DESIGN AND UNDERSTANDABILITY OF DIGITAL-AUDIO MUSICAL SYMBOLS FOR INTENT AND STATE COMMUNICATION FROM SERVICE ROBOTS TO HUMANS

Gunnar Johannsen

IMAT-Lab. Systems Engineering and Human-Machine Systems
University of Kassel
Moenchebergstr. 7, D-34109 Kassel, Germany
joh@imat.maschinenbau.uni-kassel.de
www.imat.maschinenbau.uni-kassel.de/members/joh/johhome.html

ABSTRACT

Auditory displays for mobile service robots are developed. The design of digital-audio symbols, such as directional sounds and additional sounds for robot states, as well as the design of more complicated robot sound tracks are explained. Basic musical elements and robot movement sounds are combined. Two experimental studies, on the understandability of the directional sounds and on the auditory perception of intended robot trajectories in a simulated supermarket scenario, are described.

1. INTRODUCTION

Multimedia technologies are currently available which can more systematically be applied for human-systems communication in industrial, transportation, medical, and many service domains [1], [2]. Taking the multimedia concept more seriously allows to re-integrate the most important human sensory modalities for the supervision and control of technical processes and systems. Thus, many drawbacks and restrictions of over-emphasizing the visual channel can be corrected by the additional appropriate use of auditory information.

However, before developing the next generation of multi-media human-machine interfaces for process supervision and control applications, more knowledge on suitable auditory displays needs to be collected. Even the investigation of auditory warning displays needs to be further intensified [3]. More ambitious auditory displays will go beyond the warning displays by communicating systems states and intentions by means of semantic sound symbols and sound tracks.

This paper reports on results from a recent research project[1] in which auditory displays for autonomous mobile service robots in a human-machine multi-agent environment have been investigated and developed. The service robot domain was chosen as an example for future use of auditory displays within all kinds of multimedia process supervision and control applications. The idea of the auditory displays for the service robots is to combine relevant noise signals of their movements with basic musical elements to intelligible auditory symbols. Co-operative mobile robots shall communicate their actual positions, movements and intentions as well as special states by means of non-speech audio symbolic expressions to the human.

In the next section, the design of directional sounds for robot movements as well as special robot state sounds and complete robot sound tracks will be explained. Two experimental studies, one on the understandability of the directional sounds and the other on the auditory perception of intended robot trajectories in a supermarket scenario will be described in Section 3.

2. DESIGN OF DIGITAL-AUDIO SYMBOLS AND SOUND TRACKS

2.1 Directional Sounds

Eight possible directions of motion of the robot in space are considered, namely the four main directions of Left, Up, Right, and Down as well as the intermediate directions of Down-Left, Up-Left, Up-Right, and Down-Right. Each directional sound consists of three tones. The musical basic elements rhythm and melody are used in the four main directions, independently of each other. The directional sound Up is represented by a melody upwards whereas a melody downwards denotes the sound Down. In both cases, each tone is of equal time duration. A rhythm of two short tones followed by one long one, all on the same pitch level, means Left. Consequently, a rhythm with one long tone followed by two short ones on the same pitch level expresses the direction Right. The musical elements melody and rhythm are

[1] The research project on "The Importance of Acoustical Information for the Guidance and Usage of Technical Systems" was supported by the VW-Stiftung (Volkswagen-Foundation) and the University of Kassel during my sabbatical research semester. It was carried out in Vienna, Austria from March to October 1999. My main host was the Institute for Handling Devices and Robotics at the TU Wien (Vienna University of Technology – Univ.-Professor Dr. P. Kopacek). Further, a co-operation existed with the Institute of Electro-Acoustics, Experimental and Applied Music at the University of Music and Performing Arts, Vienna.

combined in the intermediate directions with respective intermediate values of melody span and rhythm. Each of the eight directions are presented in four variations, with changed sound colour (timbre) or changed tempo. Music instruments and robot noises are used.

