Evaluation of Spatial Approaches for Improvement of Patient Speech Acquisition at Magnetic Resonance Tomography Systems

M.Friedrich[1], K. Porzelt[2], M. Bechtold[2], P.-C. Eccardt[2]
[1] University of Erlangen - Nuernberg, Germany, Email: m.friedrich@gmail.com
[2] Siemens AG, Erlangen and Munich, Germany

Abstract

During operation of magnetic resonance (MR) tomography systems, noise is produced with significant sound pressure levels. This noise may influence patient-to-operator communication as well as examinations that require synchronous recording of patient speech. Signal processing algorithms for removal of MR and other noise sources from patient speech will benefit from the highest possible input signal-to-noise ratio (SNR). In this contribution spatial approaches for improvement of the input SNR are discussed, experimental results and conclusions about the use of directional microphones and multiple microphone approaches in the MR environment are presented. Only limited improvements for for support of further signal processing could be obtained with the evaluated methods.

Introduction

Magnetic Resonance (MR) Tomography is a medical imaging method for the diagnosis of organic tissue. During examination, the patient is placed inside a large tube-like structure on a movable patient table. The examined region of the body is placed in the middle of the tube (which is also called bore). The operator of the system is situated in an adjacent room and supervises the patient and the MR device. Communication between operator and patient is possible through an intercom system, but is influenced by noise that is generated by the MR system during examinations. Signal processing can be employed to reduce the MR noise in the communication channel and improve the distorted speech signal. Though, patient speech should be acquired with the highest possible speech-to-noise ratio, which represents the input SNR for further signal processing.
Methods that employ spatial differences of the speech and noise sound fields, such as directional microphones or multiple microphone approaches, are not expected to provide the required speech enhancement for patient-to-operator communication. However, they might be useful to improve the input SNR for further signal processing stages. In this contribution, the use of some spatial approaches for improvement of the input SNR is discussed and evaluated by experiments on a MR system.

Sound fields and sources

MR systems are installed in a special room that provides a shielded enclosure for magnetic, radio frequency and acoustic fields. Depending on the actual building construction, the MR system may allocate significant part of this examination room. Walls of the enclosure may have very small absorption factors, and the MR system itself has a large reflecting surface as well. In such a case, the MR system is situated in a highly reverberant environment. Measured reverberation time on a real MR installation was $T_{60} = 0.7s$, the calculated critical distance for omnidirectional source was $r_H = 0.65\,\mathrm{m}$. Considering division of the room by the MR system, the lower frequency for diffuse field description can be as high as $f = 350 \ldots 400\,\mathrm{Hz}$ [1].

Sources of interest in this environment are patient speech and MR noise. During examination, the patient's head can be situated at any position on the patient table. The speech sound field can be described by a point source in the room with a weak directivity factor of $\gamma_s \approx 1.44$ [2]. Inside the resulting critical distance from the head, direct sound from the patient dominates, and the patient represents a local sound source. The energy density of the emitted sound decreases approximately proportional to $1/r^2$ with distance r.
On the contrary, the MR system emits the interfering noise from a large vibrating surface. Sound emission from an extended surface leads to emission of planar sound waves into the room with a frequency-dependent angle [1]. The $1/r^2$-dependency of sound energy density on distance does not hold here, instead an approximately constant mean sound pressure level has been measured for positions in front of the bore, independent on the distance from the system. Reflections of the emitted sound at room boundaries often takes place within the near field of the source. The resulting noise sound field in the room is characterized by strong acoustical modes.

Microphone positioning

Microphones for patient speech acquisition are ideally placed in close proximity to the patient's mouth, to minimize the effect of sound energy reduction with distance. This can be realized for example by using a headset with boom microphone, which allows placing the microphone directly in front of the patient's mouth. Though, in practice a fixed mounting of the microphone on the MR system might be preferred over the variable placement, for other than acoustical reasons. Furthermore, one might wish to avoid microphone placement inside the bore, in order to prevent potential problems due to strong magnetic and radio frequency fields, or possible interference of the microphones with the measurement system. In this

case, reasonable microphone positions for patient speech acquisition are located around the ends of the bore, as shown in Fig. 1 (left).
During the experiments, both microphone placements were considered and are referred to as *headphone position* and *fixed microphone position*, respectively.

