Bio-inspired, nonlinear and adaptive acoustic sensing - Study of sensor design

Claudia Lenk[1], Tzvetan Ivanov[1], Vishal Gubbi[1], Kalpan Ved[1], Martin Ziegler[1], Tobias Fritsch[2], Jan Küller[2], Daniel Beer[2]

[1] *Dept. Micro- and nanoelectronic systems, Technische Universität Ilmenau, 98693 Ilmenau, Deutschland, Email: claudia.lenk@tu-ilmenau.de*

[2] *Fraunhofer Institut für Medientechnologie, 98693 Ilmenau, Germany, Email: daniel.beer@idmt.fraunhofer.de*

Introduction

Sound processing systems, and in particular speech processing systems, typically are based on 3 main steps (see e.g. [1]), as schematically shown in fig.1:
(1) Sound detection (Transduction): For this step, microphones (nowadays often MEMS-based) with linear transfer characteristics, a self-noise floor of $20-33$ dB sound pressure level (SPL) and a dynamic range of $100-120$ dB are typically applied. Measurement microphones can offer a lower self-noise floor down to $0-7$ dB SPL.
(2) Signal conditioning and Feature extraction: In this step, signal conditioning like amplification, filtering as well as nonlinear transformation of the signal are applied. Furthermore, features of various complexity are extracted. Low complexity features are, e.g., the envelope or the zero crossing rate in time domain as well as central spheroid in the frequency domain. Often, a time-frequency representation of the signal, known as spectrogram, is calculated. More complex features are rhythm, chords, note onset or genre, amongst others.
(3) Sound/Speech processing: This steps describes classification, regression, and/or prediction of sounds/speech. It is typically based on neural networks (NN), most commonly on recurrent, convolutional or transformer [1].

Thereby, the signal conditioning and feature extraction stage has a strong influence on the performance of the sound/speech processing system, since it improves the clustering of data [2]. In most cases, feature extraction and sound processing are implemented software-based, requiring an analog-to-digital conversion beforehand.
Despite the strong increase in performance of these systems, several problems are not yet solved (see e.g. [2], in particular: (i) their performance drastically decreases for decreasing signal-to-noise ratios (or high levels of noise), (ii) these systems have a considerable latency due to the typically software- and cloud-based implementation of step (2) and (3), (iii) these systems mostly do not support local learning due to the complexity of speech models and required computation power.
One way to address these issues, which was inspired by the working principle of human hearing, is to integrate signal conditioning and feature extraction in the sensing/transduction stage itself (see fig. 1). In human hearing the most important processing steps are frequency decomposition, amplification and compressive transfer characteristics [3]. The latter two are thereby dynamically adaptable for changing inputs and acoustic environments. Thus, the signal is filtered and amplified before sensing takes place, increasing the signal-to-noise ratio

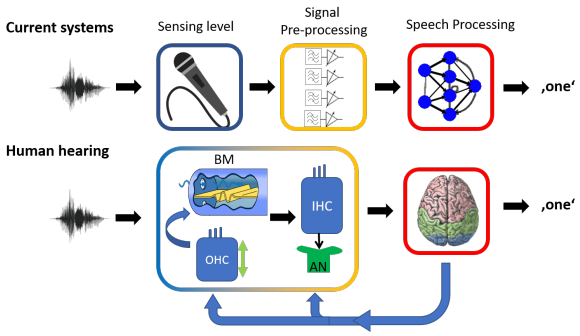

Figure 1: Comparison of automatic speech processing systems and human hearing process. While in technological systems, signal processing and feature extraction takes place after transduction of the signal ('sensing'), these steps are performed before/during the transduction process in the biological system.

considerably while decreasing the latency. Furthermore, for small signal-to-noise ratios, this effect helps taking up only the relevant signal without amplifying the noise as well. In biological hearing systems, this effect is believed to be responsible for the superior property of hearing signals with 0 dB SPL, which effectively are in the same range as the thermal noise level [3]. A sensor system incorporating these functionalities should improve the latency of the system, decrease data streaming needs and increase its overall efficiency. We developed such a sensor system [4], which will be described in detail in the next section. The goal of this study was to improve the sensing properties by re-designing the transducer.

