MEMS-based Vibration Sensor in Single-Backplate Technology

Somu Goswami[1], Christian Bretthauer[1], Andreas Bogner[1], Abhiraj Basavana[1], Sebastian Anzinger[1], Marco Haubold[1], Gunar Lorenz[1], Johann Strasser[1], Daniel Weber[1], Lorenzo Servadei[1,2], Robert Wille[2]
[1]Infineon Technologies AG, Neubiberg – Bayern, Germany
[2]Technische Universität München, Munich – Bayern, Germany
E-mail: [1]name.surname@infineon.com, [2]name.surname@tum.de

Abstract*—Quality of voice communication using True Wireless Stereo (TWS) is not ideal in noisy environments. In TWS, two outward-facing microphones per earbud and acoustic beamforming techniques are typically used to isolate the user's voice. Due to space constraints, this approach cannot be further improved by adding more microphones. Here, the use of bone-conducted voice pickup holds promise for improving voice calling quality. There are some microphone-based bone-conducted voice pickup solutions on the market that either have insufficient vibro-acoustic performance or are not ideal for use in consumer electronics due to their size. This work conceptualizes a novel vibration sensor that overcomes the limitations of existing solutions while meeting the requirements for such a vibration sensor. The proposed concept is based on industry-proven single-backplate Micro-Electro-Mechanical Systems (MEMS) technology with an attached proof-mass that allows bone-conducted voice pickup. Simulation and characterization of the fabricated MEMS and assembled prototypes show that the proposed solution can achieve a high vibration sensitivity of -29.7 dBV/g and a high Signal to Noise Ratio (SNR) of 70 dBA, in a small 3.0x2.0x0.8 mm^3 package size. Therefore, it is overall well suited to for voice communication using bone-conducted voice in TWS.

***Index Terms*—MEMS, TWS, vibro-acoustics, bone conduction, vibration sensing, voice pickup**

I. INTRODUCTION

Over the last years, True Wireless Stereo (TWS) have seen rapid uptake by the consumer electronics market. These devices offer features like active noise cancellation, transparent hearing, voice assistance and calling. High-resolution capture of the user's voice for calls or digital assistants is still a major challenge for device OEMs to address. Most of today's TWS use two of three microphones per ear to pick up the user's voice. The acoustic pickup is done by the two omnidirectional microphones, which are then processed by beamforming and noise reduction algorithms, similar to the working principle presented in [1, 2]. However, isolating the user's voice when there are multiple speakers or in a noisy environment is problematic in such implementations, resulting in degraded call quality.

Voice pickup via the bone-conducted path is not contaminated by ambient noise and provides an alternative approach to solving the problem of voice calls under noisy conditions [3,4]. In addition to the vocal folds, the human voice is also transmitted through the skull (bone), where the sound vibrates the larynx [5, 6]. Existing bone-conducted pickup solutions offer either good performance in a large package size [7, 8] or inadequate performance in a small package size [9].

This paper proposes a concept for a vibration sensor that overcomes the limitations of existing solutions. The concept is based on the single backplate Micro-Electro-Mechanical System (MEMS) technology work of Dehé et al. (2013), that has become an industry standard for mature and commercial single backplate microphone processes. It comprises of a poly silicon backplate and a membrane structure. In this work, we redesigned the single backplate MEMS structure to develop a novel MEMS that can reliably pick up bone-conducted voice. The concept and design are validated through iterative simulation and fabrication steps. The new MEMS can be used with existing building blocks of a standard microphone, namely the existing assembly process and Application Specific Integrated Circuit (ASIC) for readout.

II. RESEARCH BACKGROUND

The research work presented by Reinfeldt [10], Stenfelt et al. [11], and other publications contribute to the understanding of the behavior of bone-conducted voice [12-14]. Based on these studies, a summary of the TWS application requirements for a vibration sensor are articulated below:

- *Low acoustic sensitivity: To avoid pick up of environmental acoustic noise.*
- *Small package size: Easy mounting and positioning in a TWS, as the placement of a contact vibration sensor determines the frequency characteristics, signal strength, sound quality and intelligibility .*
- *A high bandwidth of >4 kHz: Detect the frequency components of the bone-conducted voice up to 4 kHz (higher frequencies are muffled by the skin).*
- *High vibration sensitivity: Sensing of vibro-acoustic energies even below 10 mg of bone-conducted vibrations.*
- *High SNR/low noise floor: Avoid masking of low amplitude vibrations by the sensors self-noise.*
- *Microphone like audio interface: Facilitates easy combination with microphone signals without intermediate conversions, which could increase audio delay.*
- *Simple assembly and packaging: To reduce assembly tolerances and improve reliability.*

