Proceedings of the AUV 2000, State College, Pennsylvania, USA, June 28-29, 2000

Acoustic-based position discrimination of a moving robot

Guido Tesch

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

Tel.: +81 - 93 - 512 1566
http://www.gmd.gr.jp/ guido.tesch@gmd.gr.jp

A real-world experiment is described, demonstrating the possibility of discriminating positions of a mobile robot moving through an unknown, unprepared office room, solely based on broadband audible acoustic signals. The average distance of distinguishable positions was found to be less than 15 cm. MLS measurements are used to obtain the impulse response of the room at the present position, and a modified vector distance measure is used to cluster the data via the Neural Gas algorithm. No further interpretation of the data is performed. The data-analysis method is independent of the environment or the recording system (the robot), therefore it is potentially applicable to other environments and vehicles. The method is especially interesting for the navigation of AUVs in unknown environments, as acoustic signals represent a major source of information in the underwater world, being available under the broadest range of circumstances.

Keywords: **broadband acoustic signals, acoustics, mobile robots, autonomy, exploration, navigation, MLS measurement**

1. Introduction

The overall topic of this paper is the navigation of autonomous mobile robots in unknown, unprepared environments. More precisely, the problem under investigation is the recognition of a revisited position in space and finding other previously visited positions relative to that position. The primary components of these procedures are

- position characterization, and
- position comparison.

This paper suggests a new sensor modality (omnidirectional broadband acoustics) as a complementary information source for autonomous robot navigation, which is local and rich in content. A data-analysis method for broadband acoustics is presented that focuses on local navigation systems using non-geometrical world models.

The hypothesis of this work is that broadband audible acoustic signals can be used to distinguish between positions of a mobile robot moving through an unknown, unprepared, static office room, with the perspective of using this information for navigation tasks such as recognizing revisited positions in the room. Active sensing will be used, meaning the room is excited acoustically and the resulting echoes are analyzed. The robot's navigation should be independent of sound sources in the environment, therefore any signals from such sources are viewed as undesirable noise.

After describing the data-analysis method, a real-world experiment is presented demonstrating the potential of this method. The results of the experiment are discussed, and the paper concludes with some aspects of transferring the method to AUV applications.

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

2. General background

The use of broadband audible acoustic signals to analyze the surrounding is primarily researched in the field of room acoustics [3]. From general signal processing research it is known [2] that the effect of any linear, shift invariant (LSI) system T on a signal x can be completely described by the impulse response (IR) h of that system.

$$y(n) = \sum_{k=-\infty}^{\infty} x(k) \cdot h(n-k) \quad \text{or}$$

$$y(n) = x(n) * h(n) \quad \text{, with} \quad h(n) = T[\delta(n)]$$

Here, $\delta(n)$ denotes the unit sample at index value n and $*$ denotes the convolution sum. LSI systems are often called linear time invariant (LTI) systems, if the domain for n is time.

Regarding their acoustic properties, silent environments of our world are generally considered to be LTI systems, so a fundamental topic in room acoustics is the determination of the IR for a given environment. Today, the method generally preferred for IR measurement uses a Maximum Length Sequence (MLS) excitation signal. [1][5] The following list names some properties of this method that are important for the present context.

- Practical, meaning the test signal is easy to produce
- Measurements are accurately reproducible
- High signal-to noise ratio with low crest factor
- Insensitive to transient noise (cracking, snapping)
- Fast

3. Data-analysis method

The data-analysis method that was used in the experiments reported on in this paper is described in the following paragraphs. It was implemented to verify the hypothesis stated in chapter 1 above, and is viewed as a first promising step towards an acoustic-based navigation system.

