FP7-ICT-2007-C
Objective ICT-2007.8.0
FET-Open 222107

NIW
Natural Interactive Walking

Deliverable 3.1

Contact-based Sensing Methods for Walking Interactions

Yon Visell, Alvin Law, Jessica Ip, Rishi Rajalingham, Severin Smith, Jeremy R. Cooperstock - McGill Univ.

Gianpaolo Borin, Marco Civolani, Antonio De Sena, Carlo Drioli, Federico Fontana, Pietro Polotti – UNIVR

Rolf Nordahl, Stefania Serafin, Luca Turchet, Smilen Dimitrov – AAU

Sep. 30, 2009
v. 1.0

Classification: PU

Contents

1 Contact-Based Force Sensing (McGill)

Work described within this section of the deliverable details research on the application of intrinsic force sensing methods for surface contact acquisition via a distributed array of low-cost, rigid tiles. The results are presented in detail in a report that is included in the supplementary materials for this deliverable. (The report itself is currently undergoing blind peer review for a conference proceedings, so should not be treated as public material in its present form.)

The paper presents a novel, low-cost system and set of techniques designed to enable users to interact naturally on foot in immersive virtual environments. The physical interface consists of an array of networked, rigid floor tiles distributed over an area of several square meters within a CAVE-like virtual environment simulator. Each tile within the array is capable of 3 degree-of-freedom isometric sensing of forces applied by users' feet. The mechanical constraints of this interface are exploited by our data processing algorithms in such a way that it is possible to capture foot-floor interactions with a linear resolution as low as 1-2 cm, up to 30 times better than the nominal resolution of the tile array, over an area of several square meters. The resulting position and force estimates are employed to track contact positions and forces, and thereby to allow users to interact with distributed, deformable ground surfaces, such as soil or ice, presented in a virtual environment, or to operate virtual interface elements such as navigational widgets in a hands-free way. During interaction, users receive visual, auditory, and vibrotactile feedback in the form of physically plausible interactions between the foot and the ground surface or interface. We demonstrate the application of this device to the simulation of deformable ground surfaces in an immersive virtual environment, and to the presentation of virtual floor-based interface elements (see also Deliverable 4.1).

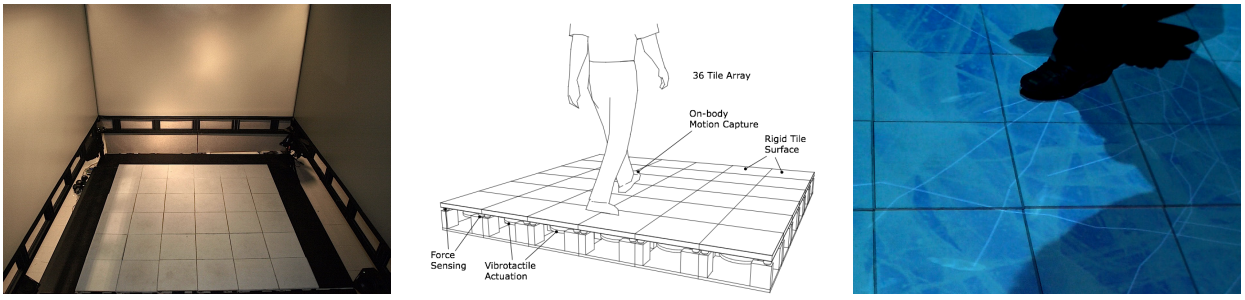


Figure 1.1: Left: Distributed floor interface prototype as installed in the authors' laboratory. The photo shows the floor interface as situated within an immersive, rear projected virtual environment simulator. Middle: Illustration of the interface. Right: Still image from video documentation of an interactive scenario in which the sensing methods presented in the report are employed to enable users to interact with a frozen pond demonstration. Cracks in the dynamically form at loci of contact between feet and floor.

Supplementary materials for this section

The report referenced above, on contact based force sensing via a distributed floor tile array, can be downloaded from the supplementary materials included with the deliverable, on the project website.

2 Acoustically-Sensitive Pavements (UNIVR)

In this chapter, we discuss the development of sensitive pavements by means of capture and analysis of the mechanical waves generated by the interaction of walkers' feet with the pavement itself. The appealing aspect of this approach is that it exploits the natural, pervasive and costless sensing-network embedded in any object because of its solid state and its elastic properties that allows the propagation of mechanical waves. In principle, known the physical properties of the "passive"object, i.e. the medium (the pavement, in our case), it should be possible to "listen" to the mechanical waves generated by the interaction with an "active" object (a shoe in our case) at some distant location in the medium and understand where the interaction took place. Something similar to what happens, when we localize a sound source by listening to it. In that case, we exploit the fact that we use a couple of sensors, the ears, distant a certain fix amount of centimeters one from the other. It is well-known that the two ears form the simplest case of a sensor array. Moreover, we are able to extract further spatial information by exploiting some other cues conveyed in the sound, as the spectral shape and others, this latter aspect being dependent from a foregoing learning process about the kind of sounds that we listen to.

