# Giving Robots a "Voice": A Kineto-Acoustic Project

Ralf Hoyer[1], Andre Bartetzki[2], Dominik Kirchner[3], Andreas Witsch[3], M.J.G. van de Molengraft[4] and Kurt Geihs[3]

[1]Artist, Berlin, Germany
direkt@ralfhoyer.de
[2]Audio Communication Group, Technical University of Berlin, Germany
andre@bartezki.de
[3]Distributed System Group, University of Kassel, Germany
{kirchner,witsch,geihs}@vs.uni-kassel.de
[4]Control Systems Technology Group, Technical University of Eindhoven, Netherlands
M.J.G.v.d.Molengraft@tue.nl

**Abstract.** In this paper we present a kineto-acoustic project based on soccer robots. The movements of robots, determined by the needs of a soccer game, are transformed into a piece of music. Therefore, the robots are equipped with microphones, speakers, and custom-designed audio modules. The amplification of microphones and speakers is adjusted to create constantly varying feedback effects. These effects evolve from the relative positions and motions of the robots. Furthermore, data from control computers are utilized for the musical sound modulation. As the sequence of movements is not deterministic the resulting musical structure is unique in each performance.

**Keywords:** sound art, live electronic music, audio signal processing, robotics.

## 1 Introduction

The development of human-machine interaction is becoming increasingly important. Extensive research exists e.g. in the field of neuroscience with the goal to emulate human-like behavior and communication for autonomous robots. This is particularly the case in areas such as service robotics, health care, and search & rescue. Enhanced acceptance of these technologies among potential users has to be established, to take on the role as a helpful partner. However, it is also observed that users develop emotional ties to robots that behave independently and intelligently. It seems like the existence of quasi-human behavior also suggests the presence of human feelings. The resulting conflicts and misunderstandings are the basis for the musical and theatrical art project *RoboMusicTheatre* with robots and humans, which will be implemented progressively over the coming years. The project *MID SIZE ROBO SOCCER MUSIC* will expand one scenic idea to an independent kineto-acoustic project for a prestigious international art festival[1]. The foundation for the sound structure is provided by

[1]Festival *Inhuman Music - Music by Accident, Animals and Machines*, taking place from 21st to 24th February 2013 in Haus der Kulturen der Welt in Berlin.

autonomous behaviors, team coordination, and dynamics of the robotic systems. Therefore, the performance is realized by two robotic soccer teams, which generate acoustic feedback loops.

The paper is structured as follows. Section 2 describes the artwork and the concept of the kineto-acoustic robotic performance. The following sections state technical details of the realization. The audio processing and the development of the needed hardware are presented in Section 3. Subsequently, key characteristics of the applied robotic technology are identified and described. The paper concludes with a discussion and outlook to future work.

## 2 Artistic Concept

Every sound is the result of motion. The sounds and tones generated by our everyday high-tech environment may be largely irrelevant, often annoying, and the same time inescapable. Nevertheless, sound is the objective of every motion in acoustic arts. However, the type of tone generator is immaterial.

Links exists between sound and motion at an even more fundamental level, for instance with electro-acoustic feedback. In this case, the frequency of the feedback is determined by the distance between microphone and speaker, the resonances in the room, and all devices in the electro-acoustic chain [1].

With this in mind, we developed the idea of using robots of the *Middle Size League* of the *RoboCup* initiative [7] to generate a polyphonic texture of sound. The emerging sound is audibly linked to the motions of the robots, seeming to give them a "voice" while moving as independent players across the soccer field.

The robots are equipped with microphones, speakers, and a custom-designed audio module to form this "robotic voice". The robot motions and individual positions lead to varying feedback effects due to the microphone amplification. Additionally, other data from the robot computers are evaluated for the sound modulation. One example is the distance to other robots in conjunction with the speed of movement. Thus, the movements of robots, determined by the flow of a soccer game, are transformed into a piece of music. Due to the unique motion flow, the sound cannot arise in any other manner and will be different in each performance.