IR estimation

- Based on MLS measurement using an ordinary loudspeaker and microphone
- Output rate and sampling rate each 10 kHz, using 12 bit signed linear coding
- 16383-point MLS signal ⇒ 3.28 s for a single measurement (The first MLS is needed to stabilize the acoustic response, therefore the first MLS usable for measurement is the second one.), 1.64 s when operating continuously (assuming the system is stabilized by one MLS at the beginning of the overall measurement)

Result: A data vector x of length 16383 containing sampling values in the integer range [-2048; +2047].

The length of the MLS signal is chosen so that no time aliasing occurs, i.e. the impulse response of the environment is assumed always to become negligible after at least 1.64 seconds. (≈ 570 m sound traveling distance in air at 25 °C) The output and sampling rate are chosen so as to assure that all components in the measurement system operate in their linear range. The 12 bit coding is given by the computer's I/O card.

Modification of the IR

- Applying a high-pass filter with cut-off frequency 200 Hz to eliminate strong 60 Hz power line noise ⇒ Frequency range is 200 - 5000 Hz (approx. 4.5 octaves)
- Use only the first 600 samples (= 60 ms or approx. 21 m sound travel distance in air), which include at least all first reflections from objects closer than 10 m.
- Compensate roughly for transmission loss using the simplified formula

  $$x(i) = x(i) \cdot \left(1 + \frac{i}{28.719127}\right) \;, \quad i \in [0, 599]$$

  This results in sample values being doubled for every meter of sound traveling distance (assuming the speed of sound to be 348.2 m/s at 25 °C)
- Apply a Gaussian smoothing filter with mean value zero and standard deviation 5.0, resulting in approximately 11 sample values being averaged together with Gaussian weighting.

We only use the first reflections of the IR as these contain all information viewed as relevant for the task in question. (Meaning distances to nearby objects in the environment.) Considering also the rest of the IR would reveal further acoustic properties (e.g. the reverberation time), but it is assumed that these are not necessary to distinguish positions within a room.

The compensation for transmission loss is necessary so that echoes arriving at different times are not unequally weighted in the distance measure as it is defined below.

The Gaussian smoothing filter has two effects: 1. it reduces the amount of white noise in the IR, and 2. it broadens sharp peaks in the IR, which represent distinct echoes (first reflections). For a further description, see the distance measure discussion below.

Distance measure

- Standard vector distance, based on the L_2-norm

$$d(x_1, x_2) = \sqrt{\sum_{i=0}^{599} (x_1(i) - x_2(i))^2}$$

This distance measure reflects the difference of energy levels in the modified IRs for each point in time. Because of the Gaussian filter applied to the IRs, strong echoes that normally appear as narrow peaks in the IR are broadened. This results in a smoother transition in the difference value when echoes become closer in time, thereby reducing the accuracy of position discrimination. On the other hand, the Gaussian filter reduces the amount of noise in the IR, thereby facilitating the correct, continuous clustering of positions in the environment. In essence, there is a trade-off to be realized between accuracy and noise reduction in function of the choice of the parameters for the Gaussian filter.

Data clustering method

- Unsupervised clustering using the Neural Gas Algorithm (NGA) [4]. Parameters: $\lambda_i = 5.0$, $\lambda_f = 0.01$, $\varepsilon_i = 0.5$, $\varepsilon_f = 0.005$, training set size 140 (3 runs in the experiment below), 30 epochs ($t_{max} = 4200$) number of neurons: 50

Unsupervised clustering is used to avoid any kind of interpretation of the data. This keeps the overall data-analysis method applicable to other application areas. The NGA is a well known, well behaved clustering method that is globally stable and converging, robust against faulty measurements, and it preserves the full dimensionality of the data. The λ and ε parameters are set to standard values, and the training set size results from the way the experiment below was realized. The number of epochs (i.e. the average number of times each IR is accounted for in the clustering algorithm) was set to a high value so as to investigate the maximum performance that can be reached within the present data-analysis method. Further investigations might result in an optimization of these parameters. The number of neurons was chosen so as to avoid under-representation as well as over-representation within the experiment.

