WEARABLE SETUP FOR GESTURE AND MOTION BASED CLOSED LOOP AUDIO INTERACTION

Tobias Großhauser
Ambient Intelligence Lab
CITEC - Cognitive Interaction Technology
Bielefeld University
tgrossha@techfak.uni-Bielefeld.de

Thomas Hermann
Ambient Intelligence Lab
CITEC - Cognitive Interaction Technology
Bielefeld University
thermann@techfak.uni-Bielefeld.de

ABSTRACT
The wearable sensor and feedback system presented in this paper is a type of audio-haptic display which contains onboard sensors, embedded sound synthesis, external sensors, and on the feedback side a loudspeaker and several vibrating motors. The so-called “embedded sonification” in this case here is an onboard IC, with implemented sound synthesis. These are adjusted directly by the user and/or controlled in real-time by the sensors, which are on the board or fixed on the human body and connected to the board via cable or radio frequency transmission. Direct audio out and tactile feedback closes the loop between the wearable board and the user. In many situations, this setup can serve as a complement to visual output, e.g., exploring data in 3D space or learning motion and gestures in dance, sports or outdoor and every-day activities.

A new metaphor for interactive acoustical augmentation is introduced, the so-called “audio loupe”. In this case it means that the sonification of minimal movements or state changes, which can sometimes hardly be perceived visually or corporeal. This is for example small jitters or deviations of predefined ideal gestures or movements.

Our system is easy to use, it even allows operation without an external computer. In some examples we outline the benefits of our wearable interactive setup in highly skilled motion learning scenarios in dance and sports.

1. INTRODUCTION
This project presents wearable sensing, embedded sonification and vibrotactiles, in the form of a wearable audio-haptic display. The auditory displays, according to Kramer, p. 7-10 [1], provides several advantages and even more in conjunction with other displays, in our case tactile feedback. These benefits are, according to [1] p. 9, just to name a few, “increase in perceived quality, enhanced realism, learning and creativity and lower computational requirements”.

Further the system provides real-time feedback in an acoustic and tactile form by means of closed-loop interactive sonification (see fig. 1) according to Hunt [2], Hermann [3] and haptic feedback according to MacLean [4] and Morris [5]. Information is conveyed acoustically as well as haptically and by useful combinations of both.

Our setup is a new approach and method for movement and posture measurements in 3D-space. The wearable sensing technologies further provides the user the possibility of on-site and real-time measurements of movements and postures and embedded data analysis. Many established motion capturing technologies for complex movements, most are less mobile and not wearable. One further disadvantage of the existing systems is their high complexity, for instance they demand high-speed cameras, multi-channel audio systems or the fixation to a special room or laboratory. Some combinations of our system and the before mentioned ones promise interesting synergetic effects, but these are not considered in this paper.

The presented wearable device is simple and robust and very cheap compared to visual screens, projectors, multi-channel audio systems on the output side or video cameras and microphones on the input side. Furthermore it is easy to use and install. The devices can be cascaded to a complex system e.g. attached to different body parts or more than one person. In combination with data processing methods, using external computers or directly implemented in the onboard chip, the wearable multi-sensor device provides new possibilities for research in motion capturing, human communication and manual learning. No complicated external cameras or CAVEs are needed. The lightweight and wearable-ability of our system allows unhindered movements in 3D space and enables applications in many fields, such as sports, arts and multi media to name a few. Similar technology has been first demonstrated in haptic augmentations and sonifications for applications in music and dance, (c) allows to relate feedback information in the tactile and acoustic medium, so that these important feedback possibilities are extended and trained supportive. Even more, external instructions from the teacher, trainer and the computer or other users can be transmitted directly and unobtrusively in these audio-haptic feedback channels.

Our sensor setup is also designed for motion and gestures in general. Several approaches of interactive gestures exist. Early approaches by Herrmann [8], Verfaille et. al. [9] show first ideas and setups. In our application section we describe the use in dance, sports and ideas about data exploration. In the case of dancers, which are used to coordinate to music, sound and rhythm, sonification can depict complex dependencies between action and reaction. Accordingly these dependencies can be understood easier through listening. Examples of dance training and learning scenarios for teacher to student or self assessment and analysis for dance motor skill learning are shown.
2. DESCRIPTION OF THE WEARABLE DEVICE

Our wearable sensor setup is similar to the work described in [7] but here the system is not tool-integrated and consists not only of a 5 degrees-of-freedom (DOF) sensor, meaning 3 axis acceleration sensors and 2 axis gyroscopes. But also of different goniometers (see fig. 3 and 4) and shoe integrated foot switches (see fig. 2).

