Proceedings 1999 IEEE Intl. Symposium on Computational Intelligence in Robotics and Automation
CIRA ‘99, pp. 274 - 281, Monterey, CA, USA, November 8-9, 1999

Acoustic-based room discrimination for the navigation of autonomous mobile robots

Guido Tesch, Uwe R. Zimmer

GMD - Japan Research Laboratory*
AIM Building, 8F
3-8-1 Asano, Kokurakita-ku
Kitakyushu-city, 802-0001, Japan

http://www.gmd.gr.jp/
guido.tesch@gmd.gr.jp uwe.zimmer@gmd.gr.jp

This work investigates the possibility of using audible broadband acoustic signals for the navigation of autonomous mobile robots in unprepared, unknown, silent, static, indoor environments. To this day, audible sound signals are not used for the navigation of autonomous mobile robots. Nonetheless, they are the second best information source for the blind (next to tactile signals) for navigating through indoor environments. Descriptive psychological experiments show that humans have astounding abilities in differentiating their relative position to walls and sensing the existence of even small objects like metal discs not bigger than a plate, solely based on acoustic information.

As there exists no theory on the practical limits of discriminating spaces based on acoustics, an experimental setup is investigated that enables conclusions about the applicability of acoustics for navigation purposes. A first, straight forward approach to a representation and metric for the acoustics of the surrounding is described and tested within the experimental setup.

Keywords: **signal processing, broadband acoustic signals, acoustics, mobile robots, autonomy, exploration, navigation**

* *The GMD - Japan Research Laboratory is a joint venture between GMD - German National Research Center for Information Technology (http://www.gmd.de/) and the City of Kitakyushu.*

1. Introduction

When a blind person wants to explore an unknown room in detail, the sense of touch is his first choice as an information source. By using hands and feet or a stick, blind people often get a good idea of geometry and furnishings of a room, enabling them to solve everyday navigation tasks. In autonomous mobile robotics, this level of exploration is typically done by using distance measurements based on ultrasonic sensors, laser range finders, or vision.

In mobile robotics, *large scale navigation* dealing with several rooms, hallways and so forth, is mostly based on some map built up using the detailed distance measurements described above. However, depending on the task to be accomplished, the blind often choose not to rely on this information source for large scale navigation, but to use *acoustics*, thereby taking a different view on the surrounding. This view can be described as less based on 3-dimensional distances, but more based on topological auditory features of the present position in space.

The idea of this article is to suggest the usage of acoustics as a complementary information source for large scale navigation in mobile robotics. More precisely:

The hypothesis of this work is that broadband audible acoustic signals can be used to distinguish between positions of a mobile robot in a silent, static, indoor environment, so that this information can be used for navigation tasks such as finding a certain room or area in a building.

2. Background and approach

Until today, audible acoustics are not used for discrimination of rooms or navigation of mobile robots. Roy et al mention in [17] the use of audible acoustics for navigation in mobile robotics, but only to distinguish between floor types to estimate the dead-reckoning error. Presently, there exists no report on any artificial system capable of discriminating rooms based on acoustics. In general, there is a lack of theory about the information content of broadband acoustic signals for room discrimination. Classical room acoustics ([5], [12], [19]) does provide some relevant factors such as reverberation time and clarity or, on a more general level, the impulse response. Still, it fails to present a metric by which these factors could be compared. Speech recognition research provides methods for quantifying differences of acoustic signals ([1], [13]), but they are highly specialized on speech and not applicable to room acoustics. Other research areas such as ‘Auditory Scene Analysis’ ([3], [4], [6], [7]) or ‘Virtual Acoustics’ ([2]) do provide interesting methods for the analysis of broadband signals of room acoustics, but also do not address the aspect of *discrimination* of acoustic signals. For the psychological research area, the corresponding general picture is summarized in [20], including the statement that at present, no empirical results are available on the question, exactly which features of acoustic signals enable the discrimination of surroundings.