4. Method of experiment

A fundamental design criteria for the experiment was to keep it close to real-world applications. As there exists no thorough theoretical background on acoustic-based navigation [6], any results can only be of interest if they have a direct connection to practical use. What follows is a detailed description of the experimental setup as well as of the procedure carried out.

General goal

A robot system potentially capable of realizing autonomous navigation through indoor environments is equipped so that it can continuously measure the IR of the environment while in motion. It is moved through an office room, thereby recording the positions at which IR measurements take place. A ground plan of the room is drawn showing these positions and how they are clustered by the data-analysis method. A verification of the hypothesis in question is considered given if the following is true:

The clustering result forms a set of clusters of positions so that

- there are at least two clusters,
- every cluster includes at least two positions,
- there is no geometrical overlap between clusters.

A system achieving this goal will be valuable for autonomous navigation, as it realizes a stable recognition of revisited positions. The degree of usefulness is determined by the number and geometric sizes of positions discriminated as well as the respective demands of the application in question.

The environment

The environment under investigation was a standard office room, measuring 5.5 m x 10 m and with a height of 2.62 m, equipped with tables, chairs, shelves, computers, some boxes and a sofa. The room was not specially prepared for the experiment, except that any acoustic queues for the robot's position were eliminated. (Any active sound sources in the environment are silenced.) The background noise in the room was a constant 23.7 dBA. The main materials used in the environment were concrete, metals (steel, aluminium), and wood. The floor was covered with a thin carpet, as well as with large wooden plates towards the upper right corner. Figure 1 shows a ground plan of the environment.

The robot / measurement system

A robot ATRV-JR from Real World Interfaces Inc. equipped with some additional components for acoustic measurements was used: An I/O card from National Instruments (PCI6071E), an electret condenser microphone from Sony (ECM-T145), and a loudspeaker from Bose (Roommate II). The position of the microphone was 120 cm above and 60 cm behind the loudspeaker, which was located at the front of the robot looking forward. The basic noise