The eight directional sounds are a subset of a more fine-grained resolution of 48 directions around the circuit of a compass card of 360°. So far, only the subset of the eight directional sounds has been tested. All the sounds were created with a powerful PC and Windows 95, Logic Audio software, a MIDI synthesizer, and a keyboard. The pure musical sounds from the synthesizer have been recorded with the Yamaha DSP Factory with its audio expansion unit under the Logic Audio software on the PC. Thus, wav-files have been produced which needed some minor audio editing.

The equivalent audio-effect sounds of robot noises have been derived from DAT recordings of the movement noises of a real service robot in the laboratory. Many different original sounds from robot movements, navigation, and related activities have been recorded from two mobile robots.[2] The variation of the directional sounds based on robot noises has been generated through several steps of audio editing from the recordings of the movements of one of these service robot. The same melody and rhythm patterns are composed by time and frequency editing as in the pure musical sound cases. The Logic Audio editor has been used for frequency transpositions of the pitch levels as requested in the different tones of the directional sounds. The Cool Edit audio editor was more appropriate for cutting and assembling the necessary time slices of each sound element to the desired directional sounds. Some amplifications with fading-in and fading-out effects have been performed for achieving a clear separation between the different three tones of each sound.

2.2 Special State Sounds

Additional sounds for robot states and situations have been newly designed. These robot states and situations are Heavy Load, Waiting, Near Obstacle, and Low Battery. The latter sound was recorded from the original robot`s indication of the low battery status. This is a continuous, quite annoying high tone. The other three sounds have been played on the MIDI keyboard. The author tried to convey the subjective impression of the meanings of these three sounds. For example, the Heavy Load sound was played with three parallel tubas as one accentuated short time-interval tone followed by one tone of a longer time duration. Figure 1 shows the score and the wave form of this Heavy Load sound.

2.3. Robot Sound Tracks

The sound tracks of the intended robot trajectories are composed of moving-straight segments and turnings. The supermarket scenario was designed in such a way that straight movements and turnings of 45 and 90 degrees are possible. The moving-straight segments are represented by the directional sounds which are described in Section 2.1. The directions Left, Right, Up, and Down are particularly used but also a few of the intermediate directions (Down-Left, Up-Left, Up-Right, and Down-Right) are sometimes possible. Down means downwards on the computer screen and towards the human subject, shown on the lower middle of the screen. Correspondingly, Up means away from the subject.

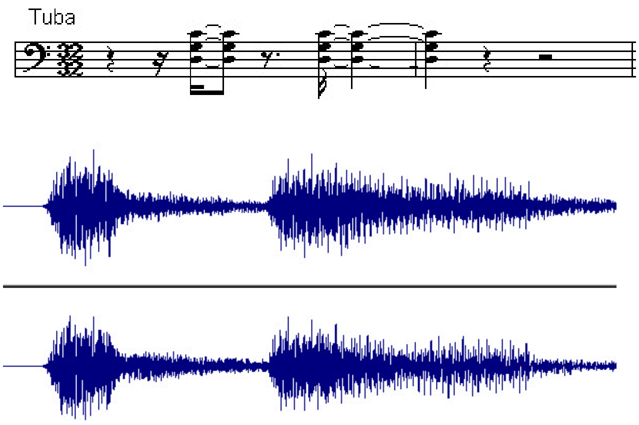

Figure 1. *Score and wave form of the Heavy Load sound.*

The turnings are derived from recordings of the original robot turning sound by transposition. They are always heard with any directional change between any kind of two complete moving-straight sections. If a complete moving-straight section consists of a number of straight segments of the same direction, the appropriate directional sound is repeated correspondingly without any Turning sound in between. A segment is defined as the straight connection between two neighbouring active decision areas; see Section 3.2.

The robot sound tracks actually used in the experiments of the supermarket scenario are the overlappings of the sound tracks of the intended robot trajectories and, during some of their segments, the additional sounds for robot states and situations as described in Section 2.2. In the second half of these experiments, the sounds of the real robot movements are also overlayed.