On the other hand, there do exist a number of descriptive publications which show that humans have astounding abilities in differentiating their relative position to walls and sensing the existence of even small objects like metal discs not bigger than a plate, solely based on acoustic information. ([16], [21]) These results, together with everyday experience from the blind, suggest that in general, it is possible to discriminate between rooms as mentioned in the above hypothesis. The questions are

- how to decide if a developed artificial system confirms the hypothesis given, and
- which representation model of the acoustic information, and which metric over the representations provide a sufficient basis for room discrimination.

The approach taken to answer these questions is experimental by nature. In the next section, a suitable real world experiment is developed which will serve as a test for a system according to the given hypothesis. Section 4 will introduce a first attempt towards a representation and metric for discrimination of room acoustics, which is then tested according to the given experiment.

3. Measuring performance: An experimental test

Consider the following experimental setup as a test to verify, whether or not a system confirms the given hypothesis. A major aspect while designing this experiment, is to keep it close to real world applications. As there is currently no theory on which to rely on when estimating the quality of results, this ‘real world view’ needs to be taken in order to develop a reasonable test condition. This aspect will be referred to repeatedly in the following descriptions.

The environment

Figure 1 shows a typical office environment, with rooms of different size and shape but similar furnishings, an empty hallway and an empty staircase. The environment is kept static throughout the experiment. Only the shaded area is used. All unused doors as well as the automatic door between the hallway and the staircase are kept closed. There exist no sound sources, the background noise level is 47 dB and is independent of position. No further preparation of the environment is done according to the experiment, so real world conditions are met.

The mobile robot

A B14 robot from RWI Inc. is used (we call it Lisa), equipped with a double Pentium Linux computer (200MHz), 64 MB RAM, a standard 16-bit sound card, and an omnidirectional microphone (Sony ECM-T140) mounted on top of the robot. No specialized sound system is used for the experiment. Three fans and a hard drive inside Lisa produce a constant noise level of about 58 dB at the position of the microphone. This rather high noise level represents a major factor in the design of the experiment, as it represents the fact that no real world application for acoustic systems can guarantee a noise free environment. The above mentioned noise level is considered a ‘worst case’ situation, and any system showing acceptable results under this condition is expected to work sufficiently when lower noise levels are given.

By employing its laser range finder, the B14 robot is capable of exploring the test environment autonomously. But, as the focus of this paper is on the processing of sound signals, Lisa’s motion is controlled by an operator using a joystick. The position of the operator remains constant with respect to the robot throughout the experiment. This setup is considered realistic with respect to possible applications like exploring the test environment autonomously and hereinaf-