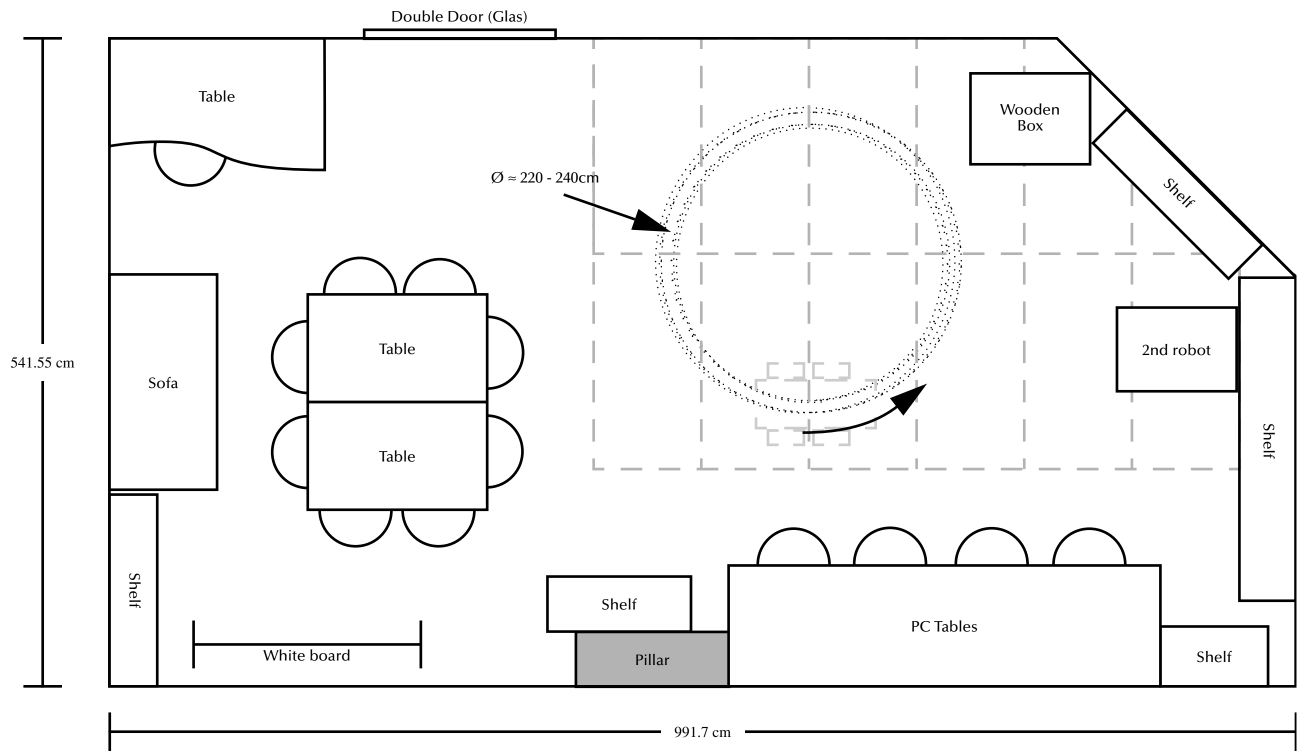

Figure 1: A ground plan of the environment for the experiment, together with the positions at which the robot started an MLS measurement. The diameter of the circle followed by the robot as well as the direction of travel are indicated.

level of the vehicle at the microphone (produced by 3 fans and a hard disc) was 45.5 dBA. The level recorded during test signal emission was 97.2 dBA in front of the loudspeaker at 1 m distance, and 86.7 dBA at the position of the microphone.

Procedure

A default starting position was defined, which is marked by a shaded robot outline in figure 1. The robot was placed at some position relative to this default starting position, and the robot then drove a complete circle with a given radius and at a given speed. During motion, the robot continuously emitted the MLS test signal and recorded the room response. After one circle was completed, the robot was stopped, repositioned by hand, and restarted. This procedure was repeated for several starting positions close to the default, and with the robot driving at either 10 cm/s or 3.3 cm/s. In addition, the system was placed at 276 positions along the circular path, with an IR measurement performed without motion at each of these positions.

The robot's odometry system was used to record the position of the robot for each point in time where an onset of a MLS in the test signal occurred (every 1.64 seconds). The error of the odometry values for positions during one circle were confirmed to be less than 5 cm. Figure 1 shows all recorded positions as small black dots along circular paths in the middle of the room.

Discussion of the procedure

The path of the robot during IR measurements was chosen so as to ensure that the path itself did not induce a cluster structure, and that there was a considerable overlap in sensor data between successive measurements. For example, if a box shaped path around the table to the left side of the room had been chosen, it would not have been surprising if clusters had formed for positions along each straight edge and for the corners. Inferring some usefulness for navigation from such results is considered problematic.

The robot's path was controlled in a way so that experimental results based on data collected at different speeds are comparable. With an autonomous, exploration-like behavior, it would have been difficult to realize this experimental condition. Nevertheless, the variation in starting position and diameter of the circular path to follow as well as the inaccuracy of the robot's motor control resulted in considerable variation in the paths, as can be seen in figure 1. The experimental procedure is therefore considered to be close to real-world applications.

Another reason for restricting the robot's path was that the sound system used was not omnidirectional. Therefore, not only its position in space but also the robot's orientation influences the sensor readings. The complexity of the problem under investigation was controlled by keeping a constant orientation for each segment of the circular path, with a maximum error of 10º. Future investigations might either make use of a purely omnidirectional sound system, or take orientation into account for the robot's navigation.

5. Experimental results

Many different combinations of training data sets and test data sets for the NGA were tested, either using only data collected with one speed or combining data from all speeds. The following results represent typical examples. Also, the parameters of the data-analysis method were varied, the results of which are reported on in a qualitative sense. For any quantitative results depicted in figures below, the parameters for the data-analysis method given in its description above were used.