Modelization of wave propagation in-air and localization of sound sources by means of wave capturing and analysis is a well established field. There is a huge amount of literature on the subject specifically about methods based on sensor arrays and data analysis (for an exhaustive overview, see, for example, [1]).

However, in-solid wave propagation is a much more tangled process as compared to the in-air case. In practice, many approaches to the problem of in-solid localization are possible and none is optimal, due to the complexity of the involved physics. According to the physical conditions and the specific task, one or the other technique can better fit the requirements of the application that one needs to develop. This subject was widely investigated by a European project funded under the FP6 within the action IST-2002- Multimodal Interfaces: TAI-CHI (Tangible Acoustic Interface for Computer Human Interaction [10]). The main objective of the project was the development of tangible human-computer interfaces based on acoustical sensing, in a similar fashion as we are aiming at here in the case of walking surfaces.

In the next sections, we provide a concise and qualitative overview of the field of in-solid acoustics and describe one approach to the localization problem that UNIVR implemented and tested: the Time Reversal (TR) method. This methodology was picked from the results of TAI-CHI project.

In-solid acoustics: an overview

Acoustic propagation in solid media is well-known to be a very complex physical phenomenon to study, characterize, and model. A wide variety of physical phenomena that are difficult to predict and account for always come in the way of a correct modelling of the wave propagation phenomenon, and interfere with the constraints of accuracy and flexibility in the development of acoustically-based localization systems. Such phenomena are partly due to the inherent physical properties of in-solid propagation phenomena (wide variety of modes of propagation, wavefront dispersion phenomena, means of propagation that are inhomogeneous and non isotropic, etc.), and partly related to the acquisition process (sensor characteristics, contact film/glue response, digitalization process, etc.). The TAI-CHI project has provided theoretical solutions and innovative approaches to overcoming these limitations.

Even if one considers an ideal model of an unbounded medium, one has to deal with more than

a single class of waves. Both Longitudinal and Transversal waves are present in an unbounded solid medium, i.e., in first approximation, far from its boundaries. For this reason, they are called Bulk waves (see Figure 2.1)

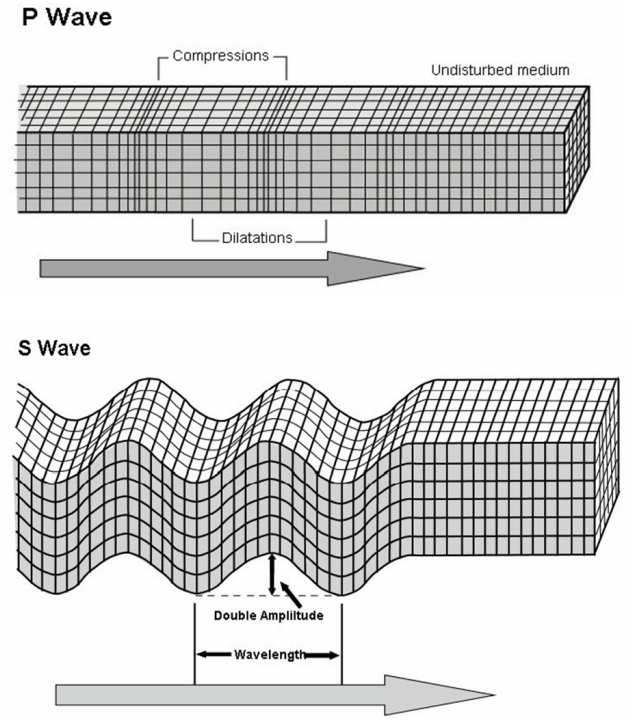


Figure 2.1: Bulk waves in unbounded solid media.

More precisely, in the Longitudinal waves (also called Pressure waves) the oscillations of the particles occur in the same direction or opposite to the direction of energy transfer. On the other side, Transverse waves (also called Shear waves) consist of oscillations occurring perpendicularly to the direction of energy transfer (see Figure 2.1).

Besides, other forms of vibrational phenomena occur in proximity of the boundaries. Surface waves (also called Rayleigh waves) appear and propagate on the surface of the solid. They are a combination of transversal and longitudinal waves that causes the particles to vibrate according to elliptic orbits. Also, another kind of surface waves can be detected, i.e. Loves waves, which result in vibrations on the horizontal plane (see Figure 2.2).

The situation gets even more complicated in the very common case of wave propagation in a medium bounded by two planar surfaces (a plate, or, in our case, a pavement). In that case, Lamb's theory applies [11]. Lamb's theoretical formulations have found substantial practical application, especially in the field of nondestructive testing [5]. Lamb's theory indicates the existence of two entire families of sinusoidal wave modes in infinite plates of a certain width. This stands in contrast with what happens in unbounded media, where there are only two wave modes, the longitudinal waves and the transverse waves.