Bio-inspired, acoustic sensor system

Our bio-inspired acoustic sensor with signal conditioning and feature extraction functionality is based on a linear acoustic sensor in connection with an high-speed feedback [4], schematically shown in fig. 2. Thereby, the transducer is a silicon beam ($350 \times 150 \times (2-5)\,\mu m^3$) with integrated deflection sensing using the piezoresistive effect and with an integrated actuator based on the bimorph effect (thermomechanical actuation principle) [7]. The feedback loop amplifies the sensing signal and adds a bias offset, before feeding the signal back to the actuator on the transducer. Feedback is realized in an FPGA structure for test purposes and in analog circuits to decrease latency decrease and increase efficiency [6]. Using this setup, it could be shown that the sensor system can undergo a Hopf bifurcation by changing either the feedback strength, corresponding to the amplification factor,

or the bias offset.

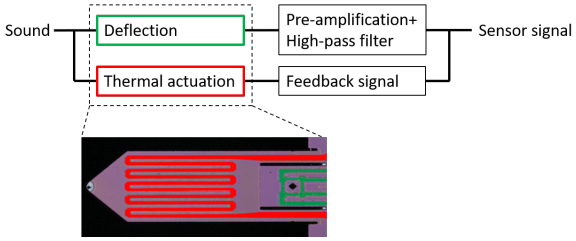

Figure 2: Schematic setup of system for bio-inspired, nonlinear, adaptive acoustic sensing.

Depending on these values, three sensing modes can be realized: (i) a passive mode with linear transfer characteristics, if the feedback is turned off, (ii) an active, linear mode, in which the gain of the transducer can be increased by increasing the feedback strength, but retaining linear transfer characteristics, and (iii) an active nonlinear mode with nonlinear (compressive) transfer characteristics, i.e. an SPL-dependent gain, which is strongly increased for low SPL. Furthermore, the sensor system performs a frequency decomposition of the signal, since each transducer is operated at its resonance. Thus it acts like a band-pass filter. Finally, it was already shown, that a dynamical adaptation of the system to different inputs can be easily implemented [6], e.g. as an amplitude dependent gain switching for highlighting the sound onset and increasing the dynamic range.
Despite these promising functionalities, the transducer design was derived for application in atomic force microscopy [5] yielding rather poor acoustic sensing properties, i.e. a self-noise level of roughly 60 dB SPL and rather large resonance frequencies of 10 kHz to more than 500 kHz, which cover only partly the audible range. Thus, to improve the acoustic sensing properties, we tailored the transducer by changing its geometry, as described in the next section.

Design considerations

The redesign shall achieve two goals: (i) decreasing the resonance frequency to enable coverage of the audible range with an array of transducers with different resonance frequencies and (ii) increase the sensitivity and decrease of the self-noise floor. The new design of the transducer is shown in fig. 3. The changes in the design according to the above listed two design goals (frequency change and sensitivity increase) are described in the following separately.
Following the Euler-Bernoulli theory for single-side mounted, rectangular beams (length≫width), one can derive a formula for dependence of the resonance frequency on the geometric parameters length l, width w and thickness d of the beam:

$$f = \frac{\omega_0}{2\pi} = \delta_n^2 \frac{d}{2\pi l^2} \sqrt{\frac{E_{\mathrm{Si}}}{12\rho_{\mathrm{Si}}}}. \qquad (1)$$

Here, E_{Si} the elasticity module, ρ_{Si} the density for Si, and δ_n a pre-factor for the n-th mode ($\delta_1 = 1.8751$). Since the thickness d can be varied only in a certain range (1 µm-

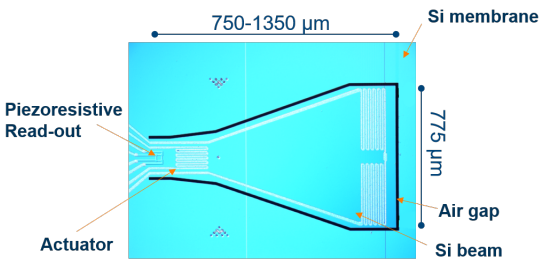

Figure 3: Sound-optimized design with increased length and larger surface for lower resonance frequency and larger sensitivity.