As an overall result of the experiments, it can be stated that no matter what speed was used, the position discrimination ability of the system, for land robots as the one used here, was clearly in the range of what often used indoor navigation sensory systems can achieve nowadays.

Speed 10 cm/s

Figure 2 shows a typical result of the clustering of positions for data recorded at 10 cm/s. The training set consisted of 140 IRs, the test set included all 276 IRs recorded at 10 cm/s. Some characteristics for this speed that become apparent when comparing results for different set constellations, parameters and other speeds are:

- The circular path is segmented into distinct areas, clustering only positions that lie geometrically close together.
- There is a slight overlap between adjacent clusters, which does not occur in this way for the other speeds.
- There are clusters with only one element. Changing the number of clusters in the NGA either results in more "singletons" or in remote positions being clustered together.

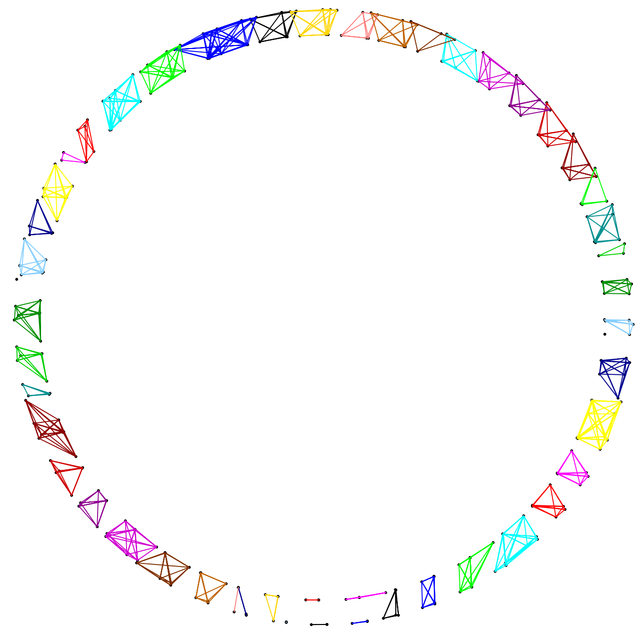

Figure 2: A typical result of the clustering of positions for data recorded at 10 cm/s.

Speed 3.3 cm/s

Figure 3 shows a typical result for data recorded at 3.3 cm/s. Again, the training set size was 140, and the test set included all 276 IRs collected at the present speed.

Some basic results that are typical for the present speed:

- The circular path is segmented into distinct areas, clustering only positions that lie geometrically close together.
- There is no overlap between clusters.
- All clusters contain at least two recordings.

This result represents a complete confirmation of the goal criteria formulated at the beginning of chapter 4. The geometric size of positions discriminated in the experiment can be estimated as 14.75 cm. The maximum geometric size of a position is 27 cm.

Speed 0 cm/s

Figure 4 shows a typical result for data recorded without motion. Training set size and test set size were again 140 and 276. The basic findings for this setup were:

- The circular path is segmented into distinct areas, and generally positions are clustered together that are geometrically close. Nonetheless, there are always positions located far apart from each other that are clustered together, with the actual