3. EXPERIMENTAL STUDIES

3.1 Understandability of Directional Sounds

The comprehensibility and the recallability of acoustic basic elements for humans was checked in the first experiments. Human subjects had to learn, understand and recall the auditory symbols of the eight directional sounds with their related meanings, as described in Section 2.1. They had to process altogether three test parts at the computer screen and heard the appropriate acoustic basic elements over loudspeakers. The experimental set-up was programmed in Delphi. A group of eight non-musicians and a control group of two professional musicians participated in the experiments.

In the first part of the experiments, all eight directions in each of the four variations are presented, thus, altogether 32

[2] http://www.ihrt.tuwien.ac.at/IHRT/English/mob_rob.htm (Laboratory for Mobile Robots – Univ.-Professor Dr. P. Kopacek)

acoustic basic elements, and that in random order. For each acoustic basic element, its respective meaning is displayed immediately afterwards.

For refreshing the subjects' memory, they are able to select all 8 directions, each in the 4 variations, again in the second part, now in any order. After each acoustic basic element, the subjects are asked for a subjective evaluation of its characteristics regarding Urgency, Expressiveness, and Annoyance; see Figure 2. For this purpose, they have to make a selection in each case, and mark three Seven-Point scales with respect to their subjective impression

from - - - (very small) to + + + (very urgent)

from - - - (very small) to + + + (very expressive)

from - - - (very small) to + + + (very annoying).

In the third part, the subjects shall then show, how well they can recognize all 8 directions independently of the presented 4 variations. They get a feedback after each of their decisions whether the respective acoustic basic element has been recognized correctly or wrongly.

All three test parts of this experiment have been repeated twice. During the second repetition of part 3, the non-musicians were between 34% and 100% (one subject) correct whereas the two professional musicians were 97% and 100% correct. The subjects have been informed in the written instructions that they should be able to retain the acoustic basic elements in their memory because the second experiment with acoustic displays in a robot-simulation environment will be based on these.

3.2 Intended and Perceived Robot Trajectories

In the second experiments, intelligible auditory symbols and sound tracks are presented in a supermarket scenario with a simulated (Windows 95 and Delphi) environment of a mobile service robot. It is assumed that the supermarket is open during seven days a week for 24-hours. A mobile service robot for cleaning and for carrying goods will inform the human subject (the customer) with sound symbols of non-speech auditory predictive displays about the trajectory of its intended movements and about the additional robot states and situations Heavy Load, Waiting, Near Obstacle, and Low Battery. These additional sounds for the robot situations have to be learned by the human subjects in a training phase at the beginning of the second experiments. The subjects can listen to these sounds in any order as often as they wish.

A floor plan of the supermarket is visualized on the computer screen. The human subject is shown in turquoise on the lower middle and the robot in different starting positions which are depending on the investigated trajectory; see Figure 3.

A matrix of decision areas has been constructed. Any intersection between a horizontal and a vertical line together with the respective nearest surrounding of this crossing, in which alternative routes can be chosen (beyond returning the same way), is determined as an active decision area in the visual floor plan of the supermarket; see Figure 3.

The human subjects are asked to recognize and to understand the intended trajectory of the robot indicating the intended directions where the robot plans to move on, as well as the overlapped additional sounds of the robot situations, from listening to the robot sound track. They have to draw the auditorily perceived trajectory into the visual floor plan of the supermarket on the computer screen and have to mark the perceived additional sounds. The subjects have been informed about the correctness of their auditory perception.

Altogether, the subjects perform four sub-experiments, each with four different trajectories. The intended trajectory and the perceived trajectory as well as the intended and the perceived additional sounds are recorded. Also, the durations of the training phase and of the drawing of each perceived trajectory are measured. The intended and the perceived trajectories can be compared in the replay mode; see Figure 3.