## 3 Sound Processing Technology

The application of networks for sound production, musical composition, and performance has a long tradition in experimental music and sound art. Related examples are the *League of Automatic Music Composers* [1] of the late 1970s or the contemporary laptop orchestras [2]. While these are still based on a traditional ensemble of individual human musicians, an ongoing evolution of concepts and pieces of art exists, where humans are only partially involved or totally absent. Without human musicians or organizers, different ways to structure music and sound production are necessary. One of the main principles of self-organization in nature is the feedback loop, which has become popular in technology based music and sound art for many decades [3]. Even some of the basic components of sound production use a feedback structure e.g., oscillators and filters.

Self-organized systems based on autonomous agents utilize feedback-loops to create emergent behavior. If the resulting actions incorporate audible effects, the acoustical output and the corresponding behavior like motion patterns might be regarded as speech, musical gestures, or structures.

In order to create a musical structure and to form an ensemble out of the team of machines, we developed a distributed system for sound processing based on portable audio modules and a structure for communication.

### 3.1 Hardware Architecture of the Audio Modules

Each mobile robot is equipped with an audio module, which can receive, process, and send audio signals or control data. The audio modules work autonomously but are connected to the hosting platform to allow reception of allocentric position and velocity. In addition, there is a main audio processing unit which can send and receive control data from the audio modules of all mobile robots.

The main component of the audio module is a single-board computer (BeagleBoard xM [4]) with a 1 GHz ARM CPU. A small electret microphone capsule perceives the sound of the environment. The microphone is connected to a built-in audio input via a pre-amplifier. This pre-amplifier includes an adjustable noise gate which helps to separate the sound of the other robots from the background noise. After being processed by the software running on the board, the audio signal gets amplified and dispensed by the speaker.

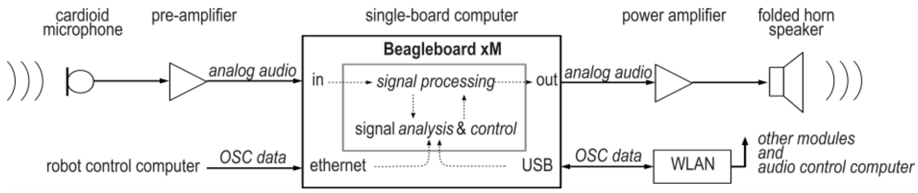


**Fig. 1.** Hardware architecture of the developed audio module.

The directivity patterns of both, the microphone (cardioid) and the speaker (folded horn), were chosen to minimize the probability of direct feedback between audio output and input of the module itself. For the same reason microphone and speaker are aligned to opposite directions. The horn shape of the speakers can be regarded as a social and aesthetical reference to the megaphones and Vuvuzelas common for football games spectators in stadiums. In order to receive context information from the hosting mobile robot an Ethernet connection is used. All audio modules are able to share features of the sound analysis and context information of the hosting platform with the central audio process via wireless network. These data are transferred using the Open Sound Control (OSC) [5] protocol.

### 3.2 Signal Processing Software and Communication

The signal processing software is written in SuperCollider (SC) [6], a programming language and run-time environment for sound synthesis, signal processing, algorithmic musical composition, and live-coding. The SC-environment consists of two components: The *server* to perform the audio signal processing and the *language* that controls processes on the server. The server is able to communicate with other instances of SC or other software on remote machines using the OSC protocol. Fig. 2 outlines the structure of the developed signal processing for this project. The audio processing of the different parts is realized as follows:

- **Preprocessing:** Due to the noisy environment the incoming raw signal has to be filtered to eliminate the noise using band-pass filters and noise-gates. We automatically adjust the level and the spectrum of the processed input signal to suppress direct feedback.
- **Feature extraction:** Sound features have to be extracted, like amplitude (envelope follower), fundamental pitch (autocorrelation pitch follower), spectral shape (spectral centroid, spectral flatness, spectral percentile), etc. These features are communicated to the other audio modules and form the basis to adapt distributed sound processing.
- **Dynamics:** In order to create an audible response dynamically to the audio input within the context of the environment, the received features are used for the selection and the control of the available sound processing algorithms.

