MULTIMODAL CLOSED-LOOP HUMAN MACHINE INTERACTION

Tobias Großhauser, Thomas Hermann

Ambient Intelligence Group
CITEC - Center of Excellence in Cognitive Interaction Technology
Bielefeld University, Bielefeld, Germany
tgrossha, thermann@techfak.uni-Bielefeld.de

ABSTRACT

The paper presents a multi-modal approach to tightly close the interaction loop between a human user and any tool in operation. Every activity of a human being generates multi-modal feedback, more or less related to the eyes (visual), the skin (sensory), the nose (olfactory) and the ears (auditive). Here we show the useful augmentation or complete creation of a nonexistent or less available feedback. As an example the performance of drilling tasks, line drawing tasks, or the complex task of bowing a violin can be considered. Some new multimodal human computer interaction technologies based on sensors and embedded systems are shown and described in this paper.

1. INTRODUCTION

Every-day or highly skilled activities have in common that for the correct execution a movement activity needs to be carried out at high accuracy in response to perceptions as they occur in real-time during the performance. While in real-world situations a mixture of senses interplays for us to generate stimuli at hand of which we can learn to coordinate and refine our actions, for some tasks certain modalities may be missing such as sound in drawing tasks and dance, or vision in drilling tasks. Or they are faint, for instance the deviation from a linear bowing movement of musical string instruments. In this paper the supportive function of feedback in different scenarios should be outlined in the meaning of man-computer symbiosis in everyday and highly skilled learning tasks. As it already was stated 1960 by Licklider [1] regarding to intellectual operations, here also operations are performed more effectively and learning processes are shortened by useful tool integrated interfaces and audio or audio-haptic feedback. Sonification and vibrotactile feedback in embedded and wearable devices show new possibilities in the field of multimodal human computer interaction. Especially in every-day and working situations, where traditional interfaces like monitors and keyboards would disturb the used “work-flow”, the embedded wearable devices provide many solutions. This are on the one hand new input possibilities with sensors like distance, pressure, acceleration sensors and gyroscopes, video cameras and microphones and on the other hand ubiquitous and adaptable output possibilities like loudspeakers and vibrotactiles.

2. FEEDBACK TYPE AND DESIGN

The feedback loop here means a system or signal that generates an output, detected by sensors to control the system or tool within itself or the human, reacting to the output. There are different approaches to the design of feedback. Bill Verplank’s more practical approach in his “Interaction Design Sketchbook” [2] with the basic question “How do you do? How do you feel? How do you know?” model of interaction describes a simple feedback loop. He states that “even the simplest appliance requires doing, feeling and knowing” which is clear if you think about e.g. the flipping of a switch or opening a door. In our system, the feedback loop between one or more humans and the computer is considered.

2.1. The Sonification Modes

Three main classifications of sonification are described by de Campo [3] are described. Sonification by “Continuous Data Representation”, by “Discrete Point Data Representation”, and by “Model-Based Data Sonification”, Hermann et al. [4]. The system described here provides real-time feedback in an acoustic and tactile form by means of interactive sonification of Hermann [5] and haptic feedback. Information is conveyed acoustically as well as haptically and by useful combinations of both.

2.2. The Applied Sonification Modes

We discern two different sonification types according to the directness of auditory feedback.

1. Continuous Sonification: This method, demonstrated e.g. by [6] and in an experiment with a rolling ball on a tiltable track of Rath [7], allows the continuous control of a movement or task in real-time. The movement, level or position of the tool is translated directly into a sound feedback. As shown in Fig. 2, this is done either by direct amplification of the sound of the tool or task itself or by rendering a synthesized sound.
2. Case-Triggered Sonification: This means that the sound is only triggered, when a certain problem or deviation appears. The sonification can be changed and turned on and off manually, so the user has permanent control. This allows the individual assignment of a specific sound or sound effect to each sensor, condition or tool, or to group useful sensor combinations. This could for instance be useful, if you use many tools at the same time, like in repair shops or operating rooms.

2.3. Sound Synthesis

Different sound synthesis models exist in the area of music technology to generate sound and music. Beside the analog sound synthesis, various additional digital synthesis methods exist. The most common ones are subtractive and additive and frequency modulation synthesis. Further synthesis methods are granular, wavetable, phase distortion, sample-based and physical modeling synthesis. Many parameters can be influenced by sensor input, such as pitch, volume and number of tones are changed according to it.