2.1. Sensor Setup

Our easily relocatable flexible sensor based system is divided into the following 3 parts: Input (Sensors), Atmel IC and output (loudspeakers and vibrotactiles). This wearable setup allows simple usage, with or without the need of an external computer. In this contribution different employed sensor types will be presented, to show the possibilities and usage in several application scenarios.

The sensors are:

- 6 DOF sensor, meaning 3 axis acceleration and 3 axis gyroscopes sensors.
- Foot switches integrated into shoes.
- Semi-flexible goniometer.

The whole setup can be fixed simple and situational on the body of any person or dancer and adapted to special training situations and problem statements. Due to the higher scanning frequencies of the sensors compared to most visual sensing-based approaches, such as for instance fast movements like jumps or even pirouettes can be examined at high accuracy.

2.1.1. Foot Switches

Several approaches of sensors in shoes and soles exist, mostly for medical observations and gait analysis like in [10] [11] and [12]. In this paper, two insole soft pad switches in each shoe are used for simple foot position and movement detection. Especially for jumps it is important to know, when the feet leave the floor.

![Insole soft pad switches](image)

In our setup, two foot switches (see fig. 2) are integrated into each shoe. This allows contact detection and weight distribution of the feet and e.g. investigations on how and when “losing the contact” to the floor during jumps. The dimension of the basic sensors we use are now between 2x2x0.5cm and 5x5x0.5cm, embedded into foam plastic.

2.1.2. Goniometers

A goniometer is an instrument which measures an axis and range of motion, or the angle or rotation of an object precisely about the attached axis between two connected arms or small sticks.

Our self made goniometers are equipped with a potentiometer and used for joint angle measurement. This is a very precise and cheap sensor, easy to fix and install. The goniometers provide a high repeat accuracy, which allows usage for longer periods. Repeat accuracy here means, that the goniometers give the same start and end value before and after a dance figure or jump, without any “drift”. They can be mounted directly on the body or into the clothing, depending on how precise the measurement has to be. In our case, the goniometers are fixed to the body and we used them, to investigate the spatiotemporal correlation between different body parts, e.g. foot and knee (see sec. 4.1.3).

2.1.3. Accelerometer and Gyrosopes

Two IDG-300 dual-axis angular rate gyroscopes from InvenSense are used. This allows the measurement of the rotation of the x-, y- and z-axis. Further the ADXL330 acceleration sensor from
InvenSense is used, a small, thin, low power, complete x-, y-, and z-axis accelerometer.

2.1.4. Hardware, Data Transfer and Battery

The basic setup is realized with an Atmel Atmega328 microcontroller with 14 Digital I/O Pins (of which 6 provide PWM output) and 8 analog Input Pins. The dimension is 0.73” x 1.70”, (1.8 x 2,5cm) allows a small form factor and makes wearability easy.

Each sensor-IC node (see fig. 5) works self-sustaining, but additional Bluetooth data transmission is possible and external periphery like computers or more complex soundsystems can be used.

A small Lithium Polymer (LiPo) battery is directly attached for power supply. The H-Bridge is an integrated electronic circuit, which is in our case used to apply a voltage to the vibration motors and changes the speed.

3. MULTIMODAL OUTPUT AND CLOSED LOOP FEEDBACK

3.1. Sonification and Sound Synthesis

Different sound synthesis models in the area of music technology exist to generate sound and music. Beside the analog sound synthesis, various digital synthesis methods exist. The most common ones are subtractive, additive and frequency modulation synthesis. Further synthesis methods are granular-, wavetable-, phase distortion, sample-based and physical modeling synthesis, just to name a few.

In this paper, the embedded synthesizer (see scheme fig. 6) is using granular synthesis similar to [13], which works on the microsound time scale. Granular synthesis is often used sample based and in analog technology. Samples are split in small pieces of around 1 to 50 ms in length. The wearable embedded synthesizer uses oscillators instead of samples and multiple grains of these are layered on top of each other all playing at various speed, phase, volume, and pitch. Most parameters can be unfluenced with sensor input, so the scope of design is manifold.

The result is no single tone, but a complex sound, that is subject to manipulation with our sensors and switches and the produced sounds are unlike most other synthesis techniques. By varying the waveform, envelope, duration, spatial position, and density of the grains many different sounds can be produced.

