ICSV13 - Vienna
The Thirteenth International Congress on Sound and Vibration
Vienna, Austria, July 2-6, 2006

THE USE OF MEMS COCHLEAR MICROPHONE "FISHBONE" FOR ARBITRARY COMPLEX FILTERING AND HILBERT TRANSFORM

Shigeru ANDO*[1], Nobutaka ONO[1], and Naoki IKEUCHI[2]

[1]Department of Information Physics and Computing, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
[2]Tokyo Electron AT Limited
1-8 Fuso-cho, Amagasaki, 661-0891, Japan
ando@alab.t.u-tokyo.ac.jp (e-mail address of lead author)

Abstract

We are studying on "Fishbone" acoustic sensor/actuator (IEEE/ASME Trans. Mechatronics, 3, 2, 98/105, 1998) which is mathematically equivalent to the Zwislocki basilar membrane model of the human cochlea. Its remarkable features are: 1) frequency-decomposed outputs, 2) log-linearity, 3) perfect efficiency, and 4) wide frequency and dynamic ranges. We propose here a novel transduction method and circuit being suitable for single-chip and wireless acoustic/vibrometric sensor utilizing micromechanical and CMOS fabrication technologies. It has digitally adjustable/selectable complex frequency characteristics for tracking and extracting a target sound from environmental noise, radio-frequency (RF) outputs with direct modulation technique on resonant beams, and Hilbert transform capability to produce real and imaginary signal pair after quadrature-demodulating the RF output. We show theoretical analysis and experimental results using a MEMS fabricated sensor chip.

1 INTRODUCTION

In the mammalian auditory system, sound is perceived after it is decomposed into numerous frequency subbands by the basilar membrane (BM) of the cochlea. Helmholtz, in his book published in 1862, likened the cochlea to a bank of high Q resonators which are tuned to different frequencies. G. von Békésy [1] found that the cochlea is more like to a waveguide on which sound vibration is traveling inward. Based on his observation, Zwislocki proposed a transmission line model [2] and expressed it as an equivalent electric circuit. The model was accurate enough to explain a lot of detailed experimental results [3]. Recent physiological studies have concentrated on nonlinear and active properties of the cochlea through extending the transmission line model [4, 5]. The artificial realization of the frequency analysis function like the cochlea has also been studied. The analog electronic cochlea [6] is a cascaded active electronic circuit of 2nd order low pass filters, and MEMBAC [7] is an array of parallel close strings made of Si. These implementations, however, are different from the real cochlea particularly in the sophisticated roles of hydrodynamic interaction of the cochlea.

We have been studying on applications of the cochlear mechanism as an efficient structure for sensors and actuators [8, 9, 10, 11, 12, 14]. We call the structure “Fishbone”, which has a Si micromachinable shape derived from the mechanical analogy of the transmission line model by Zwislocki. The remarkable features of this structure are: 1) highly sensitive and efficient sensing based on resonance, 2) integrated sensing and frequency decomposition, 3) wide dynamic range and low interharmonic distortion, and 4) close resemblance with frequency characteristics of the human auditory system. To make the best use of these features, however, we must develop a suitable transduction scheme of vibration pattern into electric signals being appropriate for specific applications. A method is the weighted addition type transduction scheme[13] by which a dynamically controllable frequency characteristics is obtained including the enhancement of frequency selectivity. In this paper, as a significant extension of it, we propose a novel transduction scheme with an arbitrary complex filtering and Hilbert transform capability. It is suitable for single-chip, RF-interfaced distributed acoustic sensor utilizing micromechanical and CMOS fabrication.

2 Cochlear Model and Fishbone Structure

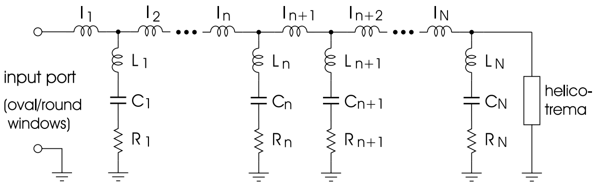
2.1 Transmission line model

The basilar membrane of cochlea, a fluid-mechanical system originally, is converted through the theory of analogy into an equivalent electrical circuit model shown in Fig.1(a). As illustrated in Fig.1(b), this model describes a process of: 1) inward transmission of the pressure and velocity (i.e. power) of the mechanical vibration input at oval and round window through the fluid-mechanical transmission line made of scala vestibuli, basilar membrane, and scala tympani, and 2) decomposition and delivery of the frequency components into basilar membrane resonators with individual resonance frequencies.