2.4. Vibro-tactile Feedback

In some situations the visual sense is occupied, the surrounding or the used tool is too loud in this cases the vibrotactile feedback is then an useful display to support the person executing a special task. One well known example of a mechanical audio-haptic feedback is the torque wrench, where you feel and hear, when reaching the adjusted torsional. In our example, the sensors, electronics and feedback is all integrated into the used tool. This is called the "tool-integrated sonification", in contrast to many other examples, where e.g. the measuring and calculating part is done on a stand-alone computer. Some recent projects show that tactile feedback is a meaningful possibility to extend existing tools, like in Grosshauser et al. [8] a violin bow and even for vision sensory substitution, like in Bird et al. [9]. In this two examples, the closed-loop tactile feedback fits perfectly in this discreet way of indication and can support body awareness over a long period of time. Beside the mechanic feedback of these used tools and that we will present later in this paper, the signals of the used mechanic vibrotactiles create passive touch cues, which are presented to the observer's skin, rather than felt in response to active movements, similar to Gibson [10].

2.5. Embedded Multi Channel Audio

Our multimodal approach, here exemplified with a cordless screwdriver, uses a 3 channel audio system (see Fig. 4). The 3 tiny loudspeakers are attached directly on the housing of the screwdriver. More loudspeakers can be used, but 3 is the minimum to indicate the direction of deviation, here the wrong angle relating to the wall, and the direction of the required movement to readjust the angle. Three-dimensional adjustment is made easier, even without looking at the screwdriver and enables e.g. blind people to "hear" the right angle relating to the wall.

2.6. Definition of Closed Loop Systems

According to Dubberly et al. [11], a "Closed-Loop-System" (see Fig. 1), does not only react and act linear, it also provides feedback to the user. In our case of audio-haptic feedback, either completely new multimodal feedback signals are generated or existing ones are amplified and manipulated. That allows in our examples e.g. to better learn coordinated activity for complex tasks.

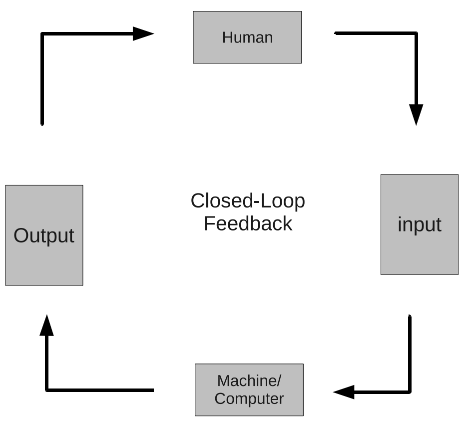

Figure 1: *Closed loop feedback scheme*

In the feedback loop in Fig. 1, the data flows from a system to a person or user and back through the system again. Adjustments are done, to achieve a specific goal, in reaction to the information from the feedback system, which is reading and comparing the sensor data. The latter depend on the used sensors and are influenced by the environment and the action of the user. Then the loop is closed and can start from the beginning.

Also a simple automated self-correcting system is integrated, meaning that under certain conditions, the system can influence or regulate itself. Non self-regulated systems are called "open loop", regulated systems are called "closed loop". The natural cycle of water for example is an open loop system, as there is no regulation about the amount or location, where it should rain or evaporate. A closed-loop system (see Fig. 2) is, for example, if the tool or machine is switched off automatically, if a certain situation occur. A more complex scenario could be, that the user leaves the correct plane or angle, the system then generates an acoustical warning, but the sound is not loud enough. The system senses that the human does not react. Now the volume has to be increased. Here the system is also regulating itself, "self-regulating", but the difference in the data and the adaptive regulation influences the state of the machine or the output directly. This is a simple self-correcting system and in more technical terms a so-called first-order cybernetic system. At the end, the machine influences the sensors, the sensors the input, and the loop is closed again.

2.7. Definition of Interaction

But is the above example really interaction? Interaction, in contrast to reaction, means according to Dubberly [11] "the transfer function is dynamic, i.e., in 'interaction' the precise way that 'input affects output' can itself change; moreover in some categories of 'interaction' that which is classified as 'input' or 'output' can also change, even for a continuous system." In our developed device, there is not only a linear coherence between "input" and "output", so the system changes itself. This means, the system does not only react, it interacts.

3. TECHNICAL SETUP AND DESIGN

3.1. The Sensors

In our exemplary use cases we use a set of many different sensors. The data from the sensors are transmitted via radio frequency or processed directly on the I/O-board. A small Lithium Polymer (LiPo) battery is directly attached for power supply. The H-bridge is an integrated electronic circuit, to apply a voltage to the vibration motors and changes the speed. Increased speed implies more urgency and attention, lower speed feels more soft. This small and light-weight sensor module can be used as a stand-alone tool, just for movement learning, or it can be clipped to any tool.