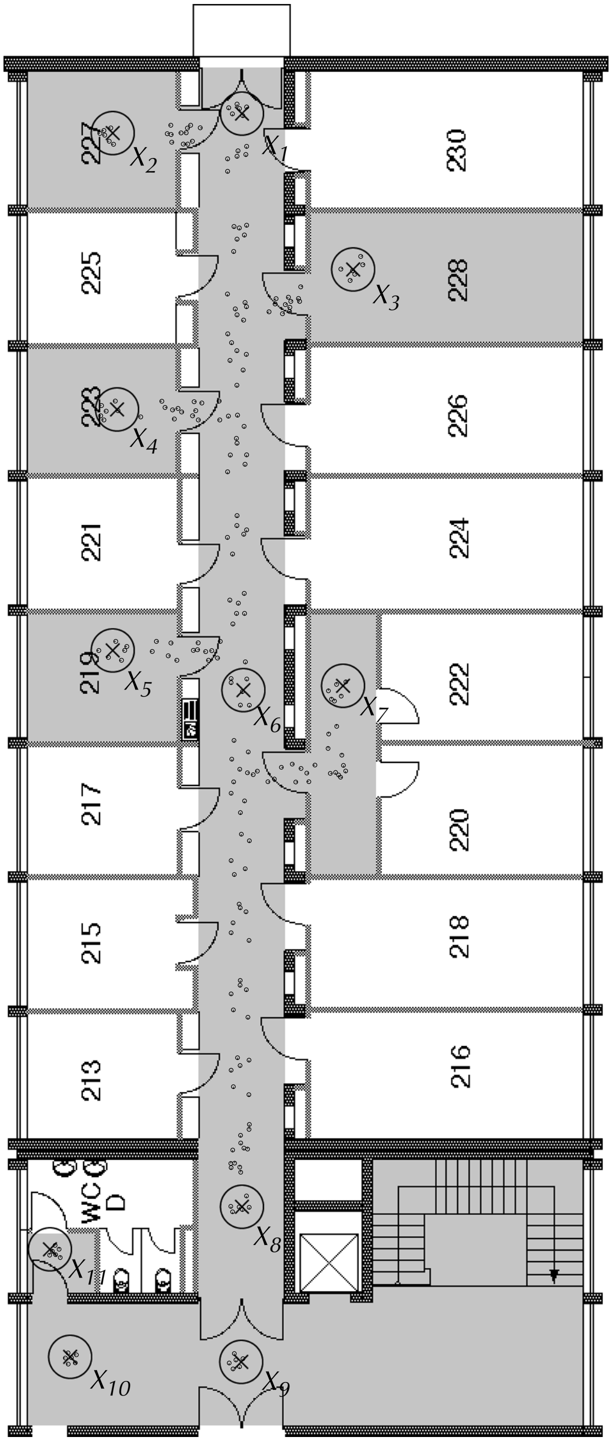

Figure 1: The environment for the experimental test. The staircase and bathroom are not used in the main experiment. Total area: 14m × 33m. Width of the hallway: 2m. Height of the ceiling: 3.25m in rooms, 2.51m in the hallway and 3.35m in the staircase and bathroom. The circled crosses (named X_1 - X_8) mark the positions to be discriminated from one another.

ter estimating the position of the robot with respect to the typical large scale division of buildings used by humans, namely rooms, hallway and staircase.

In addition to the robot itself, a sound system consisting of a laptop computer and an active loudspeaker (Bose 'Room-Mate II') is used to excite the environment acoustically. The signal used is a 10 ms frequency sweep from 1 KHz to 10 KHz, received at the position of Lisa's microphone with a maximum level of 76 dB. The position of the loudspeaker was held constant with respect to the robot.

The procedure

For the main experiment, only the rooms and the hallway, without the staircase and bathroom, are used. Section 5 describes additional tests done in the staircase and hallway. Lisa is steered through the test environment six times, thereby stopping and emitting a sweep signal at least every two meters of travelling distance. The starting point for all runs is the end of the hallway in front of room number 227. (Position X_1) For each run, a path is chosen that leads through all rooms and ends in front of the automatic door at the other end of the hallway. (Position X_8) No two recording positions across all runs are closer to each other than 10 cm. (All positions are marked by small circles in figure 1.) This way, autonomous exploration of the environment is simulated in a realistic manner. When an onset of a sweep signal occurs, the following 400 ms of recordings through Lisa's microphone are saved. For each of the six runs, a path is taken so that for every marked position (the circled crosses in figure 1), one of the saved recordings takes place in the circle surrounding that mark. This last condition for the experiment is somewhat unrealistic for free explorative motion, but it is necessary to enable a clear definition for when the given hypothesis is confirmed by the experiment.

The result of this procedure is a set of 210 recordings $A_{ij} = (p_{ij}, s_{ij})$ with $i = 1, \ldots, 6$ the run number, $j = 1, \ldots, 35$ the recording number, p_{ij} the position of the recording, and s_{ij} the recorded signal. A confirmation of the hypothesis is given, if the following conditions can be met by a clustering of the set of recordings solely based on the s_{ij}:

- All recordings made in the circled area around one marked position are members of the same cluster.
- Recordings made in the circled area around different marked positions are members of different clusters.

At the optimum, the clustering should produce at least 8 disjunctive clusters, each of them including exactly all recordings made near one of the 8 marked positions.

The choice of positions which should be distinguishable is of course somewhat arbitrary. There is no objective measure as to which positions should generally be distinguishable for the given environment. The settings made here follow the overall guideline of the experiment, namely to keep it as close as possible to real world applications. A discrimination of the marked positions provides information about which room or which region of the building the robot is in, following the notion of humans for the large scale division of buildings.

4. A representation and metric

Although there do not exist direct approaches towards the discrimination of rooms based on acoustics, several research areas provide starting points for a representation that seems suitable for defining a metric on them. The methods used in this work are taken from basic signal processing theory ([10], [14], [22]), speech recognition research ([1], [18]) and psychological research on the features of human listening ([9], [15], [23]).