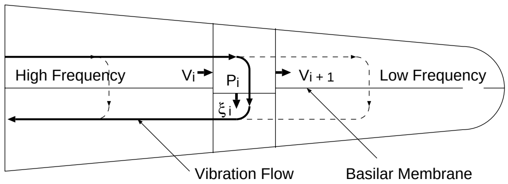
Fishbone structure proposed by us is based on the transmission line model proposed by Zwislocki (Fig.1(a)). It is a mechanical analogy of the circuit model 1(a), and, as shown in Fig.2, it consists of a single transversal beam and perpendicular resonator beams being arranged in the length order from short to long[12]. The structure itself is loss-free except for a small viscous resistance by air. The vibration energy is therefore thoroughly delivered to the resonator beams and transduced and absorbed there as electrical signal energy.

2.2 Log-linearity and pure resistivity

It is well known that a remarkable property of the basilar membrane of the cochlea is the log-linear order of the resonance frequencies (an octave interval of the resonators is uniform). This property is suitable to analyze harmonic sounds like music or speech. Another benefit of the log-linearity we have shown[**?**] is that the input impedance of the cochlea becomes purely

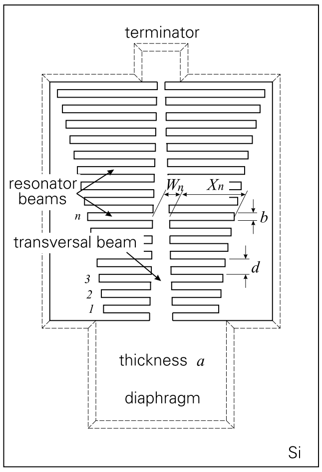


(a) Circuit model of cochlea

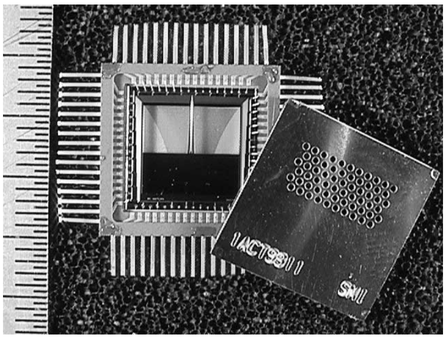


(b) Pass of vibration energy in cochlea

Figure 1: The transmission line model of cochlea. The stapes pressure is represented by the input voltage at the left side, and series inductors represent fluid inertia along the length of the cochlea. The serial resonators represent the mechanical impedance of an element of the basilar membrane.



(a) Fishbone structure



(b) photograph of a device

Figure 2: Fishbone structure (a) and the MEMS chip (b). Both sides of a transverse beam are supported by diaphragms. The frequency range is about 4kHz to 8kHz (32 channels). The 1ACT9921 used in the experiment is a same device except for a number of resonant beams (25 channels).

resistive for any frequencies if the parameters of it have the perfect log-linear symmetry. This means that the energy efficiency is equally maximized in whole frequency range.

3 WEIGHTED ADDITION TYPE TRANSDUCTION SCHEME

3.1 Piezo resistive transduction

The piezo resistive transduction[**?**, 13] is a method to obtain an electric signal proportion to the vibration amplitude from the change of resistance due to the strain of vibration. The piezo resistors are formed near the base of the resonant beam by diffusion. Let ν be a Poisson ratio of the structure, l be a length of the diffusion area, and ρ be a resistance coefficient. Then, the relative change of resistance of the resistance R is expressed as

$$\frac{\Delta R}{R} = (1+2\nu)\frac{\Delta l}{l} + \frac{\Delta \rho}{\rho} \simeq \pi E \frac{\Delta l}{l}, \tag{1}$$

where π and E are the piezo coefficient and the Young's modulus of crystal Si ($\pi E \simeq 100$).