There are a few existing concepts that address the target use-case in TWS e.g. the sensing concept presented by You, Lo et al. [15]. This concept uses three components, namely the proof-mass, the microphone MEMS and ASIC. These are assembled in a two-chamber package. When the sensor is subjected to vibration, the proof-mass in the first chamber experiences acceleration and generates a pressure load on the MEMS in the second chamber. This pressure change is coupled to an ASIC, which converts it into an acoustic signal. There are some patents and existing products in the market based on similar concepts [7]. Their performance specifications are not fully known. In principle, such a concept has the advantage of having a larger proof-mass (usually bulk metal) and large volume chambers on both the front and back sides of the MEMS microphone. This makes it possible to achieve high sensitivity and high SNR. However, the assembly and packaging of such a concept requires additional steps to bond the bulk metal and apply a bottom cover, which both increases assembly, packaging and material costs and adds more variables that potentially could impact specification tolerances for a high-volume manufacturing product.

Besides the two-chamber concepts, there is one publicly available product that has a standard microphone package concept with a piezoelectric MEMS. As per the datasheet, this concept has a smaller package size, simpler mounting and assembly steps, but the sensor has a low SNR of 59 dBA [9]. This can affect performance as the captured signal can be too noisy and in scenarios with sub-optimal skull contact or when the user is not speaking loudly enough, the speech vibrations would be masked by the self-noise of the vibration sensor.

Therefore, it can be stated that neither the two-chamber concept nor the piezoelectric solution are able to fully meet the application requirements of a vibration sensor. The work presented in this paper is aimed towards addressing this gap in technology and is explained in detail in the next section.

III. PROPOSAL

In this paper, we propose to modify existing MEMS microphone design in two ways to create a vibration sensor:

- MEMS modification: Process technology steps are changed to increase the vibration sensitivity.
- Package modification: Soundport is closed to reduce the acoustic sensitivity of the MEMS.

The modification of the package ensures that unwanted acoustic background noise is not picked up. However, closing the soundport of a standard MEMS microphone causes a significant drop in acoustic and vibration sensitivity. As such, the modification on the MEMS level is needed to increase the vibration sensitivity while keeping sufficient bandwidth.

The sensitivity of a MEMS structure is directly proportional to the membrane mass and inversely proportional to its stiffness. The resonance frequency of the MEMS structure determines the usable bandwidth, which is inversely proportional to the sensitivity. This means that our MEMS modifications address three challenges: Increasing sensitivity while maintaining sufficient usable bandwidth without increasing the self-noise (i.e. lowering SNR).

As reducing the membrane stiffness to increase sensitivity would require many changes to proven MEMS technology steps and also increase the self-noise, we decided to increase the mass of the MEMS element, specifically the mass of the moving membrane. This is done by leaving an inertial mass (proof-mass) affixed to the membrane in an additional, new MEMS process step.

Both package and MEMS modifications are built on mature and proven process technology and packaging steps. In addition, our concept is designed to be compatible with existing ASICs with a standard audio interface and that could fit in a small housing. The proposed vibration sensor concept is sketched as shown in Figure 1.

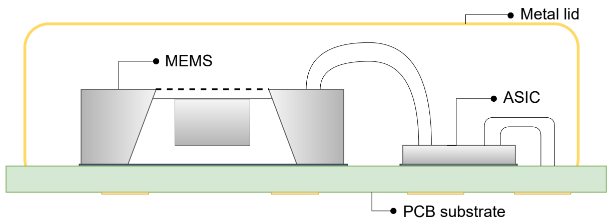

Figure 1. Proposed concept for a vibration sensor - MEMS with affixed proof-mass connected to an ASIC and encapsulated in a standard microphone housing without a soundport

To validate the concept, all ideas sketched above have been incorporated in a full-fledged design and assembled within an industrial environment. Intensive finite element simulations were done to optimize the dimension and characteristics of the proof-mass, followed by iterative fabrication steps and characterization of the assembled prototypes. This process is covered in more detail in the next sections.