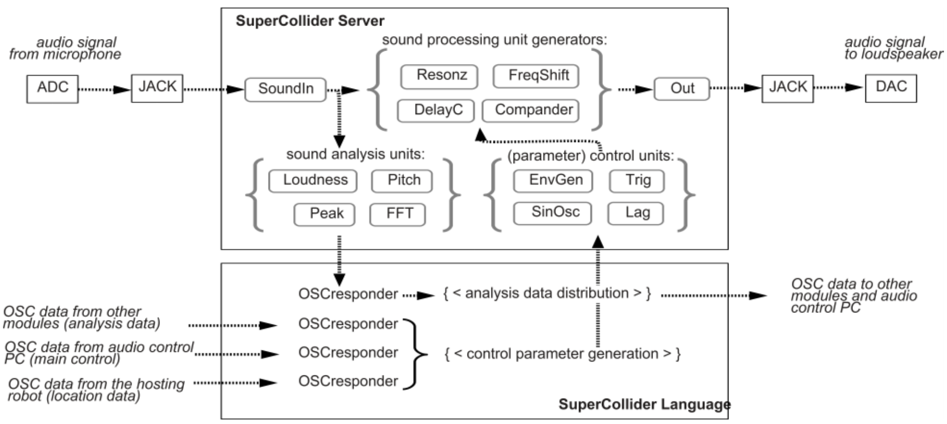


**Fig. 2.** Structure of the signal processing software SC

This framework allows the transformation of environment dynamics into an acoustic performance. In order to realize the dynamics, which result in artificial composition, autonomous mobile components are needed.

## 4 Robot Technology

The concept of the project with the goal of a polyphonic sound structure demands sophisticated technological requirements for the underlying mobile platform.

- **Autonomous system behavior:** The technology should feature a human independent (autonomous) operation.
- **Team Coordination:** Music typically consists of an interference of multiple sounds. Therefore, the applied technology should provide multiple interacting system components (agents) to synthesis this.
- **Mobility:** The dynamic of the sound structure is based on vibrant distance changes. Therefore, mobility is a key requirement of the platform.

A RoboCup MSL team of autonomous mobile robots meets these requirements. The robot size of 52x52x80cm and the ability to carry sufficient payload allows attachment of the described audio modules. Additionally, the 12x18m field size provides ideal conditions for the feedback generation with respect to the wave length of audible sound.

Two MSL teams, namely Tech United [8] from the Technical University of Eindhoven and Carpe Noctem [9] from the University of Kassel provide the robots for the performance (Fig. 3). Beside the available MSL platforms, a MSL soccer game provides a well-established setup for the project. Such a game features the dynamic game flow needed for the sound composition of the project. Both teams annually participate in MSL competitions and provide the expertise for the technical infrastructure. Following, a short excerpt of the robotic technology relevant to the stated requirements is given. More details can be found in [10] and [11].


**Fig. 3** Left: Typical situation in a MSL game (Tech United versus Carpe Noctem). Right: MSL robot from Carpe Noctem [9].

### 4.1 Autonomous System Behavior

A characteristic of autonomous mobile robots is their ability to move and act independently in their environment. In case of soccer robots the relevant context of the environment is characterized by positions and velocities of the robots, the ball, the teammates, and the obstacles. As these quantities are real valued, the number of possible states is apart from precision limitations infinite. This fact makes it impossible to model all situations explicitly. Even the simulation of a wide variety of situations is too computational expensive to cover the search space sufficiently with current computers.