3.2 Variable weighted addition

We call a transduction scheme "addition type" in which the vibration output of each resonant beam is additively accumulated and output through a single transversal wire. In this scheme, roles of the Fishbone sensor are: 1) efficient decomposition and conversion of mechanical vibration into multi-channel electrical signals, and 2) electrical and flexible synthesis of the frequency components into an electric output signal.

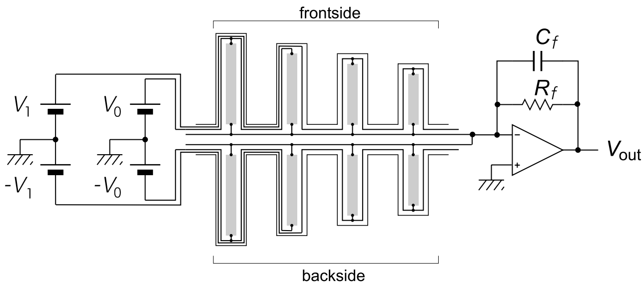


Figure 3: An addition type transduction using piezo resistors and a summing amplifier. By controlling bias voltages of the resistors, adjustment of transduction gain of each beam becomes possible.

Fig.3 illustrates a method to realize this scheme. Let a resistance of a piezo resistor on ith resonator beam be $R_i + \delta R_i(t)$, a bias voltage of the piezo resistor be V_i. Then the current flows into the imaginary ground of the operational amplifier is expressed as

$$I = \sum_{i=1}^{N} \{ \frac{V_i}{R_i + \delta R_i(t)} - \frac{V_i}{R_i + \delta R_i(t)} \} \simeq -\sum_{i=1}^{N} (\frac{2V_i}{R_i})(\frac{\delta R_i(t)}{R_i}). \tag{2}$$

Therefore, the output of the amplifier is

$$V_{out} = \sum_{i=1}^{N} (\frac{2R_f V_i}{R_i})(\frac{\delta R_i(t)}{R_i}) \tag{3}$$

where R_f is the feedback resistor. The weight $2R_f V_i / R_i$ for each vibration signal $\delta R_i(t)/R_i$ can be arbitrarily real-valued and time-varying.

4 ARBITRARY COMPLEX FILTERING AND HILBERT TRANSFORM

A problem of the weighted addition method is that the realizable impulse responses in time-domain are restricted to the symmetric or anti-symmetric ones (e.g. averaging or differential).

We consider in this section a method to obtain arbitrary frequency characteristics based on the variable weighted addition scheme. Applications that require general impulse response shapes include matched filters for target waveform, correction filters of phase distortion, etc.

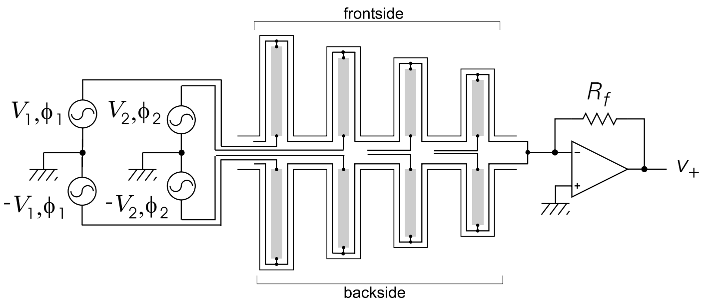


Figure 4: RF modulating transduction scheme. It changes the bias voltage of piezo resistor in the frequency Ω, amplitude V_i, and the phase ϕ_i. The amplitude and phase determine the magnitude and phase of the complex frequency characteristics of the weighted addition scheme. The carrier frequency can be sufficiently larger so that it and its harmonic frequencies are far distant from the vibration frequencies. It enables the use of binary oscillator circuits and RF transmission.