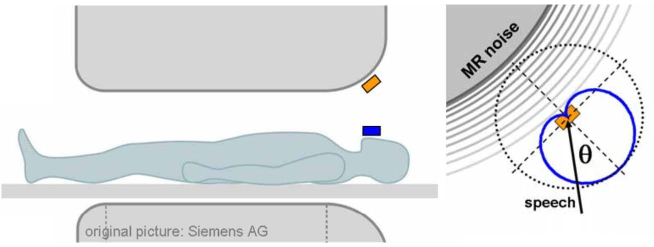

Figure 1: Left: cross sectional view of MR system with patient on movable table in bore. Evaluated microphone positions are marked: headphone position (blue) and fixed mounting position at bore end (orange). Right: directional microphone at fixed mounting position with idealized sound sources

Experimental Setup

For evaluation of spatial approaches, experiments have been conducted on a Siemens *Magnetom Avanto* 1.5T system, installed in an environment as described above. Electret microphones with different directional characteristics have been placed in headphone and fixed mounting positions on the system. Patient speech was modelled by a magnetic field compatible loudspeaker, which emits broadband stationary noise. Spectral characteristic of this emission was adapted to represent human speech according to ANSI S3.5-1997 in the frequency range of interest. The MR system was excited by stationary broadband noise as well. This method provides a known constant output level for speech, and allows long-time estimates of overall levels and spectra. Results of the experiments allow general conclusions about the use of spatial approaches. Speech and MR noise have been evaluated consecutively in all measurements, without changing the acoustical environment.

Directional microphones

Concept and gain assumptions

Directional microphones can improve the input SNR for further signal processing by reducing the energy of the interference source in the microphone signal, either by placing a local interference in directions with lower sensitivity, or by utilizing the diffuse field gain. For headphone positions of the microphone, reverberated sound and noise from its rear hemisphere are reduced. For fixed mounting, the main axis of the microphone is pointing towards patient positions in front of the bore, placing the MR surface in directions with reduced sensitivity (see Fig. 1 (right)).
An analytical estimate of the achievable gain is only possible using idealized descriptions of the sound fields. Considering free field conditions, i.e. no reflections of noise and speech signals are taken into account, noise reduction can be estimated by integration over the back hemisphere of the sensitivity function $s(\theta)$. The result is also specified as the microphone parameter *REB* (random energy-back) [3]. Speech is captured from front directions. For directional microphones, speech with incidence angle $\theta \neq 0°$ is attenuated, according to the reduced sensitivity $s(\theta)$ of the directional microphone. The overall gain $A_d(\theta)$ is calculated as

$$A_d(\theta) = 10\log\left(\frac{REB_o}{REB_d}\right) + 10\log\left(\frac{s_d(\theta)}{s_o(\theta)}\right)^2 \text{[dB]}. \quad (1)$$

Indices d and o in *eq.1* indicate values for directional and omnidirectional microphones, respectively. SNR gain estimates for directional microphones are $4.8 \ldots 10.8$ dB for cardioid microphones, and $5.7 \ldots 14.4$ dB for supercardioid microphones, depending on the patient position.
However, in realistic scenarios noise and speech can be present from all directions, due to reflections in the room and noise emission from other parts of the MR surface. An analytical prediction of the SNR gain is then not possible due to the complexity and variance of the acoustical scenario. Conclusions are therefore based on experimental results.
For discussion of input SNR improvement with directional microphones, overall sound pressure levels (SPL) of speech and noise were evaluated. Values were obtained as overall level of unweighted one-third octave bands, with band center frequencies of $315 \ldots 5000$ Hz, averaged over t=8 s.