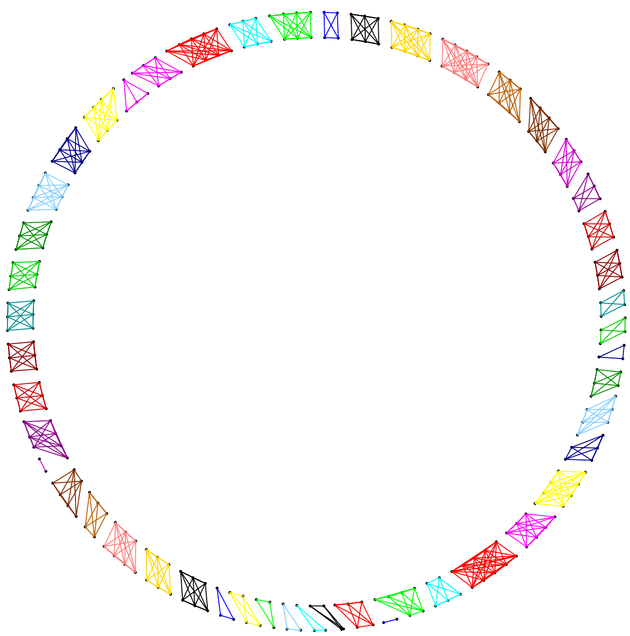

Figure 3: A typical result of the clustering of positions for data recorded at 3.3 cm/s.

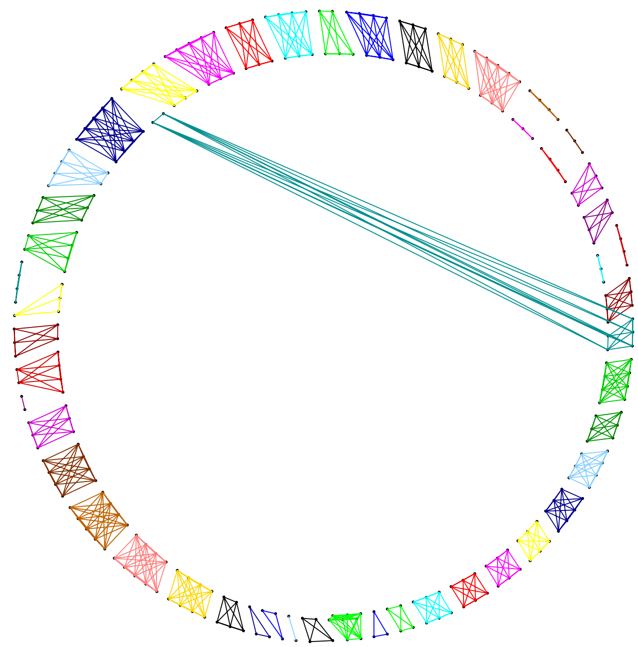

Figure 4: A typical result of the clustering of positions for data recorded without motion.

set of positions varying with the choice of training set and test set (i.e., not due to faulty measurements).

- There is no overlap between clusters containing geometrically close positions. There are cases in which more than one cluster exists that contain positions geometrically apart, which then overlap in most cases.
- All clusters contain at least two recordings.

Interference effects between echoes and resulting discontinuities in the clustering of the environment might have been the reason for the clustering of remote positions. A detailed verification of this hypothesis will be a subject of future research. Further comments on this aspect follow below.

Any speed

Figure 5 shows a typical result for the clustering of all data produced within the experiment, with the training set containing 276 recordings from different speeds, and the test set containing all 828 IRs. Some basic findings:

- The circular path is segmented into distinct areas, clustering only positions that lie geometrically close together.
- There is a slight overlap between adjacent clusters.
- All clusters contain at least 7 recordings.

Modifications of the data-analysis method

Changing the length of the IR considered:
Taking fewer data points from the IR into account resulted in clustering together of remote positions, giving rise to an overlap of clusters. For the environment in question, taking less than 600 data points into consideration would have meant not accounting for even first echoes from remote objects, although the signal-to-noise ratio for these echoes would have been high enough for them to be used for position discrimination.

Taking more data points from the IR into account resulted in either no observable change or (when taking for example 10000 data points into account) in a complete breakdown of the position discrimination, meaning that positions were clustered together in a more or less random fashion. This effect is mainly due to the fact that continuous noise in the IR estimation is more and more elevated with time by the transmission loss compensation. In the end, very late data points are dominated by (more or less random) noise, with these points in turn dominating the overall distance between IRs.