3.2. Acceleration and Tilt

An IDG-300 dual-axis angular rate gyroscope from InvenSense is used. This allows the measurement of the rotation of the x- and y-axis of the bow stroke. The x-axis rotation is an additional compensating motion for e.g. soft bowing starts. The y-axis rotation is besides other functions relevant for pressure transfer onto the bow and to balance and change articulation and volume.

The ADXL330 acceleration sensor from InvenSense is used, a small, thin, low power, complete x-, y-, and z-axis accelerometer. For the following description, every axis is important and has it's own defined plane, in which the movement is performed. Thinking in planes and rotations helps to learn complex movements, especially when the movement takes place beside your body and you the player hardly see it or control it visually.

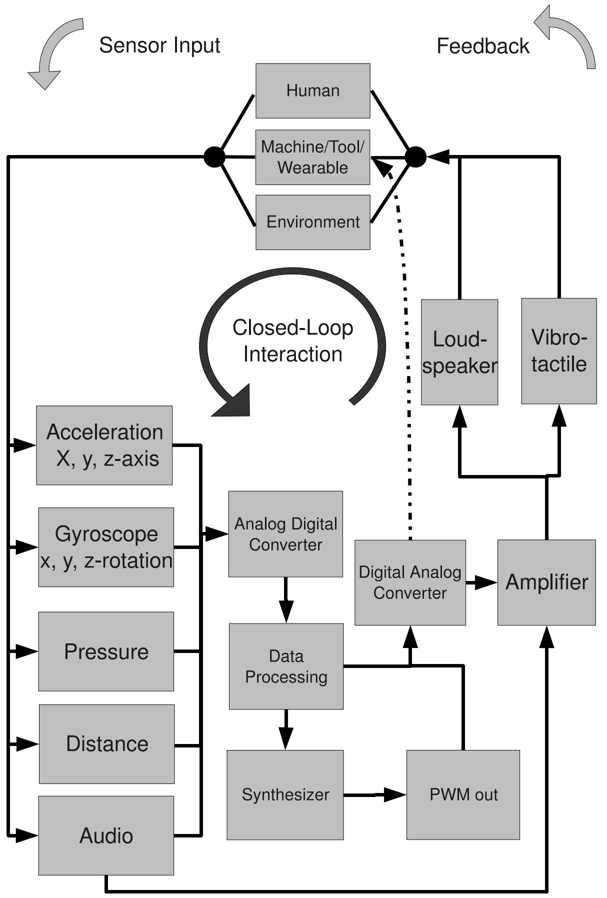

Figure 2: *interactive closed loop feedback scheme*

3.3. Goniometers

A goniometer is an instrument which measures an axis and range of motion, or the angle or rotation of an object precisely about a fixed axis between two connected elements.

Goniometers with a potentiometer are used for joint angle measuring. This is a very precise, cheap sensor and it is easy to fix and install. It can be fixed directly on the body or into the clothing, depending on how precise the measurement has to be.

3.4. Distance

Fig. 3 shows an infrared proximity sensor 2y0a21 made by Sharp. It has an analog output that varies from 3.1V at to 0.4V at a distance up to 40cm. The signal voltage is higher at close range, and decreases as the range increases. In the example below of a screw driver, drilling depth and drilling angle can be measured.

3.5. Pressure and force sensors

Similar to Koehly et al. [12] paper based force sensitive resistors (FSRs) made out of paper are used. The technique was first presented by Jensenius et al. [13] for use in low-cost music controllers. Black art paper dyed with carbon particles conducts electricity and its resistance depends on the applied pressure. It is cheaply available and easy to process. The pressure sensor is 5 x 5 x 0,5 mm, it weights only some grammes, depending on the dimensions of the surface area. In the violin example pressure sensors are used, in the drilling task pressure sensors are combined with simple switches.

Figure 3: *Sharp distance sensor*

4. EXAMPLES

In this section, we describe our approach at hand of two example scenarios. The first example is the support of learning special movements in violin playing. The second application is an augmented screw-driver as an example of an every-day used tool.