Results for fixed mounting position

Microphones with different directional characteristics were placed at fixed positions at each side of the MR system. The patient speaker, situated on the patient table, was moved to several possible positions of the patient's head. For each of the microphones, overall SPL values were evaluated for MR noise excitation, and for patient speech as function of position. For direct comparison of the microphones, speech-to-noise ratios were calculated for each of the positions, which also removed any bias from possible calibration errors.
An example result is shown in Fig. 2. Results are given as relative input SNR values, which allows direct comparison of the microphones, and evaluation of the SNR loss for remote patient positions.
An input SNR improvement by directional microphones can only be observed at positions directly in front of the microphones. At all other positions, especially for patient positions inside the bore, the omnidirectional microphone clearly provides the highest input SNR. This is due to a strong reduction of the captured patient speech level for positions beside the main direction of the microphones. For patient positions inside the bore, no direct sound path is present between the patient's mouth and microphones. Patient speech then reaches the microphones only as reflected sound with strongly reduced magnitude. On the other hand, results for MR noise reduction with directional microphones vary in the range of $-3 \ldots 6$ dB, depending on the emitted noise sound field for different excitation modes of the system. Results for different excitation of the MR are given in Table 1.

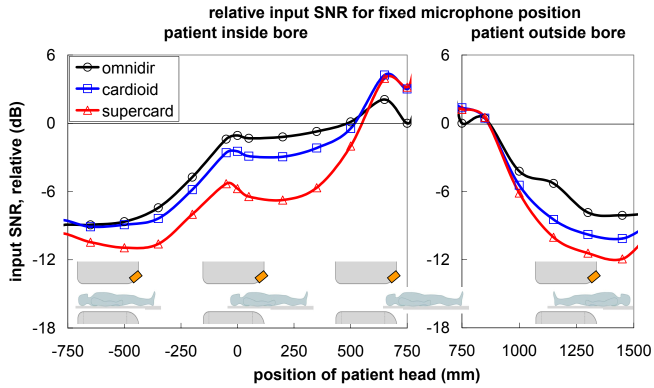

Figure 2: Input SNR comparison for directional microphones at fixed mounting position. Values are provided as relative input SNR, normalized to the result of omnidirectional microphone at position p=750 mm.

MR noise reduction	directional characteristic	
	cardioid	supercardioid
excitation 1	-3.5dB	-3.1dB
excitation 2	-5.1dB	-5.7dB
excitation 3	2.7dB	-0.3dB

Table 1: Reduction of MR noise with directional microphones: relative levels, compared to SPL of omnidirectional microphone, for different MR excitation. A positive value represents increased noise level in the directional microphone's signal.

Results for headphone position

For evaluation of directional microphones at headphone position, microphones with different directional characteristics were placed above the patient speaker. At several patient positions, overall levels for MR noise and patient speech have been evaluated, moving speaker and microphones together with the patient table. Measurement results are shown in Fig. 3.

At patient positions inside the bore, no gain can be determined for directional microphones. Variation of the input SNR is observed due to modal effects inside the bore, where both noise and speech are highly reverberated. At positions outside the bore, the directional microphone reduces capturing of noise and reverberated speech from the room. The observed SNR gain of 1.1 . . . 2.8 dB is much smaller than the diffuse field gain because of the non-uniform spatial distribution of the MR noise field. However, reduction of reverberated speech content might also be beneficial for intelligibility of the speech after noise reduction.

Multiple microphone approaches

Approaches for input SNR enhancement

For discussion of multiple microphone approaches, we consider fixed mounting of more than one microphone at the circumference of one end of the bore. In this case, speech and noise signals will be present in all microphones.

