It can be stated that SAM mimics normal hearing in a more realistic way than the n-of-m strategy does, even though the modeled SRTs show little differences between SAM and n-of-m. With increasing internal noise (worse simulated cognitive condition), however, SAM outperforms the n-of-m strategy especially in model configurations with fewer auditory nerve cells. While the two strategies deliver about the same amount of place pitch cues, SAM provides more temporal pitch cues, which may well contribute to pitch perception according to the modeled results.

Results, of course, needs to be verified with clinical studies.

ACKNOWLEDGMENT

The underlying research reported in the paper was partially funded by the grant B514-09020 of the Thuringian Ministry of Education, Science and Culture.

REFERENCES

Baumgarte, F. (**1999**). "A Physiological Ear Model for the Emulation of Masking" ORL, **61**, 294–304.

Hamacher, V. (**2004**). *Signalverarbeitungsmodelle des elektrisch stimulierten Gehörs.* Aachener Beiträge zu Digitalen Nachrichtensystemen, Mainz, Aachen, Germany.

Harczos, T., Chilian, A., and Husar, P. (**2011**). "SAM: a novel cochlear implant speech coding strategy based on an active cochlea model" *in prep*..

Hey, M., Hocke, T., Braun, A., Scholz, G., Brademann, G., and Müller-Deile, J. (**2010**). "Erhebung von Normativen Daten für den Oldenburger Satztest bei CI-Patienten" in Proc. 13. Jahrestagung der Deutschen Gesellschaft für Audiologie on CD-ROM.

Jürgens, T., Fredelake, S., Meyer, R., Kollmeier, B., and Brand, T. (**2010**). "Challenging the Speech Intelligibility Index: Macroscopic vs. Microscopic Prediction of Sentence Recognition in Normal and Hearing-impaired Listeners," in Proc. Interspeech, Makuhari, Japan, 2478–2481.

Shepherd, R.K., Roberts, L.A., and Paolini, A.G. (**2004**). "Long-term sensorineural hearing loss induces functional changes in the rat auditory nerve" Eur. J. Neurosci., **20**, 3131–3140.

Sumner, C. J., Lopez-Poveda, E. A., O'Mard, L. P., and Meddis, R. (**2002**). "A revised model of the inner-hair cell and auditory nerve complex" J. Acoust. Soc. Am., **111**, 2178–2188.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies" J. Acoust. Soc. Am., **117**, 3126–3138.

Wagener, K., Brand, T., and Kollmeier, B. (**1999**). "Entwicklung und Evaluation eines Satztests für die deutsche Sprache" Zeitschrift für Audiologie, **38**, 4–95.

Horizontal-plane localization with bilateral cochlear implants using the SAM strategy

TAMAS HARCZOS[1,2], ANJA CHILIAN[1,3], ANDRAS KATAI[1]

[1] *Fraunhofer Institute for Digital Media Technology IDMT, Ilmenau, Germany*

[2] *Institute for Media Technology, Faculty of Electrical Engineering and Information Technology, Ilmenau University of Technology, Germany*

[3] *Institute of Biomedical Engineering and Informatics, Faculty of Computer Science and Automation, Ilmenau University of Technology, Germany*

Sound source localization capability of cochlear implant (CI) users has been a popular research topic over the past few years, because it has both social and safety implications. While it is widely accepted that unilateral implantation does not provide enough information for this task, conditions, algorithms and their parameterization for the best performance in the binaural case are still in the focus of the research.

On ISAAR 2009, we presented a simulation study revealing the theoretical limits of localization performance using the widespread ACE strategy. We also gave an example of how left-right speech processor asynchrony may influence the perceived direction.

In the present paper we give an outline of a novel, auditory model based CI speech processing strategy called SAM. Furthermore, using the framework from the previous study, we show how localization performance increases when using SAM instead of ACE. We present detailed comparisons to show how factors like pulse rate, signal to noise ratio, reverberation, etc. affect horizontal-plane localization. Finally, we give a simple explanation, why, unlike other strategies, spatial perception with SAM is robust against device asynchrony.