IV. SIMULATION

In order to simulate the characteristic behavior of the vibration sensor, a quarter MEMS Finite Element Model (FEM) based DoE was designed on a simulation platform (MEMS+®). The outcome of the simulations exhibited a MEMS sensitivity of -29.9 dBV/g at 1 kHz and 1g (1g = 9.81 m/s^2) along the z-axis (out of plane). From the simulations, it can be estimated that the MEMS resonance is well above the usable bandwidth and from the MEMS noise calculations, the MEMS SNR is estimated to be 70.7 dBA relative to 1g for a bandwidth of 100 Hz to 4 kHz.

From the simulations it can be inferred that the concept can reach the vibro-acoustic requirements i.e. high sensitivity and high SNR. It is also able to meet the bandwidth requirements relevant for bone-conducted voice pickup. It has to be noted that the simulation only takes into account the MEMS behavior with certain parametric assumptions. Therefore, in order to validate the simulation results and characterize the overall system behavior, the MEMS are fabricated and assembled together with existing microphone ASICs in a 3.0x2.0x0.8 mm^3 package.

V. Characterization Results

The vibration sensor is assembled according to the procedure described in [17], but instead of a pre-molded cap, a metal lid is used (as is the standard for MEMS microphones now). Figure 2 shows the prototype samples at different stages of assembly. After assembly each prototype sample is electrically verified for connectivity and signal before characterization.

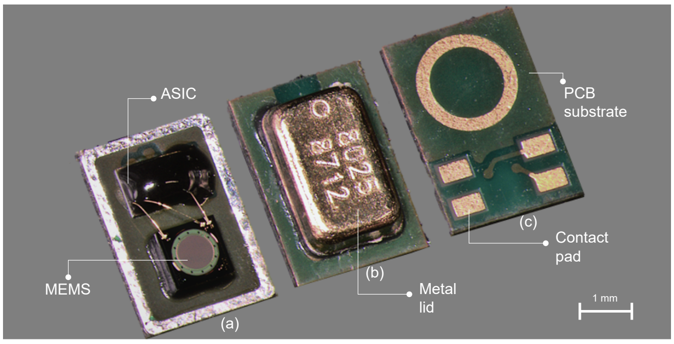

Figure 2. Micrograph of vibration sensor prototypes (a) MEMS and ASIC on a PCB substrate, (b) metal lid enclosing the MEMS and ASIC in a 3.0x2.0x0.8 enclosure and (c) bottom side of the PCB substrate with electrical contact pads and enclosed soundport. The PCB substrates are designed a bit bigger for re-use of existing set-up.

The vibration sensor prototypes are characterized using a test-fixture as shown in Figure 3. The test-fixture comprises of a reference accelerometer (B&K 4533-B) fitted on top of the vibration shaker (B&K mini shaker 4810). The Device Under Test (DUT) is mounted on a PCB and fixed on top of reference accelerometer with a 3D-printed fixture and a screw. A signal generator is used to excite the shaker and the acceleration is picked up by the reference accelerometer and the DUT. The signals acquisition of the sensors and equalization of the shaker response are carried out using an acoustic analyzer (Audio Precision® APx525). The parameters measured include noise FFT, sensitivity and frequency response behavior. A set of three prototypes with the same MEMS design characteristics are measured.

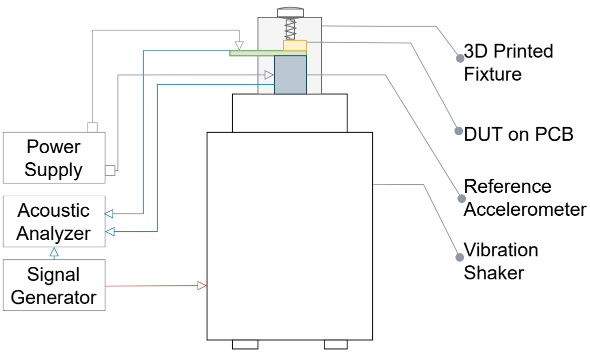

Figure 3. Set-up used to characterize vibration sensor prototypes

The resonance peak of the frequency response curve as shown in Figure 5 is measured to be outside the bandwidth of interest, i.e. a usable bandwidth of 4 kHz is attained. The average of the measured sensitivity of the prototypes at 1 kHz and 1g is -29.7 dBV/g as shown in Figure 5. The pure noise of a vibration sensor prototype is shown in Figure 4 and its corresponding SNR for a bandwidth of 100 Hz to 4 kHz is 70 dBA. The standard deviation between the sensitivity of the three measured prototypes is 0.25 dBV/g.