Beamforming algorithms could be used in such a configuration, but are not further investigated here. Results are expected to be comparable to the directional microphone results for fixed mounting position. With strongly

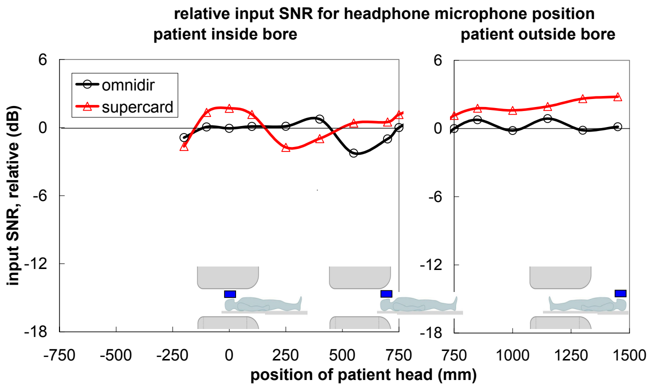

Figure 3: Input SNR comparison for directional microphones at headphone position. Values are provided as relative input SNR, normalized to the result of omnidirectional microphone at position p=750 mm.

varying MR noise reduction and with attenuation of patient speech level for positions aside the main sensitivity direction of the beam, no distinct SNR improvement is expected for the majority of patient positions.

Other approaches are known for multiple microphones, which exploit spatial differences of sound sources by a postfilter. For example, Martin [2] presented a two-microphone approach with adaptive postfilter for noise reduction and speech enhancement in noisy reverberant environments. The performance of the algorithm is strongly dependent on correlation properties of the microphone signals for different sources.

For unknown sound fields, spatial correlation properties can be explored by analysis of the coherence function of two microphones with known distance. Experiments have been conducted on the MR system, using microphones with distance $d = 50\,\text{cm}$.

Sound field analysis using the coherence function

The magnitude squared coherence of two microphones signals x_1 and x_2 is defined in the frequency domain as

$$|\Gamma_{x1x2}(f)|^2 = \frac{|S_{x1x2}(f)|^2}{S_{x1x1}(f)\, S_{x2x2}(f)} \tag{2}$$

The cross-power spectral density $S_{x1x2}(f)$ and the auto-power spectral densities $S_{x1x1}(f)$ and $S_{x2x2}(f)$ are estimated from Discrete Fourier Transforms of signal blocks of finite length L, e.g. by applying Welch's method for periodogram estimation.

The coherence function provides a frequency domain description of spatial correlation properties within a sound field. Values are in the range of $|\Gamma_{x1x2}|^2 = 0$ for completely uncorrelated signals, and $|\Gamma_{x1x2}|^2 = 1$ for ideally correlated microphone signals x_1 and x_2 . Correlated signal content in the microphone signals can be related by a linear system. A more detailed discussion of the characterization of sound fields using the coherence function can be found e.g. in [4] and [5].

Estimates of coherence for sound fields of MR noise and patient speaker were obtained on a MR installation. Since both the noise and speech source are stationary signals in the experiment, estimation can be done over very

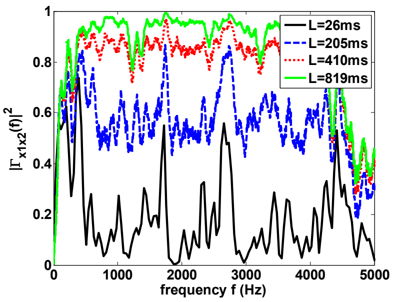

Figure 4: Magnitude squared coherence of MR noise for different estimation block lengths L.

long periods, providing results for extremely large L and with small variance.

In reverberant environments, results are dependent on the estimation block length L. For the chosen excitation signal, only signal content within the actual blocks at a time instant is considered linearly related, while signal content outside these frames is described as uncorrelated. Measured coherence results are biased towards zero and can only provide characterization of the sound fields for a defined analysis length L. However, this allows conclusions about signal processing approaches that relate the input signals by linear systems of length L.

Results for MR noise evaluation with different block lengths L are shown in Fig. 4. Since MR noise is emitted from an extended surface, and multipath reflections are present in the highly reverberant environment, a coherent description of the noise is only obtained for very large estimation block lengths.