INTRODUCTION

Over the past decade, cochlear implants (CIs) have become a widely accepted alternative for treatment of people with severe to profound hearing loss. While bilateral cochlear implantation (BI) is offered to a growing number of individuals, not all BI-users are 100% satisfied.

One possible cause for the dissatisfaction is the missing ability to robustly localize sound sources. The trend is to use $<$1K/s channel stimulation rate (CSR) and $\leq$9K/s total stimulation rate (TSR) with n-of-m strategies like ACE, which, in fact, allows for only very limited localization performance based on temporal cues, as shown e. g. in Harczos *et al.* (2010).

Proceedings of ISAAR 2011: Speech perception and auditory disorders. 3rd International Symposium on Auditory and Audiological Research. August 2011, Nyborg, Denmark. Edited by T. Dau, M. L. Jepsen, J. Cristensen-Dalsgaard, and T. Poulsen. ISBN 87-990013-3-0. EAN 9788799001330. The Danavox Jubilee Foundation, 2012.

Furthermore, the most common (and generally only) aim of CI fitting is to yield better speech perception rates. Still, most BI-recipients can localize sound sources to some extent (Grantham *et al.*, 2007), but only few can localize well (Seeber and Fastl, 2004).

The primary goal of this study is to evaluate horizontal-plane localization with the SAM strategy (see under), taking only interaural time differences (ITDs) into account. These have long been deemed unusable by CI recipients, but some studies have proven otherwise. For one of the most recent ones see Drapal and Marsalek (2010). Furthermore, factors being responsible for good or bad localization ability with SAM are searched for, and a performance-comparison between SAM and ACE is given.

METHODS

The SAM strategy

SAM (Stimulation based on Auditory Modeling) is a novel CI speech processing strategy (Harczos *et al.,* 2011), incorporating active cochlear filtering (basilar membrane and outer hair cells) along with the mechanoelectrical transduction of the inner hair cells. An overview of SAM is shown in Fig. 1. Through its functional design several psychoacoustic phenomena like compression, adaptation and realistic cochlear delays are accounted for inherently. The coder, unlike in common strategies, is not restricted by a pre-defined channel stimulation rate and it activates stimulating electrodes in a stochastic manner.

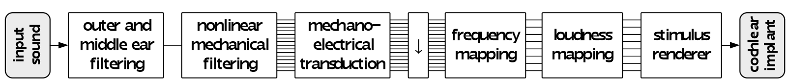

Fig. 1: SAM system overview and signal path.

Design of the experiments

The design of this simulation study was borrowed from Harczos *et al.* (2010), which can be looked up for details. An overview of the experiment design is shown in Fig. 2 and a short description of each step is given below.

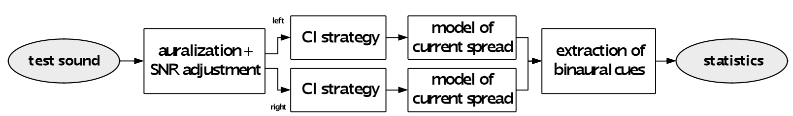

Fig. 2: Experiment design.

Two distinct test sounds are used: a TIMIT (see Zue *et al.*, 1990) speech signal snippet (referred to as *Timit* throughout this paper), and noise bursts (*pulses*) like in Seeber and Fastl (2004).

Auralization is done via the fast image method described by McGovern (2009). The distance between the sound source and the microphones is fixed at 3.0 m. Two wall-reverberation settings are tested: r=0.001 (anechoic) and r=0.7 (office-like).

Noise type and noise ratio are varied in the *SNR* (signal-to-noise ratio) *adjustment step*. Former can be white noise (*WN*) or babble noise recorded at a train station (*TS*), and latter can be 5 dB, 20 dB or clean (no noise at all).

The used CI strategy can be ACE or SAM. Tested channel stimulation rates of ACE are 720, 900, 1200, 1800, 2400, 3200 and 3500/s, and the number of selected spectral peaks (*N*) is varied between 1 and 8. SAM's selectable total stimulation rate ranges from 1440/s to 28K/s, which is the same range as with ACE.