10 µm) with a tolerance of 1 µm, tuning the resonance frequency of the transducer should mainly rely on changing the length l followed by a subsequent fine-tuning by variation of thickness d. Thereby, to achieve frequencies of $\geq$ 1 kHz, a transducer length of up to 1350 µm is sufficient (with $d \geq 2$ µm), while for $f < 1$ kHz sensors with $l = 2600$ µm were fabricated.
To increase the sensitivity of the sensor, the force exerted on the transducer by the sound pressure had to be increased. Since the transducer operates as a pressure gradient sensor, the force can be calculated by $F = \int \Delta p(x, y) dA$ with the sound pressure drop Δp between front and back side of the transducer and A the transducer surface. Thus, increasing the sensor surface, in particular by changing its width, will yield a larger force on the transducer for the same SPL, since it increases (i) the surface, the force acts on, and (ii) the path for the sound pressure around the sensor, thus increasing the pressure drop Δp. To obtain this, first, the beam shape was modified to a more triangular-like shape with a larger width compared to the previous design and second, a membrane was added around the sensor to reduce the acoustic short circuit effect and increase the pressure gradient (see fig. 3). Membrane and beam are separated by an air gap, which has several advantages compared to the case of a sensor with the width of the membrane. First, it reduces the mass and spring constant of the sensor, which increases the deflection and sensitivity for a similar force. Second, it reduces the likeliness of the excitation of torsional modes (which become more likely the larger the width of the sensor is compared to its mounting). If the gap is smaller than the thermoviscous layers, it acts like a sound soft boundary and the sound path will be around the complete system of sensor and membrane. Thereby, the thickness of the thermoviscous layers is strongly frequency-dependent: $d_{\mathrm{visc}} = \sqrt{2\mu/\omega\rho_{\mathrm{air}}}$ with the dynamic viscosity μ, the sound frequency ω and the density of air ρ_{air}. We choose a gap width of 15 µm, thus, the air gap should be acting as soft sound boundary in the audible range. In the following sections, the sensing properties of the new design are presented obtained from measurements in an anechoic chamber and simulations performed in Comsol.

Acoustic sensing properties

To obtain sensing properties like self-noise floor, sensitivity and gain, we performed measurements of the sensor

system in an anechoic chamber (see fig. 4). For this pur-

Figure 4: Measurement setup for determining the self-noise floor and the sensitivity of the sensor system in an anechoic chamber.

pose, the loudspeaker (Neumann KA310A) was placed in a distance of 4.35 m to the sensor system, while for reference measurements another microphone (MT Gefell MV203 + MV221) was placed 5.5 cm beside the sensor system. The sound field around the sensor system was tested to be homogeneous within a tolerance of 1.5 dB SPL. To obtain the sensitivity or gain value, a frequency-sweep signal with a duration of 2 s and a frequency range of 2 or 3 kHz around the resonance frequency was applied. The root-mean square (RMS) amplitude of the driving signal was normalized to 1 V. The SPL was varied by changing the dB full scale value for the driving voltage of the loudspeaker. Sound pressure levels of $15 - 75$ dB SPL were tested, and the sensing amplitude was measured at resonance.

In fig. 5 the obtained gain of the sensor system is shown for the old design (upper graph) and the new design (lower graph). For both designs, the above described three operation modes in dependence of the feedback strength are observed. For the old design, we obtained the following values for the gain: 0.04 V/Pa for the passive mode and maximal 0.14 V/Pa for the active linear mode. The gain in the active linear mode can be adjusted between the passive value of 0.04 V/Pa and 0.14 V/Pa as maximal value by changing the feedback strength. Thus, while the passive mode gain is comparable to standard measurement microphones, the gain can be strongly increased by active operation (up to 5×). In the active non-linear mode, a maximal gain of 0.23 V/Pa is observed for a minimal SPL of 59 dB SPL, which decreases with a slope of 0.015 (V/Pa)/(dB SPL). The black dashed line in the graph indicates the minimal self-noise floor, obtained in the active non-linear mode. We determined a self-noise floor of 59 dB SPL, which increases to 62 dB SPL in the passive mode.

For the new design, the gain increased strongly compared to the old design, i.e. ≈ 3 V/Pa in the passive mode, maximal 25 V/Pa in the active nonlinear mode and maximal 49 V/Pa in the active non-linear mode. This corresponds to an increase in gain by a factor $75 - 200$ due to the redesign of the sensor. The self-noise floor decreased to minimally $18-20$ dB SPL in the active nonlinear mode and $26 - 28$ dB SPL in the passive mode. This corresponds to a decrease by $34 - 41$ dB SPL in comparison to the values of the old design.