4.1 Principle

As shown in Fig.4, Let the resonant frequency of the ith beam be ω_i, the resistance be R_i, the resistance change be $\delta R_i(t)$. The common ends of the piezo resistors (front side and back side) are connected to the current-input voltage-output amplifier with the transimpedance R_f. The other ends are driven by the sinusoidal voltages with the frequency Ω, amplitude V_i, and phase ϕ_i and $\phi_i + \pi$ for the front side and back side, respectively. Then the current flowing from the ith beam to the amplifier is expressed as

$$\begin{aligned} \frac{V_i \cos(\Omega t + \phi_i)}{R_i + \delta R_i(t)} - \frac{V_i \cos(\Omega t + \phi_i)}{R_i - \delta R_i(t)} &\simeq -\frac{2V_i}{R_i} \cos(\Omega t + \phi_i) \frac{\delta R_i(t)}{R_i} \\ &\equiv -H_i \cos(\Omega t + \phi_i) F_i(t, \omega_i), \end{aligned} \tag{4}$$

where $H_i \equiv 2V_i/R_i$ is the magnitude coefficient of modulation, $\cos(\Omega t + \phi_i)$ is the carrier and phase of modulation, and $F_i(t, \omega_i) \equiv \delta R_i(t)/R_i$ is the vibration input of the beam. The sum of the current from all beams are expressed as

$$g(t) = -\sum_{i=1}^{N} H_i \cos(\Omega t + \phi_i) F_i(t, \omega_i) \tag{5}$$

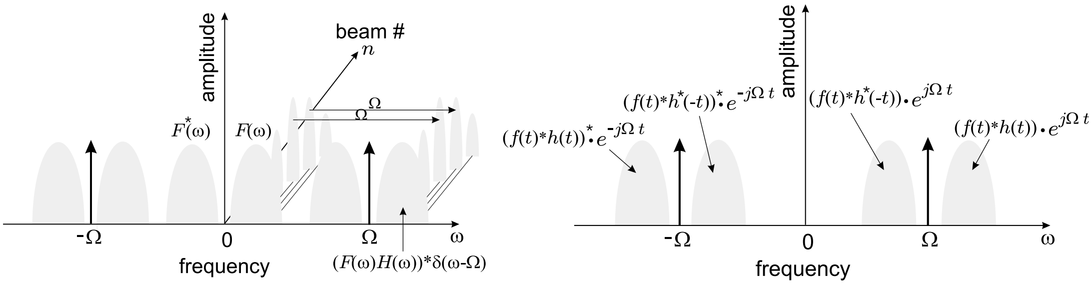
where N is the number of beams. Since $F_i(t, \omega_i)$ has a very narrow spectrum according to the frequency decomposing nature of the Fishbone structure, we can express them as

$$\begin{aligned} F_i(t, \omega_i) &\simeq \Re\{F(\omega_i) e^{j\omega_i t}\} = \frac{1}{2}\{F(\omega_i) e^{j\omega_i t} + F^*(\omega_i) e^{-j\omega_i t}\}, \\ H_i \cos(\Omega t + \phi_i) &\simeq \Re\{H(\omega_i) e^{j\Omega t}\} = \frac{1}{2}\{H(\omega_i) e^{j\Omega t} + H^*(\omega_i) e^{-j\Omega t}\}, \end{aligned}$$

where $F(\omega)$ is the spectrum of input sound (wideband) and $H(\omega)$ is the desired complex frequency characteristics. Since $F(\omega), H(\omega)$ are nonzero only in $\omega > 0$ from definition, the output current is rewritten as

$$\begin{aligned} g(t) &= -\frac{1}{4}\sum_{i=1}^{N}(F(\omega_i)e^{j\omega_i t} + F^*(\omega_i)e^{-j\omega_i t})(H(\omega_i)e^{j\Omega t} + H^*(\omega_i)e^{-j\Omega t}) \\ &\simeq -\frac{1}{2\Delta\omega}\{\Re\{e^{j\Omega t}\int_0^\infty F(\omega)H(\omega)e^{j\omega t}d\omega\} + \Re\{e^{j\Omega t}\int_0^\infty F^*(\omega)H(\omega)e^{-j\omega t}d\omega\}\}, \end{aligned}$$

which implies the analytic (Hilbert transform is in the imaginary part) input signal $f(t) = \mathcal{F}^{-1}\{F(\omega)\}$ is passing through the filter with complex characteristics $H(\omega)$, and then appears as the RF modulated signal at the upper sideband of the carrier frequency Ω. The lower sideband consists of the filter output of $f(t)$ by the impulse response $-h(-t) = \mathcal{F}^{-1}\{H(-\omega)\}$.