Therefore, we need team behavior, which is able to generalize and is usually realized by techniques known from control theory. Their goal is to converge to a global control target, like approaching the ball or to score a goal. A control cycle in this case is divided in three parts. First, sensing the environment e.g., by a camera and extracting high-level information to specify the current situation. Second, selecting the robots behavior, this includes the team coordination process. Third, the chosen motion command has to be transformed to low-level actions like motor velocities to execute the selected behavior.

As the global system – including own and opponent team dynamics – is too complex, we cannot use a model in order to prove convergence. Thus, the flow of a soccer match can be assumed as a chaotic system with some fix points. Practically this means that two similar situations with equal appearance mostly result in different outcomes. This is even true without the incorporation of learning or randomized behavior. In a MSL soccer game, e.g. fix points are manifested by the half-time.

However, to reach the goal to give the robot a "voice" this behavior is beneficial, as we can achieve highly complex team dynamics and unforeseen situations. Combined with the dynamics of the sound generation as described in Section 3, the generated sound structure appears as an expression of musical creativity formed by the robots.

### 4.2 Team Coordination

In order to enforce a convergent strategy, we realize the plan and behavior selection step by a strategy description that allows modeling from a global perspective. Therefore, we use ***A** **L**anguage for **I**nteractive **C**ooperative **A**gents* (ALICA) [12]. The language describes tasks for an arbitrary number of robots within a plan. A plan consists of a set of states connected by transitions similar to a state machine. Every state can in turn include a plan or an atomic behavior, like *kick* or *move to point*. For each transition a condition has to be specified that triggers a state-switch if evaluated as true. ALICA achieves synchronization either explicitly by synchronized transitions, or implicitly based on the underlying common knowledge base, which is in case of the Carpe Noctem robots the knowledge of active states of the other robots and the situation specified in the world model. As the world model shares not only local information of the robot itself but the common belief of the world state, synchronous transition triggering can be achieved.

The communication between the different robot processes is realized by using the ***R**obot **O**perating **S**ystem* (ROS) [13]. It provides a middleware for inter-process communication, build chain, and tool support for debugging of robot platforms. Thus, ROS supports the challenging task to maintain robots.

Using this framework we specified a plan structure for playing soccer with the control target to score goals. Furthermore, it determines regulations for the chaotic game flow as described in Section 4.1 and therefore the dynamics of the resulting sound structure.

### 4.3 Mobility

The stated autonomy results in position changes, which form the foundation of the envisioned dynamic polyphonic sound structure. Therefore a high degree of mobility supports the dynamics in this sound structure and is an essential part of the projects realization. As velocity and agility are inherent elements in MSL competitions, a lot of effort is spent from the community to realize and optimize drive systems of the used robots. The technology of omnidirectional drives [14], realized with multiple (typically three to four) omnidirectional wheels (see Fig. 3), has been established as a quasi-standard in MSL RoboCup to improve robots maneuverability [8] [9] [14].

The wheel consists of a series of rollers attached to its lateral surface that are perpendicular to the rolling direction. This allows an almost frictionless translation in direction of the rotational axis of the wheel.

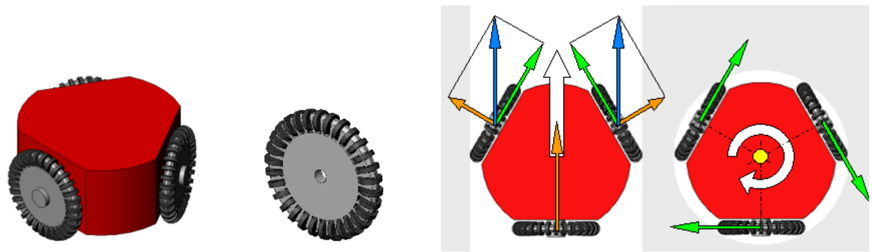

**Fig. 4.** Left: omnidirectional drive system in 120° configuration. Middle: Model of an omni wheel used in the omnidirectional drive system. Right: Omni drive with different global movements (white arrows) and the needed wheel velocities (green: propulsion velocity, orange: slide velocity, and blue: effective velocity) [15]

Each wheel velocity is individually controlled. The combination of all wheel velocities defines the global robot movement dependent on its geometry. For some illustrative robot movements (white arrows) the required wheel velocities (green arrows) are depicted in Fig.5. A detailed description of the control mechanisms is presented in [14] [10]. Using such a drive mechanism allows the robot to move in arbitrary directions and independently rotate. This results in a drive system that allows very agile movements and enables the dynamic position changes required for the polyphonic sound generation in this project.