Essentially, choosing the length of the IR to be considered depends on the signal-to-noise ratio for echoes, which in turn is determined by the overall transmission loss of the acoustic signal, the noise present for the environment under investigation, and the energy level of the emitted test signal.

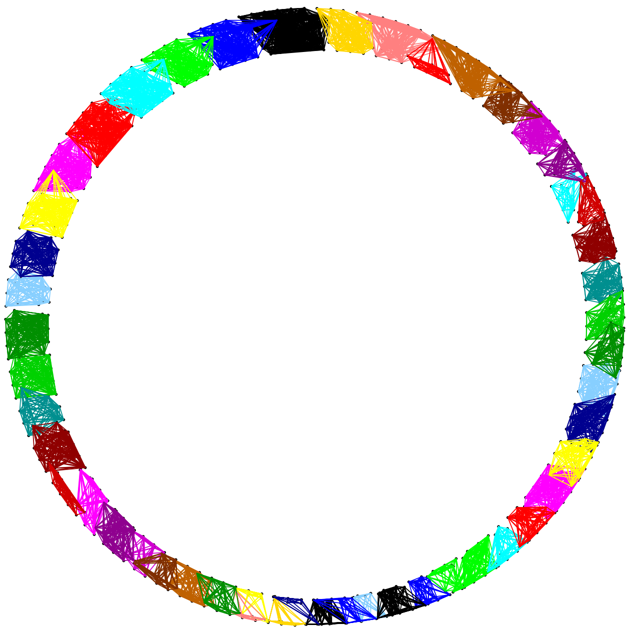

Figure 5: A typical result of the clustering of positions for all data collected in the experiment.

Changing the Gaussian smoothing:
Applying no smoothing to the IR data resulted in a high degree of randomness in the clustering for most positions, whereby some positions in small areas were always forming separate clusters, enabling recognition of those regions. Increasing the number of points averaged together with Gaussian weighting had the effect of decreasing the randomness of the clustering and increasing the number of areas being discriminated, with an optimum reached at about 11 points. Further increasing the Gaussian smoothing resulted in the merging of areas either close to each other or at remote locations.

These findings confirm the statements made in chapter 3 for the Gaussian smoothing and for the distance measure. Similar to choosing the length of the IR to be considered, the choice of the best parameter value for the Gaussian smoothing depends on the level and kind of noise present in the application under investigation.

6. Discussion of results and outlook

With the experiment described above it could be shown via construction that it is possible to discriminate positions of a mobile robot moving through an office room based on broadband acoustics, with an accuracy that makes the system valuable for robot navigation. This represents a confirmation of the hypothesis in question, for the environment and the robot system used in the experiment. Best results are achieved with slow motion of the robot, with only a small decrease in performance when combining data from all speeds of the experiment, showing the applicability of the system to real-world applications.

The occurrence of interference effects

If the signal used to excite the environment has an autocorrelation who's absolute value is significantly above zero, then even smallest changes in the robot's position can result in large changes in the recorded signal, as strong echoes arriving at the microphone close in time might then arrive with an overlap that suddenly results in the echoes cancelling each other out. Such interference effects could be the reason for the decrease in the system's performance when making measurement without motion as compared to measurements in motion. Moving during the IR measurement is assumed to reduce interference effects, as in this case the measured IR is actually a blend of many IRs, for which typically only few show interference effects.

Nevertheless, the results reported on in this paper do not represent a scientific proof for this statement. The relevance of interference effects for the method's performance depends on the environment in question. The more often strong echoes arrive at the microphone close in time, the more likely interference effects will occur. Informal tests carried out with the same robot system in different environments showed that for example, driving through a narrow corridor with the same distance to both side walls was a setup in which frequent, strong interference effects were observed. These effects were reduced when the robot moved during the measurements. Presently, a formal implementation of this test cannot be presented — a deficiency to be dealt with in future research.