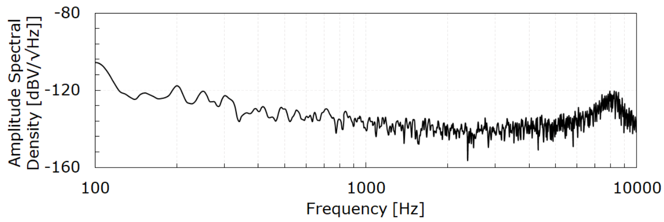

Figure 4. Pure noise measurement of a vibration sensor prototype (Amplitude Spectral Density).

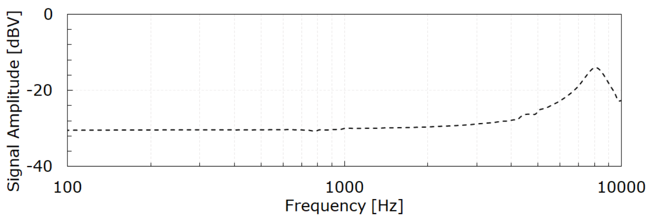

Figure 5. Frequency response measurement of a vibration sensor prototype

The characterization results confirm that the proposed concept can be fabricated and reliably characterized in a shaker setup. The vibro-acoustic performance, especially the high vibration sensitivity and high SNR, also confirm the simulation results. The low standard deviation between prototypes shows that the proposed design provides initial evidence of stable and repeatable performance. In addition, the results confirm that the MEMS design is compatible with existing microphone ASICs, enabling standard audio readout.

VI. Conclusions

Bone-conducted voice pickup is a promising means to solve the problem of voice communication in noisy environments in a TWS. This paper presents a novel vibration sensor concept that fills the technology gap by overcoming the drawbacks of existing solutions. Through simulation and characterization of the developed prototypes, it was shown that the proposed vibration sensor can achieve wide bandwidth, high vibration sensitivity and high SNR. All this combined with a standard audio interface, small size, and easy assembly and packaging makes our concept well-suited for bone-conducted voice pickup in TWS.

REFERENCES

[1] Goldin, Autodirective Dual Microphone, AES Convention, 2003.

[2] McCowan, Sridharan, Adaptive Parameter Compensation for Robust Hands-free Speech Recognition using a Dual Beamforming Microphone Array, 2001.

[3] Wesley S, Henry Y et al., Bone-conduction Pickup Transducer for Microphonic Applications, United States Patent: US8983096B2, 2013.

[4] Shin, Kang et al., Survey of Speech Enhancement Supported by a Bone Conduction Microphone, ITG Symposium, 2012.

[5] Shimamura, Tamiya et al., A Reconstruction Filter for Bone-Conducted Speech, Proc. of the Mid-west Symposium on Circuits and Systems, 2005.

[6] Kondo, Fujita et al., On Equalization of Bone Conducted Speech for Improved Speech Quality, IEEE International Symposium on Signal Processing and Information Technology, 2006.

[7] Mögelin, and Post, Vibration Sensor with Low-frequency Roll-off Response Curve, United States Patent: US20180058916 A1, 2018.

[8] Lee, Seong, et al., Optimization and Performance Evaluation of a Transducer for Bone Conduction Implants, IEEE Access 2020.

[9] VA1210 Bone Conductor Sensor – Analog Piezoelectric MEMS Voice Accelerometer with Hardware Voice Activity Detection, Technical specifications, Vesper Technologies Inc., 2020.

[10] Reinfeldt, Doctoral Thesis-Bone Conduction Hearing in Human Communication - Sensitivity, Transmission, and Applications, 2009.

[11] Stenfelt, Goode, et al., Bone-Conducted Sound: Physiological and Clinical Aspects, Otology and Neurotology, Inc., 2005.

[12] Rahman, Shimamura, Pitch Characteristics of Bone Conducted Speech, Proc. of the European Signal Processing Conference (EUSIPCO), Denmark, 2007.

[13] Tran, Letowski et al., The effect of Bone Conduction Microphone Placement on Intensity and Spectrum of Transmitted Speech Items, 2013.

[14] McBride, Tran et al., The Effect of Bone Conduction Microphone Locations on Speech Intelligibility and Sound Quality, Journal of Applied Ergonomics, 2011.

[15] You, Lo et al., Design and Implementation of a Novel Skull Vibration Sensing Module for Bone Conduction Microphone, IEEE MEMS, 2020.

[16] Dehé, Füldner et al., The Infineon Silicon MEMS Microphone, AMA Conferences, 2013.

[17] Dehé, Silicon Microphone Development and Application, Sensors and Actuators A 133, 2006.