The current spread is modeled by an exponential decay function. Settings for current spread extent (λ) are 0 (no current spread), 0.5 mm and 2.0 mm.

The localization itself happens through generalized cross-correlation (GCC), using 30 ms window size, without overlapping. Mean and standard deviation of the localized degrees per window are calculated for each test file and used for further statistics. For each sound source position, the difference between the measured mean direction and the real direction can be evaluated as the mean error. Similarly, the magnitude of the standard deviation over a test file reveals, whether the localization has high certainty or high ambiguity, see Fig. 3.

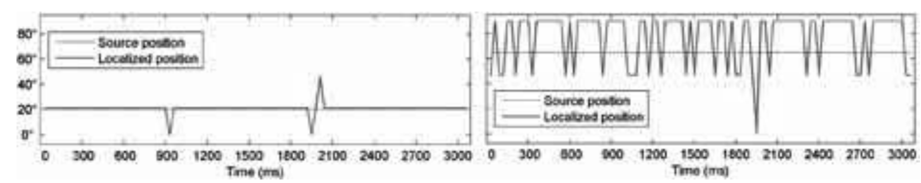

Fig. 3: Speaker at 20° (**left**) and 65° (**right**), anechoic room, no noise. GCC-localization delivers in both cases the right mean for the direction. Left: high certainty (low standard deviation), right: high ambiguity.

Electrodograms

In this study, the output of the cochlear implant speech processing algorithms are always stored in the same matrix format, where the y dimension represents the CI electrodes, and the x dimension provides the possible time slots with the given total stimulation rate. (The matrix storage format ignores pulse-specific information like pulse width, phase gap, etc., but they are assumed to be identical among the strategies to be compared.)

An ACE vs. SAM comparison is presented in Fig. 4. In the figure, it can be seen that most electrode channels of ACE contribute only little to the ITD-based localization. In contrast to that, all channels of SAM provide ITD information, which can be used for localization.

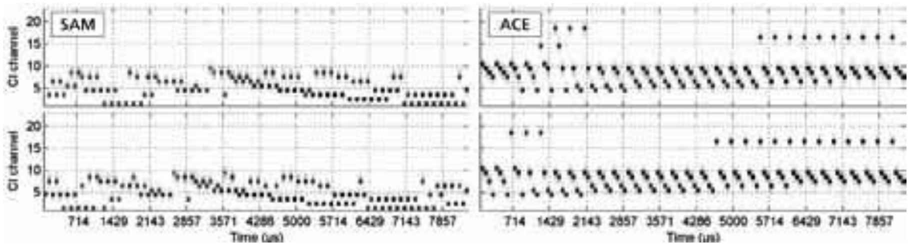

Fig. 4: Electrodograms showing a short snippet of the stimulation patterns of left and right CI for a speech signal (*Timit*) source located at 65°. *TSR*=14K/s, λ=0.5 mm. (ACE: *N*=4, *CSR*=3500/s.).

RESULTS

This chapter presents the most important outcomes of the study. The results shown in the following section are based on the assumption that the two CI processors are perfectly synchronized. Estimations of the effects of missing synchronization to the localization performance are presented in the second section.

Synchronized processors

One characteristic of the localization performance is the error between the real and the localized direction of the sound source, as a function of the real direction. Fig. 5 presents the results of such a test for both strategies, so that they can be compared easily. In the first case (top row of Fig. 5), a low pulse rate (*TSR*=4800/s) scenario is tested, while in the second case (bottom row of Fig. 5) high total pulse rate (*TSR*=14000/s) is employed. Note that in all presented comparisons the total pulse rate is the same for both CI strategies, whereas, for ACE, $TSR=CSR\cdot N$ holds.

Results with the common *TSR*=7200/s are not presented here, but it can be stated that those are better than in the case of *TSR*=4800/s and worse than with *TSR*=14000/s.

Please note that the ACE-calculations are done with the untypically low *N*=4 setting. While this setting is valid and possible, in the practice –known to the authors–, values of 8 or 10 are used. Those are typically combined with channel stimulation rates of 720/s or 900/s. All four combinations of these values lead to the complete loss of ITD-information, i. e., zero ITD-based localization ability.