Let $s : [1, 17640] \subset \mathbb{N} \to [-32767, 32768] \subset \mathbb{N}$ be a discretely sampled time series representing a 400 ms 16-bit recording (with 44.1 KHz sampling rate) of the room echoes produced through a sweep as defined in the experiment. The onset of the sweep is at $s(1)$.

Moving to the frequency domain

The representation for the acoustics of the surrounding is calculated in the following way. Let

$$\hat{s}(m,t) = C_m \cdot \left| \sum_{n=t-T/2}^{t+T/2-1} s(n) \cdot w(T/2+n-t) \cdot e^{2\pi i n m/T} \right|^2$$

with

$\frac{m}{T \cdot \Delta}$, $m = 0, \ldots, \frac{T}{2}$ the frequency in Hz

($\Delta = 44100^{-1}$ the sampling interval) and

C_m a normative constant per m

be the *periodogram* of the short-time Fourier transform of s at time t, based on a windowed T-point Fourier transform.[1] Let the window function be

$$w(n) = \frac{1}{2} \cdot \left[1 - cos\left(\frac{2\pi n}{T-1}\right)\right] \text{ for } 0 \le n \le T-1$$

[1] *For the calculation, be $s(t) = 0$ for $t \notin [1, 17640]$.*

which is the discrete Hanning window. $\hat{s}$ is calculated for values of t with a shift of t_{shift}, resulting in a matrix of size

$$\left(\frac{T}{2}+1\right) \times \left(\left\lceil \frac{17640}{t_{shift}} \right\rceil\right)$$

containing positive real values. With $T = 1024$ and $t_{shift} = 32$, this results in a 513×552 - matrix S which will serve as the basis for the signal representation. The resolution provided in the matrix for the acoustic signal is 0.726 ms in the time range and 43.07 Hz in the frequency range.

Modifications according to psychoacoustics

For noise reduction and data compression, the signal representation is modified using heuristics in compliance with well established psychoacoustic research results. The number of frequency values is reduced by summing up values according to the Bark scale. (described in [15]) A similar procedure which is explicitly used for sampled data can be found in the psychoacoustic model 1 of the MPEG standard as described in [11]. Thereby, all frequencies whose perception through humans influence the perception of a certain center frequency are grouped together. When all 513 frequency values are grouped this way without overlapping and following the MPEG specification, it results in 26 values for each time step.

Next, all values are moved to the dB-scale, and absolute minimum perception thresholds are applied as specified in [11]. The result is a matrix $\hat{S}$ with

$$\hat{S}(1,t) = max\left[thres_1, 10 \cdot \log_{10}\left(\sum_{i=2}^{2} S(i,t)\right)\right]$$

(corresponds to 43 Hz or 0.425 Bark)

… (continued approx. according to the Bark scale)

$$\hat{S}(5,t) = max\left[thres_5, 10 \cdot \log_{10}\left(\sum_{i=9}^{11} S(i,t)\right)\right]$$

(corresponds to 345 - 431 Hz or 3.3 - 4.1 Bark)

… (continued approx. according to the Bark scale)

$$\hat{S}(26,t) = max\left[thres_{26}, 10 \cdot \log_{10}\left(\sum_{i=362}^{465} S(i,t)\right)\right]$$

(corresponds to 15547 - 19983 Hz or 24 - 24.6 Bark)

Overall, what is basically done is to move to the frequency domain, explicitly represent time and reduce the amount of data in a reasonable way according to the human perception of acoustic signals.

The metric

Here, we take a very straight first approach, as there is only little known on how to compare acoustic room signals. One of the most widely used family of metrics in acoustics research is based on the L_2-norm. The analogous definition for the described representation is

$$d(\hat{S}_1, \hat{S}_2) = \sqrt{\sum_{m=1}^{26}\sum_{t=1}^{552}\left|\hat{S}_1(m,t) - \hat{S}_2(m,t)\right|^2}$$

This corresponds to evaluating the difference of $\hat{S}_1$ and $\hat{S}_2$ according to distribution of signal energy over frequency and time.