Transfer of results to real-world applications in other environments

The following list names criteria that need to be met by the environment under investigation and the sound system used, so that the data-analysis method developed in this paper is applicable.

- Echoes reflected by objects in the environment, relative to which the navigation is supposed to take place, must have an energy level providing a sufficient signal-to-noise ratio in the measurement recordings. Factors that influence this criterion are the source energy level of the test signal, the transmission loss characteristics of the environment, the level of noise in the environment and measurement system, and the sensitivity of the receiving system.

- The signal emitted into the environment must have an autocorrelation low enough so that the occurrence of echo interferences is minimized. The importance of this criterion depends on the environment under investigation. The main factor determining the level of autocorrelation of the emitted MLS signal is the bandwidth and linearity of the acoustic production system.

The acoustic sensor data is used by the data-analysis method without any kind of interpretation. The parameters of the data-analysis method can be set prior to run-time of the system, based on the analysis of basic characteristics of the environment in question. Consequently, the performance of the system is not dependent on the nature of the environment. It is therefore transferable to other domains, such as underwater environments, as long as the above listed criteria are met.

An aspect important for the transfer of the findings to real-world applications is that the orientation of the vehicle has to be taken into account. The amount of increase in complexity for the navigation task is directly related to the directivity of the acoustic signal emitted, and in how far the influence of the vehicle itself on arriving echoes changes with its orientation. The experimental setup chosen in this work involved an acoustic system with a directionality typical for small-sized loudspeakers. The system's performance with free orientation was not tested. In general, the data-analysis method will be applicable, though the question of how accurate the system's position discrimination would be cannot be answered at this time.

Outlook

The GMD - Japan Research Laboratory is presently designing and building a new small-sized research AUV that will be equipped with a sound system for acoustic-based navigation underwater. Future experiments will be performed to confirm the practicability of the data-analysis method to this application domain.

The present distance measure as well as the modification of the data vector are focused on methods and approaches that suggest themselves from the view of acoustics and signal processing. All components of the data-analysis method result from well established research results, with their combination and the choice of using a Gaussian smoothing filter being the core novelties. Further refining this method and analyzing its characteristics will be topics of future research.

The setup chosen in the experiment involved a quieting of the environment, so as to eliminate any acoustic cues for the robot's position apart from the room's response to the test signal. Consequently, dealing with interfering noise from sound sources in the environment was not a topic of this research. Future investigations are needed to analyze the systems performance under noisy conditions.

References

[1] Borish, Jeffrey; Angell, James B.
An efficient algorithm for measuring the impulse response using pseudorandom noise.
J. Audio Eng. Soc., Vol. 31, No. 7, pp. 478-488, July/August 1983.

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

[3] Kuttruff, Heinrich
Room acoustics.
3rd Ed., Elsevier Science Publishers Ltd., New York, NY, USA, 1991.

[4] Martinetz, T. M.; Schulten, K. J.
A "neural-gas" network learns topologies.
In T. Kohonen, K. Mäkisara, O. Simula, J. Kangas, (Eds.), Artificial Neural Networks, Proceedings of ICANN'91, International Conference on Artificial Neural Networks, Vol. 1, pp. 397-402, North-Holland, Amsterdam, 1991.

[5] Rife, Douglas D.; Vanderkooy, John
Transfer-function measurement with maximum-length sequences.
J. Audio Eng. Soc., Vol. 37, No. 6, pp. 419-444, June 1989

[6] Tesch, Guido; Zimmer, Uwe R.
Acoustic-based room discrimination for the navigation of autonomous mobile robots.
Proceedings 1999 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA '99, pp. 274-281, Monterey, USA, November 8-9, 1999.

Guido Tesch, born 1970, studied computer science at the University of Bonn, Germany, where he received his diploma in 1998. Since 1999, he has been working at the GMD - Japan Research Laboratory. His main research interests are in autonomous robust systems and acoustic-based navigation.