Besides the gain and self-noise level, we analyzed the signal distortion introduced by the sensor system. Therefore, a single tone at the respective resonance frequency of the transducer ($f = 3.79$ kHz) was applied and the response measured at the resonance frequency and the higher harmonics. For feedback strengths larger than approximately half the critical value, the second harmonic response occurs, while for lower feedback strengths only the first harmonic is observed. This indicates the onset of (slightly) nonlinear behaviour. The corresponding total harmonic distortion value (THD) increases from 0.2 % at the first occurrence of the second harmonics peak in the frequency response for $a = 0.25$ to roughly 1 % at the highest feedback strength in the active nonlinear mode ($a = 0.5$). For larger feedback strengths, self-excited oscillations are observed due to the system undergoing the Hopf bifurcation (as described above). In this range, multiple higher harmonics occur in the frequency response, yielding an increase of the THD value to 1.36 % for a feedback strength $a = 0.56$ slightly above the critical value ($a = 0.51$).

Finally, to obtain the directionality and to study the damping and influence of the surrounding membrane of the sensor system, numerical simulations of the system were conducted using COMSOL Multiphysics. The simulation setup is shown in fig. 6. Plane wave excitations are used. Geometric non-linearities as well as a deformed

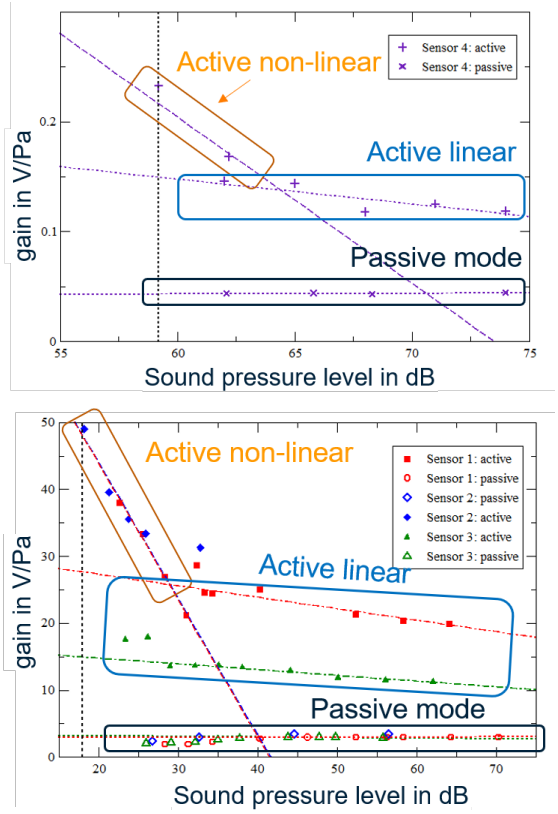

Figure 5: Gain of sensor system for old transducer design (upper graph) and new transducer design (lower graph) obtained from 4 different sensors and for the three operation modes: passive (black box), active linear mode (blue box) and active non-linear mode (orange box). The black dashed line indicates the self-noise floor for the active non-linear mode. The dotted, dash-dotted lines are linear fit to the data points.

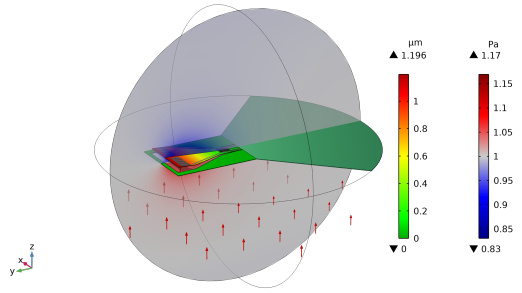

Figure 6: Setup for the simulation: the beam is mounted on one side (fixed constraint). Sound input is given as plane wave in free field (using spherical wave radiation), indicated by the arrows.

mesh are included to avoid mesh distortion for large displacements. To model damping, a mechanical loss factor is used as well as thermoviscous damping, due to the small size of the transducer. Thermoviscous damping yields a decrease of the quality factor from roughly 11000 to $\approx$ 200, an effect known for oscillating microbeams in gas or fluid environment [8] and for sound propagation in/near small structures [9]. Furthermore, the simulations reveal the occurrence of a resonance by the surrounding membrane. Nevertheless, the resonance is clearly separated from the transducer resonance and yields only a negligible sensing signal.