4.1. Highly skilled tasks

There are many applications of tool supported or tool based highly skilled tasks. This could be a scissors or scalpel or a musical instrument. In most of these applications, the linear guidance of an object in 2D or 3D space is necessary. This could occur in bowing movements or while guiding a scissors or scalpel in a surgery. A recognition of jitter or deviation from a given line could be indicated by interactive sonification or tactile feedback. In the following, a short example of the tool-mounted feedback in the field of musical instrument learning is shown.

4.2. Musical instrument learning

The mixture of acceleration, pressure and goniometer-based sensing allows the precise measurement of a violinbow, and thereby and exercising without the musical instrument. Similar to Grosshauser et al. [14] the feedback is directly and interactive according to the movement of the arm or bow.

The following scenarios are basic extractions of beginners' violin lessons. Depending on the age of the pupil or student, different approaches exist. One of these is the breakdown and fragmentation of a movement into several simpler action units, based on the ideas of Conrad von der Goltz [15]. In our scenarios, a simple bow-stroke is decomposed. This is even trained from time to time by advanced students and professional musicians to develop their skills and physical awareness. Similar methods of deconstructing complex movements exist in the areas of dance and sports. The sensor and the real-time sonification gives us the possibility to train these simplified movements and adding step by step more and more complexity. In other words, this means the combination of simplified movements to more complex ones. The single and combined movements in the following cases can be performed simultaneously or successively, with or without instrument.

Problem: Adding a second plane, the y-plane with zero deviation of the y-axis to the exercise, drawing a virtual straight line.

Pedagogical aspect: Understanding the "virtual straight line" of bowing movement.

Idea: If you move your hand exactly along one direction so that you draw a perfect line into the air beside your body, complex compensating movements of the hands and arms are necessary. If you try this with a pupil the first time, it is not only hard to understand the movement without seeing your hands, also practicing in front of a mirror is difficult, because every change has to be side-inverted.

Result: Students learn to move the hand on defined straight lines, without looking to it.

Concerning the described issue, the sonification is provided directly from the position, where the fault occurs, e.g. on the frog of the bow. in different ways. The spatial position is defined according to our hearing experience through the sound source directly integrated into the sensing area. The deviation of a given plane or constancy of the movement is observed real-time.

4.3. Everyday tasks

We present every-day task sonification for task such as like drilling, using a cordless electric screwdriver. These are situation, where it could be useful to add/support or replace the visual sense. Especially while drilling and screwing, to see from different views, if the drilling machine is horizontally and vertically in the right position, mostly a 90 deg. angle to the wall. Also all other angles can be obtained by presetting them. A demo video showing a sound augmented drilling machine support is available on our website at http://www.techfak.uni-bielefeld.de/ags/ami/publications/GH2010-MCL.

In many tools, the loudspeaker have the advantage of a small form factor than display, which facilitates the mounting and integration. Especially in drilling situations, intervisibility is not always possible and auditive cues guide and support the user to fullfill the task, even in difficult situations.

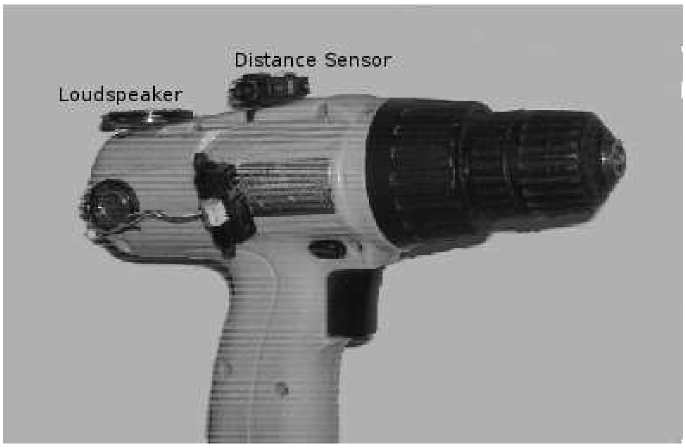

Figure 4: *Picture of drilling machine with distance sensors and 2 of the 3 loudspeakers*

5. DISCUSSION

In this paper an easily relocatable flexible sensor based system is presented for motion capturing and multi modal real-time feedback. Simple usage, even without the need of an external computer is possible. In this contribution two different sensor setups have been presented to demonstrate the possibilities and usage in typical everyday training situations.

Different sensors have been applied to sense task-relevant information. The sensors can be directly mounted on any tool and are coupled with an interactive sonification and 'haptification' approach, utilized to give real-time feedback. The idea of fitting up every-day and special tools with unobtrusive additional features to simplify or augment their usage opens up an interesting application field - especially as embedded sonification in combination with loudspeakers is feasible with very small form factors. And last but not least it is cheap, especially simple loudspeakers are in the price range of some cents and thereby much cheaper than LCD displays or similar feedback devices.