First results within the test experiment specified in section 3 for this model and metric showed that the noise level was too high as to get a reasonable amount of information from the distance values. There are many ways to cope with this problem, some of which will be discussed at the end of this article. A very simple method is to reduce the dimensions of the models by summing up values. A metric definition that accomplishes just that is

$$d_t(\hat{S}_1, \hat{S}_2) = \sqrt{\sum_{t=1}^{552}\left|\sum_{m=1}^{26}\hat{S}_1(m,t) - \sum_{m=1}^{26}\hat{S}_2(m,t)\right|^2}$$

$$d_f(\hat{S}_1, \hat{S}_2) = \sqrt{\sum_{m=1}^{26}\left|\sum_{t=1}^{552}\hat{S}_1(m,t) - \sum_{t=1}^{552}\hat{S}_2(m,t)\right|^2}$$

$$d_v(\hat{S}_1, \hat{S}_2) = \sqrt{\alpha_t \cdot d_t(\hat{S}_1, \hat{S}_2)^2 + \alpha_f \cdot d_f(\hat{S}_1, \hat{S}_2)^2}$$

What is done here, is to reduce the matrix to two vectors, one representing the distribution of signal energy over time and the other over frequency. Model instantiations are compared within each vector type via L_2-norm and a weighted distance of the results is taken as the overall distance of the model instantiations. The weights are chosen according to the given problem and represent in how far one wants to give priority to frequency changes or to changes in time structure of the recorded signals. For our experiments, we chose $\alpha_t = 1$ and $\alpha_f = 0.01$.

5. Experimental results

Experimental results consist of a clustering of the recordings A_{ij}. Obviously, research in data analysis should be able to contribute to the problem in question. Still, we chose to focus only on the acoustic signal modeling and metric to keep the interpretation of results manageable. To ensure that we are not observing new features of some clustering method but mainly the characteristics of the chosen model and metric for acoustic signals, a well known, universal clustering algorithm is used, following [8]. We chose the hierarchical clustering with ‘group average link’ as a standard procedure with overall good performance. This algorithm requires first collecting all recordings and then clustering them all together. In light of the requirement of staying close to real world applications, this procedure *alone* for all six runs is not a realistic line of proceeding. Therefore, only three runs are clustered this way, thereby simulating an ‘exploration and calibration phase’. The only parameter in this procedure is the final number of clusters to be built, representing the accuracy with which we would like to discriminate positions in the environment. The emerging clusters are arbitrarily labeled by the operator. The remaining three runs are then added to the cluster structure on a one by one basis, thereby enabling an immediate response of the system as to which cluster (with its label) matches best the recording just observed.

In the following, we present a typical result of the experiment which has been reproduced using any order of the six runs for clustering. Here, the number of clusters to be built in the hierarchical clustering was set to 13, as any higher number of clusters produced higher error rates without improvement according to classification of recordings, and a lower number of clusters resulted in simple merging of the cluster structure reported on in the following.

- The position X_8 could be discriminated from any other position in the environment with an error rate of 0%.

This result represents a complete confirmation of the given hypothesis for the position viewed. For all other to be distinguished positions, the classifications had certain error rates that are reported in the following.

- Position X_1 could be distinguished from all other marked positions, but not from all positions in the experiment. The corresponding cluster also contained numerous recordings made throughout the hallway.
- Positions X_2, X_3, X_4, and X_5 could not be distinguished. All Recordings from the rooms 219, 223, and 227 were clustered together with high stability. Recordings from room 228 were either clustered together with these rooms or with recordings from the hallway.
- Position X_6 could be distinguished from all other marked positions with a classification error in two recordings. The corresponding cluster also contained numerous recordings made throughout the hallway.
- Position X_7 could be distinguished from all other marked positions with a classification error in one recording, which belonged to the cluster of X_6. The cluster of X_7 only contained recordings made in the anteroom of 220/222.

Taking a more broad view on the result of the clustering, the following labels could be matched by the recordings contained in a cluster with the error rates reported.