Finally, the simulations were used to study the directionality of the sensor system. Therefore, the angle, describing the direction of the incoming wave in the xz-plane, was varied between $0 - 180\,^\circ$ in $5\,^\circ$ steps. The pressure amplitude was set to 1 Pa. The result is shown is polar plot in fig. 7. As expected for a gradient pressure sensor, the system shows a figure-of-eight directionality. Here, a decrease by ≥ 10 dB compared to the zero-degree input is observed in the range of $75 - 105\,^\circ$.

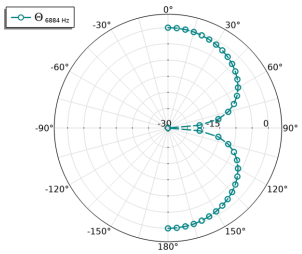

Figure 7: Directionality, shown as polar plot, of the sensor system obtained from Comsol simulations.

Conclusions

Here, we presented the re-design of the transducer of our bio-inspired, adaptable acoustic sensor to improve the acoustic sensing properties. Particularly, the goals were to decrease the resonance frequencies to the range of 0.5−10 kHz to enable the coverage of the audible range. This was achieved by increasing the length of the transducer to 1350 µm or 2600 µm, respectively. The second goal of the re-design was to increase the sensitivity and decrease the self-noise floor of the sensor. Therefore, (i) the surface of the sensor was increased by increasing its width, (ii) its spring constant was decreased by using a more triangular-like shape and (iii) an additional membrane was added increasing the path for the sound wave. All these changes together resulted in an increase of the gain by a factor of 75 (passive mode) to 200 (active nonlinear mode) and a reduction of the self-noise floor by 24 dB (passive mode) to 39 dB (active mode). In summary, the sensor exhibits sensing properties comparable to measurement microphones while additionally offering sensor-based feature extraction like frequency decomposition of the signal and nonlinear transformation, which are important for the performance of sound processing systems. Integration of these functionalities into the sensor may help address current issues regarding performance reduction of speech processing for noisy signals and efficiency of speech processing systems.

Acknowledgment

Funded by the Deutsche Forschungsgemeinschaft (German Research Foundation)-Project-ID 434434223 - SFB 1461 and the Carl-Zeiss-Stiftung in the project 'Memristive Materials for Neuromorphic Engineering (MemWerk)'.

References

[1] Papastratis, I.: Speech Recognition: a review of the different deep learning approaches, https: theaisummer.com (2021).

[2] Araujo, F. A. *et al.*: Role of non-linear data processing on speech recognition task in the framework of reservoir computing, Sci. Rep. 10 (1), 1-11 (2020).

[3] Hudspeth, A.J.: Integrating the active process of hair cells with cochlear function, Nat. Rev. Neurosci. 15 (9), 600-614 (2014).

[4] Lenk, C. *et al.*: Enabling Adaptive and Enhanced Acoustic Sensing Using Nonlinear Dynamics, 2020 IEEE International Symposium on Circuits and Systems (ISCAS), 1-4 (2020).

[5] Rangelow, I. W. *et al.*: Review Article: Active scanning probes: A versatile toolkit for fast imaging and emerging nanofabrication, J. Vac. Sci. Technol. B 35 (6), 06G101 (2017).

[6] Durstewitz, S., Lenk, C., Ziegler, M. : Bio-Inspired Acoustic Sensor with Gain Adaptation Enhancing Dynamic Range and Onset Detection, 2022 IEEE International Symposium on Circuits and Systems (ISCAS) (2022).

[7] Ivanov, T.: Piezoresistive cantilevers with an integrated bimorph actuator, Thesis urn:nbn:de:hebis:34-1153 (2003).

[8] Stauffenberg, J. *et al.*: Determination of the mixing ratio of a flowing gas mixture with self-actuated microcantilevers, J. Sens. Sens. Syst. 9, 71-78 (2020).

[9] Küller, J., Zhyhkar, A., and Beer, D.: Sound Propagation in Microchannels, Fortschritte der Akustik - DAGA 2020 (2020).