For patient speech signals, coherence values are close to one only for patient positions in front of the microphones. Measured coherence is strongly reduced for other positions, especially inside the bore, where reverberated signal content dominates over direct sound in recorded speech signals.

Conclusions for multiple microphone approaches with postfilter

The two-microphone-algorithm with postfilter in [2] reduces uncorrelated (noise) signal contents, and estimates correlated (speech) signal content. Achievable noise reduction and minimum clean speech estimation error can be evaluated as a function of coherence [2]. Due to the dependency of coherence results on the block length L, conclusions for multiple microphone approaches are given for different estimation block lengths, referring to algorithms that relate the input signals by a linear system of same length L.

Measurement results for noise and speech signals are given in Fig. 5. For small block lengths (e.g. L=26 ms), MR noise signals are mainly uncorrelated, and attenuation of the noise signal is achieved. However, due to the low coherence of the speech signal, relevant damping of speech is expected as well, which limits the overall SNR gain. Furthermore, speech distortion is likely due to the frequency dependent gain of the algorithm, which is a result of large coherence differences over frequency.

For larger block lengths L, coherence values of both speech and noise are increasing, as shown in Fig. 5 for

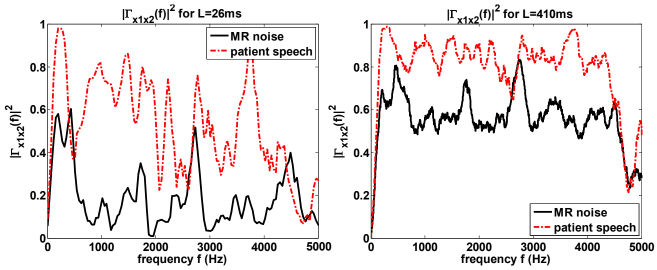

Figure 5: Magnitude squared coherence of MR noise and patient speech (patient position inside bore). Results are obtained for small block length L=26ms (left) and large block length L=410ms (right).

L=410ms. However, for patient positions with dominating reverberation content in the speech, coherence values are lower than one even for very long block lengths, which limits the achievable speech enhancement. Noise attenuation by the adaptive postfilter is strongly reduced, due to the distinct increase of coherence for MR noise. Therefore, only small input SNR improvement is expected for large filter dimensions of the algorithm as well.

Due to these results, and for the expected input SNR on MR systems during patient examinations, the discussed multiple microphone approach with postfilter is not considered to be applicable for input SNR improvement in the MR case.

Summary

The use of spatial approaches for input SNR improvement of patient speech acquisition at MR systems has been discussed and evaluated by experiments on a MR installation. Results for directional microphones and multiple microphone approaches do not indicate distinct improvements for all possible patient positions. For headphone mounting, directional microphones might provide a small SNR gain and reverberation reduction at positions outside the bore. At evaluated fixed mounting positions however, neither directional microphones, nor the discussed multiple microphone approaches are expected to be beneficial. Considerable speech-to-noise improvement can not be obtained for the support of other signal processing approaches that can improve the quality of the distorted patient speech.

References

[1] G. Müller and M. Möser, Taschenbuch der technischen Akustik, Springer, 2004.

[2] R. Martin, Freisprecheinrichtungen mit mehrkanaliger Echokompensation und Störgeräuschreduktion, Fakultät für Elektrotechnik, RWTH Aachen, 1995.

[3] S. Weinzierl, Handbuch der Audiotechnik, Springer, 2008.

[4] R. Martin, Small Microphone Arrays with Postfilters for Noise and Acoustic Echo Reduction, in Microphone arrays: signal processing, techniques and applications, Springer, 2001, pp. 255-279.

[5] G.W. Elko, Spatial coherence functions for differential microphones in isotropic noise fields, in Microphone arrays: signal processing, techniques and applications, Springer, 2001, pp. 61-86.