- The hallway could be discriminated from any other parts of the environment with an error rate of 5.24% (11 of 210 recordings were misclassified when choosing cluster labels accordingly.)
- 'Position X_8' (1 cluster), error rate 0% (0 of 6 recordings misclassified)
- 'Anteroom of 220/222' (1 cluster), error rate 6.7% (1 of 15 recordings misclassified)
- 'Middle part of hallway + approach coming from rooms' (1 cluster), error rate 9.5% (6 of 63 recordings misclassified)
- 'Room 219, 223, 227 + approach coming from the hallway' (3 clusters), error rate 14.6% (6 of 41 recordings misclassified)
- 'Hallway at 227/225 and between 213 and 217 + approach coming from rooms' (3 clusters), error rate 19.0% (12 of 63 recordings misclassified)

Considering the last five examples mentioned, if the remaining 22 recordings are considered misclassified, this results in an overall error rate of 22.38% within the classification given by the cluster labels.

Extension of the experiment

The environment used in the experiment is considered difficult for position discrimination, as differences of rooms in size and shape are small and building materials provide damping of acoustic signals for the sake of a quiet working environment for the employees. Especially, rooms 219, 223, and 227 could not be distinguished by the model and metric in use. In order to check the performance of the given system in a probably 'easier' environment, the experiment is expanded to the hallway and the bathroom of the same floor. The new recording positions are shown in figure 1, together with three to be distinguished positions, marked by circled crosses.

Results show that recordings made at positions X_9, X_{10}, and X_{11} could easily be discriminated from each other as well as from recordings made at X_8, thereby confirming the hypothesis in question for the given environment. Distance values between recordings made at one position are comparable to those observed in the main experiment, whereas distance values between recordings made at different positions are larger by a multiplication factor of up to 10.

6. Discussion

The results presented in this article show that it is possible to build an artificial acoustic system capable of distinguishing positions of a mobile robot in a silent, static, indoor environment. Several factors can be identified that play a role in forming these results.

- The test environment can be considered difficult for a sound discrimination task, as differences between rooms are minor. The extension of the experiment, using the staircase and bathroom, shows that depending on the environment, even with high noise levels, the simple approach taken can provide useful information for navigation.
- The model, metric and clustering are made up of basic methods. This leaves plenty of room for further refinement in all respects.
- Besides the structure of the environment and the characteristics of the model and metric, a major factor for the error rates reported is indeed the noise level. Figure 2 gives an idea of the signal-to-noise ratio and the colorfulness of the noise that had to be handled. Many applications will have much lower noise levels, so performance will in general be *at least* as good as in the experiment described.

As previously mentioned, there are several ways to cope with the problem of high noise levels and bad signal-to-noise ratios. The approach taken in this work is very simple in nature. Other simple methods require restrictions on the characteristics of the noise which in general seem to be unrealistic to the author, as for example requiring certain statistical properties. A more sophisticated and well studied method would be to implement an optimal filter, based on an estimation of the frequency spectrum of the noise and knowledge about the test signal used. ([10], [14], [22]) The problem with this approach is that, for the application scenario in question it is not possible to estimate the frequency spectrum of the noise *at the same time* as the recording of the test signal is done. This is because the test signal will dominate recordings at this time made *at any place* in the surrounding. So the estimation of the noise spectrum has to be based on some other time interval, loosing the possibility of directly implementing an optimal filter as suggested in most signal processing literature. These approaches assume that the corrupted signal is formed by adding the estimated frequency spectra of the noise and signal, which requires a correct estimation not only in amplitude but also in the phase shift of the noise and signal in question. As long as one does not require the noise in the application to meet certain constraints apart from not exceeding some