3.2. The Two Basic Sonification Modes

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

1. Continuous Sonification: This method allows the continuous control of a movement or parts of it in real-time. The “shaping of a figure” is translated directly into a sound feedback. Especially the filter-like sound composition sounds appealing and sounds similar to popular musical effects are used to listen to anyway.

2. Case-Triggered Sonification: This means, the sound only appears, if a certain problem or deviation appears. The
sonification can be changed and turned on and off manually, so the dancers have permanent control. This allows the individual assignment of a specific sound or sound effect to each sensor or condition, or to group useful sensor combinations.

3.3. Wearable Embedded Sonification

Our integrated and wearable devices have at least one built-in loudspeaker. If acoustical feedback occurs, the position in the 3D-space is automatically given through the sound emitted by the device. In result, no complex pointers or 2D- or 3D-sound systems are necessary to point to the relevant position. The spatial hearing of the humans allows exact and fast location of the sound source, without having to turn the head or to change any corporal position. Figuratively, every device is an active moving sound source, meeting the human habit of hearing and reacting to noises and sounds in everyday life and environments. The directional characteristic of the built in loudspeakers allows even the acoustical recognition of the gyration of the wrist, which would hardly be possible to simulate in virtual sound environments.

3.3.1. Pulse-Width Modulation, digital to analog conversion and amplification

For audio out, the Pulse-Width Modulation (PWM) outs are used. Pulse-width modulation uses a rectangular pulse wave whose pulse width is modulated resulting in the variation of the average value of the waveform. A standard digital to analog converter circuit from [14] is used to receive the analog voltage (see fig. 7). This voltage is amplified with a transistor to drive the loudspeaker and is easy available for around 1 euro. Suitable vibration frequencies are around 250 Hz, since fingers and skin are most sensitive to these frequencies (see [15]). In this paper we present a new developed active vibrotactile feedback system, easy to use, lightweight and very flexible attachable to manifold objects and body parts. In this case two vibration motors are fixed to our board.

![Figure 7: Digital analog converter with amplifier](image)

3.4. Vibro-Tactile Feedback

3.4.1. The Vibration Motor

Several vibration devices were taken into account, including simple vibration motors, solenoide piezo-electric elements and voice coils. Besides the simple control, weight and form factor, the availability and price have been important criteria for the choice. The left vibration motor in fig. 8 with the dimensions 5x15mm, lightweight and cylindric shape seemed to be the best compromise. Furthermore, this kind of motor is typically used in mobile phones and is easy available for around 1 euro. Suitable vibration frequencies are around 250 Hz, since fingers and skin are most sensitive to these frequencies (see [15]). In this paper we present a new developed active vibrotactile feedback system, easy to use, lightweight and very flexible attachable to manifold objects and body parts. In this case two vibration motors are fixed to our board.

![Figure 8: Several vibration motors](image)

3.4.2. Listening with the Skin

We call “listening with the Skin” the awareness of local distributed and dynamically triggered vibro-tactile feedback. The vibrations are short rhythmic bursts between 40Hz and 800Hz, which is the sensitive range of the mechanoreceptors in the fingers. The distance between the two motors is big enough for easy identification which one is vibrating. The amplitude and frequency can be varied independently. This allows to evoke more or less attention, increasing and decreasing of the vibration and at least 4 significant combinations between the two motors: (1) both motors on, (2) motor 1 on, motor 2 off, (3) motor 2 on and motor 1 off and (4) both motors off. As described in Bird [16] the touch-sense feedback channel is extended and the awareness of the vibrotactile feedback is increased and trained.

3.5. Multi Channel versus Direct Sound

Compared to existing standard audio setups, especially multi channel systems, our wearable device is very simple, but very easy locateable in the 3D listening space. A simple example is, if you try to locate an alarm clock just by hearing the alarm, you know very simple and exact, where it is and the sound comes from. On the other hand, finding the exact position of a sound source in a stereo or multi channel sound field, is much more difficult and dependent of the position of the listener. If there are more than one persons, trying to describe the same source, it is already nearly impossible. If you perform this tasks with headphones, it is easier, but usually headphones are not applicable in many situations.

More advanced technologies like 3D Audio, Spatial Audio, and WFS systems improve partly the stability of the sound source, but again, the complexity and form factor of the equipment does not fit into the idea of a new, unobtrousive wearable interface.