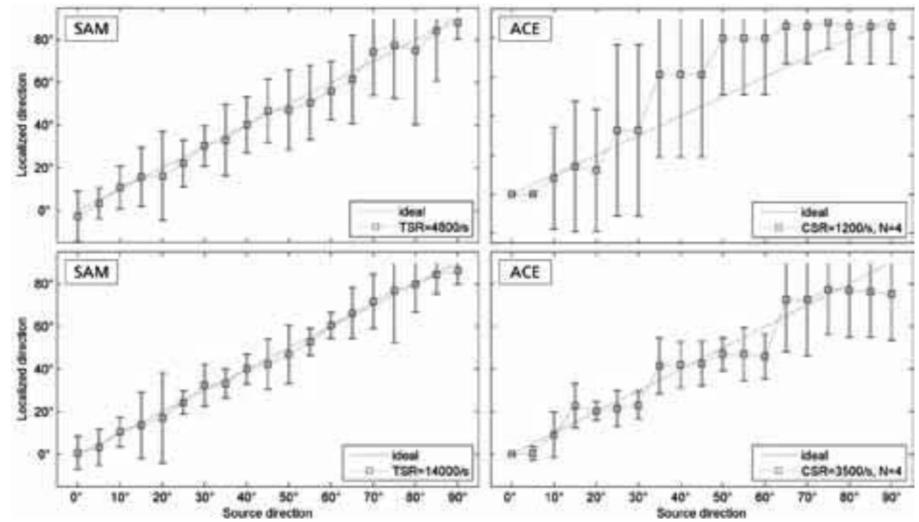

Fig. 5: Mean localized angles and mean of standard deviations for a speech signal (*Timit*) located at 0°, 5°, 10°, ..., 90°. λ=0.5 mm. **Top:** *TSR*=4800/s (ACE: *N*=4, *CSR*=1200/s). **Bottom:** *TSR*=14K/s (ACE: *N*=4, *CSR*=3500/s). Perfect synchronization of left and right devices is assumed. Note the large error and the step-like behaviour with ACE.

The effect of changing the pulse rate and that of various current spread settings is presented in Fig. 6. It can be stated that the extent of current spread (at least the tested settings) does not substantially affect localization ability. Merely the localization certainty decreases. An increase in the pulse rate, on the other hand, tends to reduce localization errors.

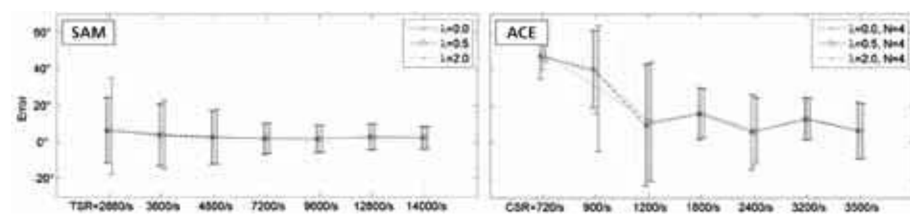

Fig. 6: Mean errors and mean of standard deviations for speech signal (*Timit*) averaged over angles between 5 and 90°, plotted as a function of stimulation rate, for various current spread settings.

Fig. 7 presents the outcomes of the factor analysis regarding reverberation level, SNR, noise type and signal type. A characteristic difference is that –in terms of localization error– SAM always performs better than ACE. As expected, adding background noise or reverberation increases the error.

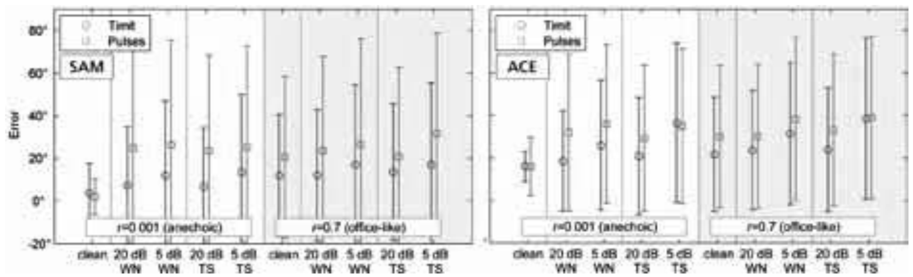

Fig. 7: Mean errors and mean of standard deviations averaged over angles between 5 and 90°, plotted as a function of reverberation level, SNR, noise type and signal type. *TSR*=7200/s (ACE: *CSR*=1800/s, *N*=4), λ=0.5 mm.