## 5 Conclusion

This paper presents a kineto-acoustic project that encourages the reflection about emotional ties in an autonomous setting (see video [16]). Mobile autonomous soccer robots and distributed polyphonic sound processing are the two technologies to realize the artwork. The sound processing units mounted on a Middle Size League robot are equipped with a microphone and a speaker. The distances between the robots and their velocities determine the modulation and amplitude of various acoustic feedback loops. Other data can be evaluated for sound processing.

The resulting polyphonic "voices" of the robots build a texture of sound. This has a structure that can be appreciated as a form of avant-garde electronic music. Usually, these kinds of artworks have human creators and inherently bear human expression. The listener of the presented performance may infer expressive musical values in this music, even though it was autonomously generated by technology.

In this performance sound originates from motion. In future work we plan to investigate the reverse order. Therefore, transformations have to be identified that determine the movements to generate specific sounds. The goal is to allow robots to generate musical structures, sounds, and tones in the same autonomous fashion they solve soccer tasks. This would take us beyond music into the area of choreography.

## References


1. Reich, S.: Pendulum Music. In: Writings About Music (pp. 12–13). The Press of the Nova Scotia College of Art. ISBN 0-919616-02-X 1.(1974)
1. Föllmer, G: Linien der Netzmusik, In: Neue Zeitschrift für Musik, 5/2004
2. Toplap, http://toplap.org, [Online; accessed 23-September-2012]
3. de la Motte-Haber, H: Systeme mit Selbstregulation, In: Kunsttexte.de 3/2011, Auditive Perspektiven, www.kunsttexte.de [Online; accessed 23-September-2012]
4. beagleboard.org, http://beagleboard.org, [Online; accessed 24-September-2012]
5. opensoundcontrol.org, http://opensoundcontrol.org, [Online; accessed 24-September-2012]
6. SuperCollider, http://supercollider.sourceforge.net, [Online; accessed 24-September-2012]
7. Kitano, H., Asada, M., & Kuniyoshi, Y. (1997). RoboCup A Challenge Problem for AI. AI Magazine, 18(1), 73–85.
8. Kirchner, D., Witsch, A., Saur, D., et al. (2012). Carpe Noctem – Team Description Paper.
9. Hoogendijk, R., Merry, R.J.E, et al.: Tech United Eindhoven Team Description 2012.
10. Siegwart, R., Nourbakhsh, I. Introduction to Autonomous Mobile Robots (p. 321). Cambridge: MIT Press. (2004).
11. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach (Second., p. 1081). Pearson Education. (2003).
12. Skubch, H., Wagner, M., Reichle, R., Geihs, K.: A modeling language for cooperative plans in highly dynamic domains. In: Mechatronics 21, pp. 423-433, (2011)
13. Quigley, M., Gerkey, B., et al.: ROS: an open-source Robot Operating System. In: ICRA, Open-Source Software workshop, (2009)
14. Rojas, R., Förster, A. Holonomic control of a robot with an omnidirectional drive. In: KI-Künstliche Intelligenz, 7., BöttcherIT Verlag, (2006).
15. Omnidirectional drive, http://de.wikipedia.org/wiki/Omnidirektionaler_Antrieb, [Online; accessed 26-September-2012]
16. Kineto-acoustic project video, http://www.das-lab.net/content/giving-robots-voice-kineto-acoustic-performace, [Online; accessed 04-December-2012]