The developed device can not only be fitted with more loudspeakers for multi channel audio out, even more than one wearable device itself can be fixed on the body or clothes. In this case, more
different sounds from more directions can be provided and produce interesting interferences.

3.6. The “Audio Loupe”

A conventional loupe or magnifier glass is a [17] “type of magnification device used to see small details more closely”. In this paper we introduce the “audio loupe”, a acoustical magnifying glass for motion, posture and gesture in an auditory form. This means: Acoustical time stretching of fast motion, like an echo, and, similar to visual zooms, a acoustic zoom in- and out function to magnify or demagnify positions or movements. If for example a constant movement varies in speed, the deviation is not perceived physically, but measured and sonified. similar to an “loupe” are developed. This approach does not exclude visualisation at all, but combinations of several feedback channels in future projects might provide additional help for understanding and learning. The magnifying works precisely, especially with the goniometers. Smallest deviations of a continuous movement are detected or smallest movements in stagnant postures. This is important in dance and coordination training as described in sec. 4.1.

4. APPLICATIONS

One basic idea of this new interactive interface was, to receive a 3D audio-haptic feedback in the most easy but realistic, precise and useful way. In the end, the user and performer should be able to set up the device alone, without the support of a technician. This will help to increase the acceptance of this new technologies and methods. In the following, two applications are described for dancing and the data exploration. Further interesting fields would be learning and improvising music, similar or additional to the applications and systems from Beilharz [18] and Bevilacqua [19].

4.1. Dance

As dancers are used to coordinate to music, sound and rhythm, sonification in this case can depict complex dependencies between action and reaction. Accordingly these dependencies can be understood easier through listening. Examples of dance training and learning scenarios for teacher to student or self assessment and analysis for dance motor skill learning are developed. The most reliable data in fast motion scenarios and jumps are the goniometer data. The calculated data of the accelerometers and the gyroscopes still have a certain drift and an infeasible repeat accuracy. The professional system of XSense [20] is expensive, too large housings, and not flexible enough, especially if additional sensors are needed. The 6-DOF Board is used for tilt detection and acceleration measurements of jumps.

4.1.1. Constancy of Motion

The specific task described in the following section was the constant speed of motion in specific planes and lines and stretching.

The data in fig. 9 show two slow motion arm movements, the first one a correct movement and the second one an incorrect. The upper line is the goniometer data, the two lines below are the acceleration data. Here we have the problem, that it is hard to see the difference between the good and bad example, but the sensor data mapped to sound in realtime creates useful feedback. The sensor data of differing speeds in the same motions are sonified with the following possibilities:

- Sonification position of the upper arm.
- Continuous sonification of the angular rate changes of the elbow, or the deviation of the ideal speed.
- Continuous sonification of the angular rate changes of the wrist, or again the deviation.

In other examples, the device provides feedback, if a certain point or e.g. height of the hand or foot is reached. This means a simple way of controlling the quality of the exercise or right amount of stretching is reached. Here Continuous Sonification is used (ref. sec. 3.3).

4.1.2. Group Dancing and Synchronisation

In group dancing situations synchronised motion is an important issue and difficult issue to train. It is simplified by measuring the speed of the angular rate changes. Differences in speed are displayed in acoustic or tactile form. Here Case-Triggered Sonification is used (ref. sec. 3.3).

Figure 9: Sensor data of constant correct and incorrect arm movement

- Sonification position of the upper arm.
- Continuous sonification of the angular rate changes of the elbow, or the deviation of the ideal speed.
- Continuous sonification of the angular rate changes of the wrist, or again the deviation.

Figure 10: Sensor data of two synchronised jumps

The fig. 10 shows a recordings of coordinated movements. Here again, sonification of the turning points indicates the synchronisation with one or more dancers. This data are recorded with two dancers and data transmission via Bluetooth to a standard laptop computer. This allows later examination of the data, but in real life training situations, real-time audio or tactile feedback increases training efficiency.

Several aspects of synchronisation and alikeness are clearly displayed by sonification:

- synchronised starting points at the beginning of the jumps
- timing of the landing
4.1.3. Jumps

Another exemplary scenario is a jump in different variations. The jumps are not only of good or bad quality, they include small faults and deviations causing different results. In dance movements and jumps it is usually not only a question of “correct” or “incorrect”, it is more a “thinking” about complex dependencies between many parameters to find a final coherent result (see fig. 12).