Asynchronous processors

Because of the block-by-block processing in ACE, the lag between the (unsynchronized) left and right CI may get into the millisecond range. (Fig. 8 shows a special case with untypically high *TSR*. Standard settings can cause more lag and worse localization performance.) This leads to ITD-based localization errors. Given that the internal clock speed of the two CI devices differ only a little, the direction of a fixed sound source may be perceived as slowly and periodically changing (see Fig. 8).

The filter bank in SAM is based on the time domain simulation of the auditory system and processes the input sound on a sample-by-sample basis. This method yields a maximum lag-range of $\pm(TSR^{-1}/2)$ seconds (shaded in Fig. 8), which leads to less localization error in unsynchronized systems.

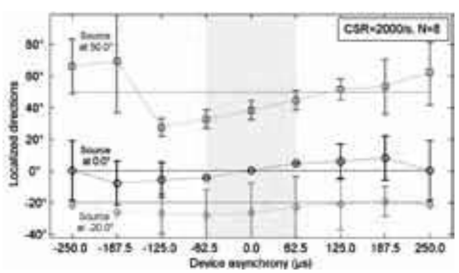

Fig. 8: Mean localized directions and mean of standard deviations for selected source directions using ACE, plotted as a function of left-right device asynchrony. *CSR*=2000/s, *N*=8, λ=0.5 mm. Shaded range highlights possible asynchrony with SAM.

SUMMARY

It has been shown that horizontal-plane localization of sound sources is working well with SAM over a wide range of SAM's possible parameterization. The amount of cues for ITD-based localization preserved by SAM clearly exceeds that preserved by ACE, when compared using the same total pulse rate. Noise and reverberation seem to have less negative impact on the localization performance with SAM.

It has been illustrated, furthermore, that in the real-world scenario, where the CI processors are not synchronized, time-domain filtering of SAM makes the lag between the left and right devices less critical.

ACKNOWLEDGMENT

The underlying research reported in this paper was funded by the grant B514-09020 of the Thuringian Ministry of Education, Science and Culture.

REFERENCES

Drapal, M. and Marsalek, P. (**2010**). "Stochastic Model Shows How Cochlear Implants Process Azimuth in Real Auditory Space," Chinese Journal of Physiology, **53**, 439–446.

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Labadie, R. F., and Haynes, D. S. (**2007**). "Horizontal-Plane Localization of Noise and Speech Signals by Post-lingually Deafened Adults Fitted With Bilateral Cochlear Implants," Ear and Hearing, **28**, 524–541.

Harczos, T., Werner, S., Szepannek, G., and Brandenburg, K. (**2010**). "Evaluation of cues for horizontal-plane localization with bilateral cochlear implants," in Proceedings of ISAAR 2009: *Binaural Processing and Spatial Hearing*. 2nd International Symposium on Auditory and Audiological Research. Elsinore, Denmark. Edited by J. M. Buchholz, T. Dau, J. Christensen-Dalsgaard, and T. Poulsen. ISBN: 87-990013-2-2. (The Danavox Jubilee Foundation, Copenhagen), 37–46.

Harczos, T., Chilian, A., and Husar, P. (**2011**). "SAM: a novel cochlear implant speech coding strategy based on an active cochlea model," *to be published.*

McGovern, S. G. (2009). "Fast image method for impulse response calculations of box-shaped rooms," Applied Acoustics, **70**, 182–189.

Seeber, B. U. and Fastl, H. (**2004**). "Localization cues with bilateral cochlear implants," J. Acoust. Soc. Am., **123**, 1030–1042.

Zue, V., Seneff, S., and Glass, J. (**1990**). "Speech database development at MIT: TIMIT and beyond," Speech Communication, **9**, 351-356.