In the last two of these sub-experiments, the sound tracks of the trajectories are composed of the same sound symbols of the intended trajectories and the overlayed additional sounds for the robot situations. However, the sounds of the real robot movements are now also overlayed. This makes the scenario even more difficult for the subjects but it is also more realistic. In a real-world human-robot environment, the real robot movements are also always heard. They are presented in real time whereas the overlayed intended trajectories are auditory predictor displays and, thus, faster than real time.

The experimental results with the same eight non-musicians and the two professional musicians showed large differences in their auditory perception. The same three of these ten subjects who performed best in the first experimental study made only very few errors with the perception of the robot sound tracks of all trajectories as well as the additional sounds for the robot states and situations.

4. CONCLUSIONS

The design of directional sounds and robot sound tracks has been accomplished with basic musical elements and recorded robot noise signals. The experimental studies showed that the sound symbols and sound tracks are recallible and understandable, at least for more musical people. Positive training effects have been observed with all human subjects. The investigated digital-audio sound tracks are feasible means of communication in human-machine interaction. Further research needs to consider also other application domains.

5. REFERENCES

[1] Borys, B.-B., and Johannsen, G. "An experimental multi-media process control room," *Proc. Annual Conference 1997: Human Factors and Ergonomics Society Europe Chapter, Advances in Multi-media and Simulation*: 276-289, Bochum, 1997.

[2] Johannsen, G. "Cooperative human-machine interfaces for plant-wide control and communication," In J.J. Gertler (Ed.) *Annual Reviews in Control* 21: 159-170. Pergamon, Elsevier Science, Oxford, 1997.

[3] Stanton, N.A., and Edworthy, J. (Eds.) *Human Factors in Auditory Warnings*. Ashgate Publishing, Aldershot, 1999.

Figure 2. *Selection of eight directional sounds and their subjective evaluation.*

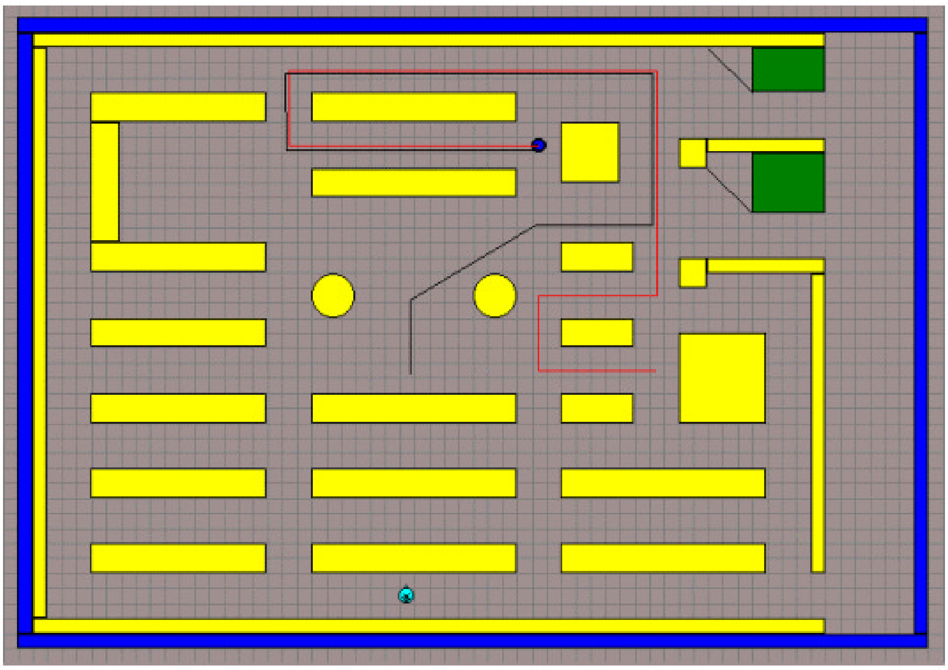

Figure 3. *Replay of intended and perceived trajectories in the supermarket scenario.*