows of 2 s, with common-mode compensation activated. The delay was varied from -400 μs to 400 μs, in 2 μs increments. The recorded digital delay estimate τ_1 as a function of actual delay is shown in Figure 9. The power consumption at this operating frequency was measured to be 30 μW for the analog part, and 1.8 μW for the digital part, for a total

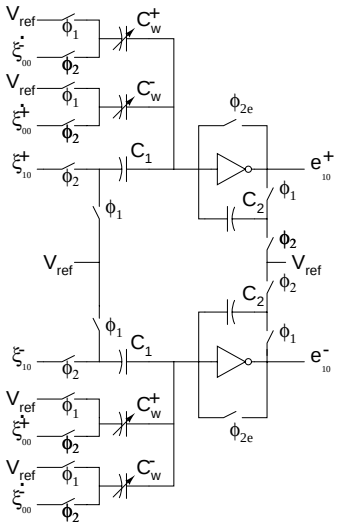

Figure 6. Switched capacitor realization of SS-LMS learning rule for time-delay estimation.

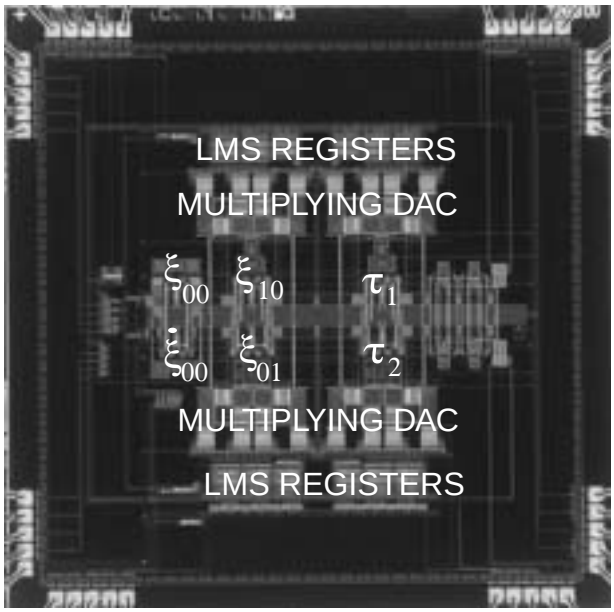

Figure 7. Micrograph of 3×3 mm^2 chip in 0.5 μm CMOS technology.

of 32 μW. Both power and temporal resolution scale with the clock speed, and ns-range time delays can be resolved at sub-mW power levels.

4. Conclusion

A micropower adaptive VLSI chip for gradient-based bearing estimation in miniature acoustic arrays was presented, that combats problems of gain mismatch and $1/f$ noise originating from the sensor array and preamplifiers. The obtained time-resolution leads to sub-degree angular resolution in localization of acoustic sources. The chip is suitable for integration with a MEMS microphone array, for small, compact, battery-operated applications in surveillance and hearing aids.

Acknowledgments: This work was partly supported by ONR N00014-99-1-0612, ONR/DARPA N00014-00-C-0315 and N00014-00-1-0838. The chip was fabricated through the MOSIS service.

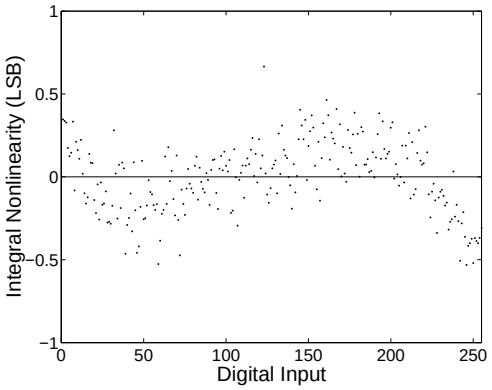

Figure 8. Multiplying D/A converter characteristic

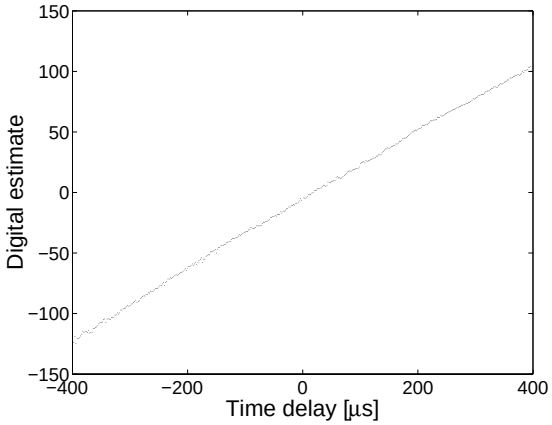

Figure 9. Digital estimation of delay τ_{10} for synthetic 200Hz sine wave signals

[1] J. Lazzaro and C.A. Mead, "A silicon model of auditory localization", *Neural Computation*, vol. **1**, pp. 47-57, 1989.

[2] N. Bhadkamkar and B. Fowler, "A sound localization system based on biological analogy", *IEEE Int. Conf. on Neural Networks*, vol. **3**, pp. 1902-1907, 1993.

[3] P. Julian, A.G. Andreou, P. Mandolesi and D. Goldberg, "A low-power CMOS integrated circuit for bearing estimation", *Proc. IEEE Int. Symp. on Circuits and Systems (ISCAS'2003)*, Bangkok, Thailand, 2003.

[4] C.C. Enz and G.C. Temes, "Circuit Tecniques for Reducing the Effects of Op-Amp Imperfections: Autozeroing, Correlated Double Sampling, and Chopper Stabilization", *IEEE Proceedings*, vol. **84** (11), pp. 1584-1614, 1996.

[5] M. Stanacevic and G. Cauwenberghs, "Mixed-signal gradient flow bearing estimation", *Proc. IEEE Int. Symp. on Circuits and Systems (ISCAS'2003)*, Bangkok, Thailand, 2003.