Figure 11: temporal, local and rhythmic dependencies

This means, certain values are sorted e.g. knee angles or time between lift off and point of return in the air. These previous points are sonified during the next, or the further following jumps. This enables the dancer to compare different trials from the past with your current jump in real-time. In both mentioned examples, only the maximum values are sonified, to hear the difference between the trials exactly. This acoustical repitition of the past in this case jumps, allows efficient online support and investigation possibilites during the active training phase.

The basic setup consists of 2 Goniometers, 2 foot switches, the accelerometer and gyrometer board. In fig. 12 only two of the 10 data channels, meaning two sensors of the left leg sensors are plotted. The upper line is the vertical acceleration, the lower line is the angle of the knee. It is quite hard to see the differences of the correct example (first jump) and the incorrect eaxmple (second one) and only post exploration of the data is possible. With realtime, onboard embedded sonification, the differences are more easy to investigate during the training. This is supported by recommender systems like in sec. 3.7.

Fig. 12 shows the sensor data of differing speeds and correct (the first jump) and incorrect (the second jump) body springiness distribution in the same motions, with the following sonification possibilities:

- Sonification of the knee bending.
- Audio cue, when the food leaves the floor.

![Figure 12: 2 jumps, good and bad version, acceleration versus knee angle](image)

- Over all acceleration sonification.

This measuring method also allows dancers extensive “offline” analysis of their movements, if the sensor data are saved. Sonified variations of body parts and joints, different trials with several changes of certain parameters, positive and negative progress and dependencies between all of them are shown and sonified. Audio feedback of different trials, again similar to a loupe, for professional dancers in an auditory form will provide more possibilities in the future, the longer this setup is evaluated.

Different useful combinations of important parameters in 3D-space and temporal flow are sonified. Also positive or negative skill developments between different trials and rhythmic and temporal synchronization of motion sequences are explored and sonified. Some combinations are:

- Combination of single motion points and sequences (of different trials).
- Combination of different trials with important positive or negative changes.
- Combination of different trials with important changes within longer and shorter sequences.

A further idea is the reduction of a complex movement sequence to small steps and working out of a personal “best case” scenario, to be achieved later again and more and more often after a certain amount of trials.

4.2. Sports/Every-Day Postures and Gestures

Walking with insole e.g. realised by Benocci et al. [21] and Kong [10] with several pressure sensors in the sole of the shoe. Our system is simplier, as we don’t need the pressure data for our investigations. Walking, Running, “Rhythmische Gleichmäßigkeit und/oder Abweichung davon. Recommender System etc.

- Rhythmal regularity of the single steps
- Regularity of the step size
- Measurement of the constancy of the upright acceleration

Fig. 13 shows data of a leg movement while running with small rhythmical deviations from a steady running flow. The “acoustic augmentation” of the running shows the rhythmical regularity and, even more important, the flow of the motion. For examples a typcall symptom of fatigue is irregularity of the leg.
motion, which can be sonified with continuous sonification or case-triggered sonification, if it reacts to abrupt angular rate changes.

Figure 13: Footsteps while running, with knee angle versus walking speed measurement

5. CONCLUSION

Sonification and haptic feedback addresses, besides the visual sense, a wide range of feedback channels available. For that end we presented a multimodal audiohaptic and wearable sensor/actuator system to support human activity for many possible applications. The described way of the integration of many sensors and output possibilities are expected to have a positive effect in many learning scenarios and multi-sensorial perception. The audio-haptic feedback possibilities demonstrate that changes in movement - here in 3D-space - can be signaled unobtrusively and quite intuitively using combined haptics and audio as indexical and information carrying sign. Even real-time correction or an overdone correction can be shown.

Our first impression is that the continuous sensor data based closed-loop audio-haptic feedback described above works well and is quite efficient to direct the attention to improper executions. As promising prospect, the system may for example lead to learning aids for visually impaired people, especially as they are more biased to use their non-visual senses to compensate the lacking visual information.

The feedback helps to understand quite intuitively, how a special and complex movement is executed and trained. Further developments in augmenting both areas, the sensor and the feedback side, will show how learning processes can be improved and adapted to situated demands in everyday life situations. Especially the wrist-mounted device with the multi-modal feedback and multi sensory input is adaptable to different scenarios such as in sports, music, dance, games and many more interaction scenarios. Also interactive music systems for improvisation are considered with this setup.

The ”Audio loupe” is a promising method in the field of high level dance and motor skill learning especially for examination and monitoring of progress and hard to understand complex movements and dependencies.

6. ACKNOWLEDGEMENT

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

7. REFERENCES