(a) decomposition-modulation-addition process (b) spectral components of output

Figure 5: Spectral structures of the arbitrary complex filtering and Hilbert transform. Around the carrier frequency Ω, two components can be obtained. One is the desired complex-filtered output and the other is the filtered result of the conjugate filter. The HF bandpass filter (e.g. ceramic filter for RF receiver) can easily isolate the desired frequency component.

The procedure is illustrated in Fig.5(a). The sound vibration is decomposed into narrow band vibrations by Fishbone. The each one is then converted to a current signal and multiplied by $H_i \cos(\Omega t + \phi_i)$. It magnifies the amplitude by H_i, changes the phase by ϕ_i, and shifts the center frequency by Ω. The addition circuitry sums up all the signals into the output, which is in the upper sideband the sound vibration filtered by $H(\omega)$ followed by the modulation with a constant carrier Ω. A noticeable difference is between the above scheme and the simple filtering followed by the modulation. In the latter case, negative frequency components are multiplied by $H^*(\omega)$ whereas it is multiplied by the same $H(\omega)$ in the former as shown in Fig.5(b). By rewriting the lower sideband component as

$$e^{j\Omega t}\int_0^\infty F^*(\omega)H(\omega)e^{-j\omega t}d\omega = e^{j\Omega t}\{\int_0^\infty F(\omega)H^*(\omega)e^{j\omega t}d\omega\}^* = \{f(t) * h^*(-t)\}e^{j\Omega t},$$

it is verified that the component generates the time-reversed impulse response of the desired one. Therefore a simple AM demodulation, which is efficient in the latter case, collapse the result by the simultaneous generation of the desired output and the time-reversed impulse response output. The isolation of the upper sideband is necessary before demodulation.

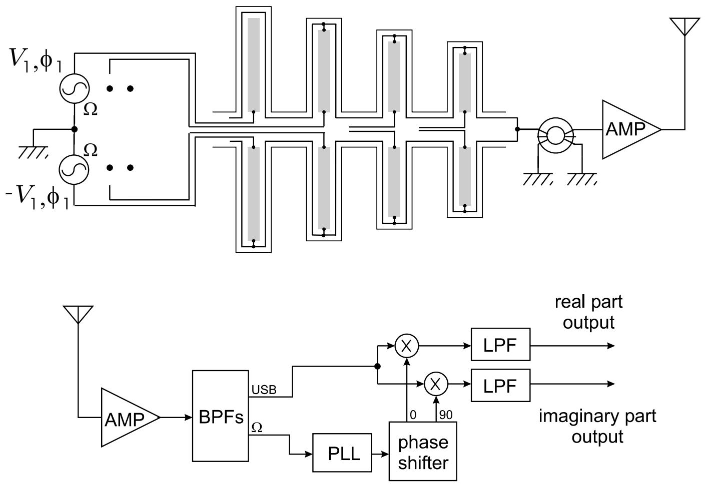
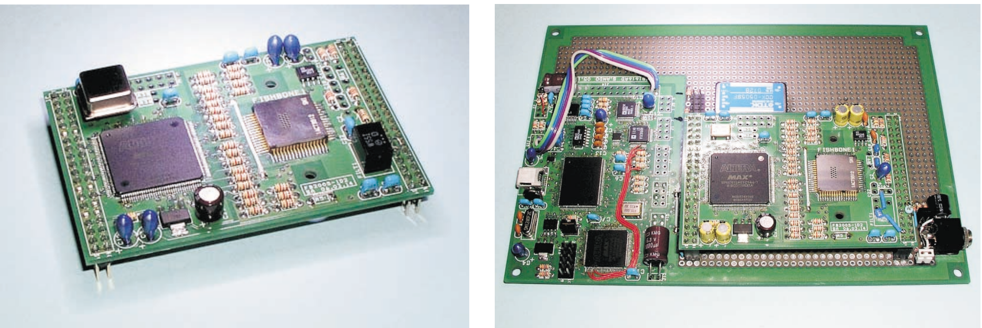


Figure 6: RF-interfaced arbitrary complex filtering microphone. The upper figure (a) is the transmitter unit, and the lower one (b) is the receiver unit with Hilbert transform outputs.