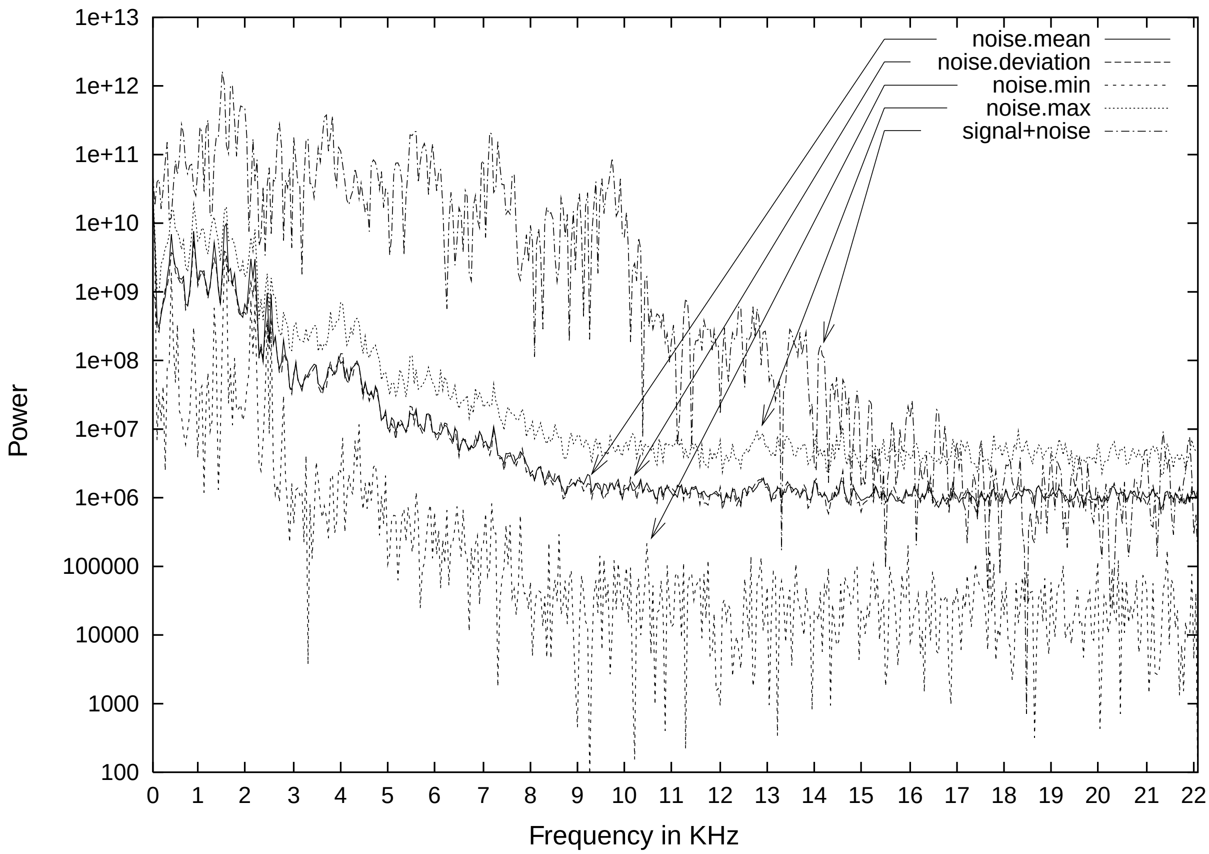

Figure 2: Example for the frequency characteristics of the noise in the experiment. Values are based on DFTs of 552 successive, overlapping windows with size 1024 and overlap 32 of 400 ms of signal data with no onset of a test signal. This matches with the data the representation of the test signals is based on. “signal+noise” indicates the frequency vector used in the metric calculation for the developed procedure for a typical 400 ms recording.

energy level, knowledge about the phase of the noise at the time of the signal recording is necessary to implement an optimal filter.

A different, and also very simple way of handling noise would be to do several signal recordings at the same position and average the values of their time-frequency-representations. The practicability of this method depends on the characteristics of the application viewed and could very well have been used in the experiment described in this article. This approach might be tested in future experiments.

Overall, the work described in this article suggests that further investigation in the topic of room discrimination based on broadband acoustics can provide a substantial gain in performance for navigation systems, depending on the characteristics of the application in question.

References

[1] Michèle Basseville, *Distance measures for signal processing and pattern recognition.* Institut National de Recherche en Informatique et en Automatique (Le Chesnay, France), INRIA, No. 899, September 1988.

[2] Jens Blauert, *Spatial hearing: The psychophysics of human sound localization.* Rev. Ed., MIT Press, Cambridge, MA, US, London, 1996.

[3] Albert S. Bregman, *Auditory scene analysis: Hearing in complex environments.* In: S. McAdams, E. Bigand (Eds.), "Thinking in sound: the cognitive psychology of human audition.", p. 10-36, Oxford University Press, Oxford, England, 1993.