We plan to conduct long-term user studies with this prototypes and we currently investigate more scenarios of competitive and useful closed-loop audio-haptic feedback. Finally, we are very convinced that we can easily adapt the system to other everyday activities and even to other highly skilled fields such as movement training in sports, e.g. the smooth shift of body balance as it is demanded in movements from Tai Chi or in dance, which is also of relevance in case that the equilibrium sense is impaired. Many further scenarios are imaginable, where the closed-loop feedback system can help to better learn, understand and perform complex movements.

6. ACKNOWLEDGEMENT

The work has been supported by CITEC – Center of Excellence in Cognitive Interaction Technology.

7. REFERENCES

[1] J. C. R. Licklider, "Man-computer symbiosis," in *IRE Transactions on Human Factors in Electronics*, 1960, vol. HFE-1, pp. 4–11.

[2] *Interaction Design Handbook*, Bill Verplank, https://ccrma.stanford.edu/courses/250a/lectures/IDSketchbok.pdf, 2003.

[3] Alberto de Campo, "Toward a data sonification design space map," Montreal, Canada, 2007, Schulich School of Music, McGill University, pp. 342–347, Schulich School of Music, McGill University.

[4] T. Hermann and H. Ritter, "Listen to your data: Model-based sonification for data analysis," in *189–194, Int. Inst. for Advanced Studies in System research and cybernetics*, 1999, pp. 189–194.

[5] Thomas Hermann and Andy Hunt, "Guest editors' introduction: An introduction to interactive sonification," *IEEE MultiMedia*, vol. 12, no. 2, pp. 20–24, 2005.

[6] Guillaume Lemaitre, Olivier Houix, Yon Visell, Karmen Franinović, Nicolas Misdariis, and Patrick Susini, "Toward the design and evaluation of continuous sound in tangible interfaces: The spinotron," Duluth, MN, USA, 2009, vol. 67, pp. 976–993, Academic Press, Inc.

[7] Matthias Rath and Davide Rocchesso, "Continuous sonic feedback from a rolling ball," Los Alamitos, CA, USA, 2005, vol. 12, pp. 60–69, IEEE Computer Society Press.

[8] Tobias Grosshauser and Thomas Hermann, "Augmented haptics - an interactive feedback system for musicians," in *Haptic and Audio Interaction Design, 4th International Conference, HAID 2009, Dresden, Germany, September 10-11, 2009, Proceedings*, M. Ercan Altinsoy, Ute Jekosch, and Stephen A. Brewster, Eds. September 2009, vol. 5763 of *Lecture Notes in Computer Science*, pp. 100–108, Springer.

[9] Jon Bird, Paul Marshall, and Yvonne Rogers, "Low-fi skin vision: a case study in rapid prototyping a sensory substitution system," in *BCS HCI '09: Proceedings of the 2009 British Computer Society Conference on Human-Computer Interaction*, Swinton, UK, UK, 2009, pp. 55–64, British Computer Society.

[10] J. J. Gibson, "Observations on active touch," *Psychol. Rev., vol. 69, no. 6, pp. 477–490*, 1962.

[11] Hugh Dubberly, Paul Pangaro, and Usman Haque, "What is interaction? are there different types?," *Interactions*, vol. 16, no. 1, pp. 69–75, 2009.

[12] Rodolphe Koehly, Denis Curtil, and Marcelo M. Wanderley, "Paper fsrs and latex/fabric traction sensors: methods for the development of home-made touch sensors," in *NIME '06: Proceedings of the 2006 conference on New interfaces for musical expression*, Paris, France, France, 2006, pp. 230–233, IRCAM — Centre Pompidou.

[13] Alexander Refsum Jensenius, Rodolphe Koehly, and Marcelo M. Wanderley, "Building low-cost music controllers," in *CMMR*, 2005, pp. 123–129.

[14] Tobias Grosshauser and Thomas Hermann, "The sonified music stand – an interactive sonification system for musicians," in *Proceedings of the 6th Sound and Music Computing Conference*, Alvaro Barbosa Xavier Serra Fabien Gouyon, Ed. Casa da Musica, Porto, Portugal, 2009, pp. 233–238.

[15] Conrad v. d. Goltz, *Orientierungsmodelle für den Instrumentalunterricht*, G. Bosse Verlag, Regensburg, Germany, 1974.