4.2 Transmission and demodulation method

The above theory shows that the upper sideband and the lower sideband carry different outputs with individual meanings. Separation of them can generate two useful signals. The separation is not necessarily performed near the Fishbone. In the RF transmission scheme illustrated in Fig.6(a), the sensor transmits the modulated signal as it is, and the receiver separates the upper and lower sideband components, and then demodulates orthogonally the component into Hilbert transform pair of real and imaginary signals. The in-phase and quadrature-phase sinusoids necessary for demodulation are recovered by a PLL circuit locked with the carrier.



(a) Fishbone and modulation circuit (b) USB-based controller board

Figure 7: Experimental system of the arbitrary complex filtering microphone. The board consists of a Fishbone chip (1ACT9921) and a modulation waveform generator. The board size is about 5cm×7cm.

4.3 Experimental System

A system for experimental verification using a Fishbone chip 1ACT9921 (Tokyo Electron Ltd.) and an FPLD for modulation waveforms generation is shown in Fig.7(a),(b). The bias voltages are rectangular waveforms with a frequency 449kHz and variable phase and amplitude in 128 levels in $[-\pi, \pi]$ and in two levels $(0, 3.3V)$, respectively.

5 POSSIBLE APPLICATIONS

The first target will be tracking and extraction of a sound source in very noisy environment. The advantage of Fishbone sensor, i.e., low interference between spectral components and large dynamic range, will become clear in this situation. The second target will be the distributed and networked sound detection in resonant and/or heterogeneous environment such as in a pipe, near a reflector, or in a car. By dynamically adapting the frequency characteristics of the sensor, it can not only recover the original sound but also make a good use of the frequency characteristics as a novel source of environmental information.

References

[1] G. von Békésy, “*Experiments in Hearing*” McGRAW-HILL, pp.485-534, 1960
[2] J. J. Zwislocki, Acta Otolaryngol. Suppl., vol. 72, pp.1-76, 1948.
[3] W. S. Rhode, J. Acoust. Soc. Am., vol. 49, pp.1218-1231, 1971.
[4] E. Zwicker, Biol. Cybernetics 35, pp.243-250, 1979.
[5] E. Zwicker, J. Acoust. Soc. Am., vol. 80, pp.146-152, 1986.
[6] R. F. Lyon, and C. Mead, IEEE Trans., Acoust., Speech, and Signal Processing, vol. 36, no.7, pp.1119-1134, 1988.
[7] D. Haronian and N. C. MacDonald, Proc. Transducers’95, pp.708-711, 1995.
[8] K. Tanaka, M. Abe, and S. Ando, Proc. 14th Sensor Symposium, pp.97-100, 1996.
[9] M. Harada, N. Ikeuchi, S. Fukui, T. Osa, K. Tanaka, and S. Ando, Proc. 15th Sensor Symposium, pp.99-102, 1997.
[10] K. Tanaka, M. Abe, S. Ando, Proc. IEEE/ASME Int. Conf. Advanced Intelligent Mechatronics, A035, 1997.
[11] S. Ando, K. Tanaka, and M. Abe, Proc. Transducers’97, Chicago, pp.1027-1030, 1997.
[12] K. Tanaka, M. Abe, S. Ando, “A novel mechanical cochlea “Fishbone” with dual sensor/actuator characteristics,” IEEE Trans. Mechatronics, vol.3, no.2, pp.98-105, 1997.
[13] Y. Kusakabe, M. Abe, S. Ando, N. Ikeuchi, and M. Harada, Proc. 16th Sensor Symposium, pp.249-252, 1998.
[14] M. Harada, N. Ikeuchi, S. Fukui, H. Toshiyoshi, H. Fujita, and S. Ando, Trans. IEEJ, vol.119-E, no.3, pp.125-130, 1999.
[15] N. Ono, Y. Zaitsu, T. Nomiyama, A. Kimachi, and S. Ando, Trans. IEEJ, vol.121-E, no.6, pp.313-319, Jun. 2001.
[16] N. Ono, A. Saito, and S. Ando, Trans. IEEJ, vol.123-E, no.3, pp.92-97, 2003.
[17] N. Ono, A. Saito, and S. Ando, Proc. Transducers’03, pp.939-942, Boston, 2003.
[18] N. Ono, T. Arita, Y. Senjo, and S. Ando, Proc. Transducers’05, Seoul, 2005.