[4] Guy J. Brown, Martin P. Cooke, *Computational auditory scene analysis.* Computer Speech and Language, Vol. 8, p. 297-336, 1994.

[5] Lothar Cremer, Helmut A. Müller, *Die wissenschaftlichen Grundlagen der Raumakustik, Band I.* 2nd Ed., S. Hirzel Verlag, Germany, 1976.

[6] Richard O. Duda, *Connectionist models for auditory scene analysis.* In: J. D. Cowan, G. Tesauro, J. Alspector (Eds.), "Advances in Neural Information Processing Systems. -6-", p. 1069-1076, Morgan Kaufmann, San Francisco, CA, US, 1994.

[7] Alon Fishbach, *Primary segmentation of auditory scenes.* Proc. of the 12th IAPR Intl. Conf. on Pattern Recognition, Vol. 3, p. 113-117, 1994.

[8] William B. Frakes, Ricardo Baeza-Yates (Eds.), *Information retrieval: data structures and algorithms.* Prentice Hall, New Jersey, US, 1992.

[9] Harold L. Hawkins, Teresa A. McMullen, Arthur N. Popper, Richard R. Fay (Eds.), *Auditory Computation.* Vol. 6 (Springer handbook of auditory research), Springer-Verlag New York, Inc., 1996.

[10] Monson H. Hayes, *Statistical digital signal processing and modeling.* John Wiley & Sons, New York, NY, US, 1996.

[11] ISO/IEC Standard, *Information technology - Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s - Part 3: Audio.* International standard ISO/IEC 11172-3, First Ed., 1993.

[12] Heinrich Kuttruf, *Room acoustics.* 3rd Ed., Elsevier Science Publishers Ltd., New York, NY, US, 1991.

[13] Chin-Hui Lee, Frank K. Soong, Kuldip K. Paliwal (Eds.), *Automatic speech and speaker recognition. Advanced topics.* Kluwer, Dordrecht, NL, 1996.

[14] Sanjit K. Mitra, James F. Kaiser (Eds.), *Handbook for digital signal processing.* John Wiley & Sons, New York, NY, US, 1993.

[15] Brian C. J. Moore, *Hearing.* (Handbook of perception and cognition, 2nd Ed.), Academic Press, Inc., San Diego, CA, US, 1995.

[16] Charles E. Rice, Stephen H. Feinstein, Ronald J. Schusterman, *Echo-detection ability of the blind: Size and distance factors.* J. of Experimental Psychology, Vol. 70, No. 3, p. 246-251, 1965.

[17] Nicholas Roy, Gregory Dudek, Paul Freedman, *Surface sensing and classification for efficient mobile robot navigation.* Proceedings of the 1996 IEEE Intl. Conference on Robotics and Automation, Vol. 2, p. 1224-1228, IEEE, New York, NY, US, 1996.

[18] C. L. Searle, *Representing acoustic information.* In: Whitman Richards, "Natural computation.", p. 309-318, MIT Press, Cambridge, MA, US, 1988.

[19] Gilbert A. Soulodre, John S. Bradley, *Subjective evaluation of new room acoustic measures.* J. of the Acoustical Society of America, Vol 98(1), p. 294-301, July 1995.

[20] Thomas A. Stoffregen, John B. Pittenger, *Human Echolocation as a Basic Form of Perception and Action.* Ecological Psychology, Vol. 7, No. 3, p. 181-216, 1995

[21] Michael Supa, Milton Cotzin, Karl M. Dallenbach, *"Facial vision": The perception of obstacles by the blind.* The American Journal of Psychology, Vol. 57, No. 2, p. 133-183, April 1944.

[22] Saeed V. Vaseghi, *Advanced signal processing and digital noise reduction.* B. G. Teubner, Stuttgart, 1996.

[23] E. Zwicker, H. Fastl, *Psychoacoustics: Facts and models.* Vol. 22 (Springer Series in Information Sciences), Springer-Verlag, Berlin, 1990.

Guido Tesch, born 1970, studied computer science at the University of Bonn, Germany, where he received the diploma in 1998. This article is based on his diploma thesis.