Lamb waves are similar to longitudinal waves, with compression and rarefaction, but they are bounded by the sheet or plate surface causing a wave-guide effect. As in Rayleigh waves, the particle motion in Lamb waves is elliptical (see Figure 2.3). There are two distinct types of plate waves: extensional (symmetric, s) and flexural (antisymmetric, a), each of which have an infinite number of modes (s0, s1, s2, ... , and a0, a1, a2, ...) at higher frequencies. The symmetrical and antisymmetric zero-order modes, however, are special, since they are the only modes that exist over the entire frequency spectrum from zero to indefinitely high frequencies. As the frequency is raised, the higher-order wave modes make their appearance in addition to the zero-order modes. Each higher-order mode appears at a resonant frequency of the plate, and exists only above that frequency.

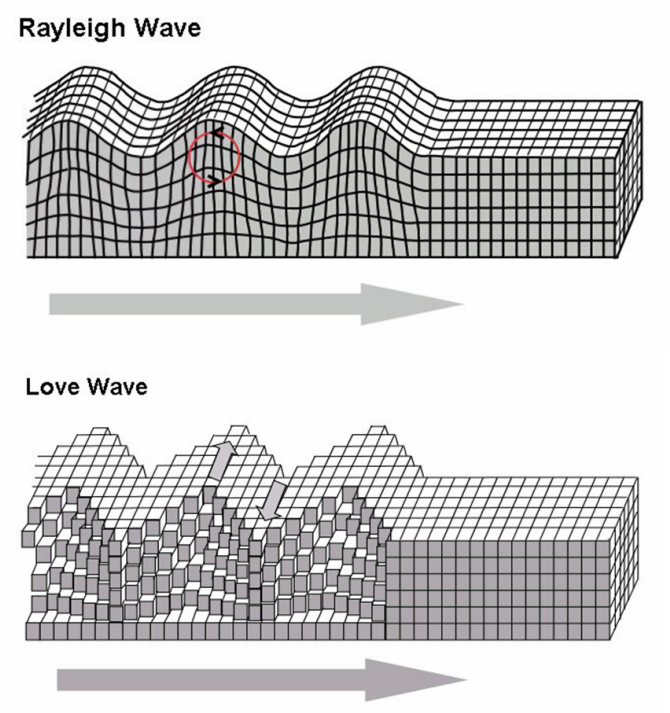


Figure 2.2: Surface waves in solid media.

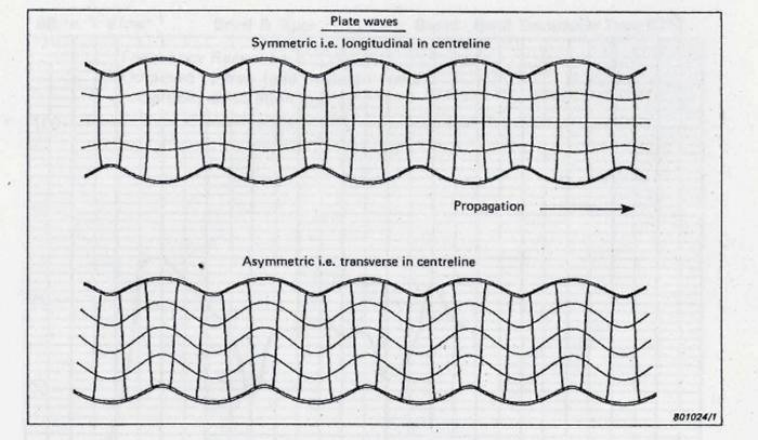


Figure 2.3: Lamb waves in solid plates: the two fundamental modes.

Furthermore, Lamb waves exhibit strong phase velocity dispersion effects that is the propagation velocity depends on the frequency (or wavelength), as well as on the elastic constants and density of the material (see Figure 2.4). This phenomenon is central to the study and understanding of wave behavior in plates. Dispersion means that the waveform is not preserved while the wave is propagating.

Moreover, the velocity of the Bulk components is higher than that of the Surface components. On the other side, the energy of the latter is much bigger, and the relevant signal that arrives at a sensor is composed by surface waves. This means that, the relevant signal is anticipated by a low energy signal due to the bulk waves, which constitute a disturbance element for the measurement process.

The totality of these phenomena limits the possibility of a reductionist approach to the localization problem. Whereas geometrical methods such as TDOA are relatively robust in-air, their application to the in-solid case can be critical, especially in case of non-homogeneous, non-isotropous and highly dispersive materials.

On the other side, the non-linearity of the propagation can be overturned into a point of force. In fact, given a certain excitation in a certain point of the medium, the continuously changing waveforms become an univocal characterization of the event at any other point of the medium, where the propagating wave is recorded. By means of some pattern recognition technique is, thus, possible to identify the excitation point on the basis of the detection and classification of the waveform arriving at a certain sensor. The latter approach, its advantages and drawbacks will be discussed in the following sections.

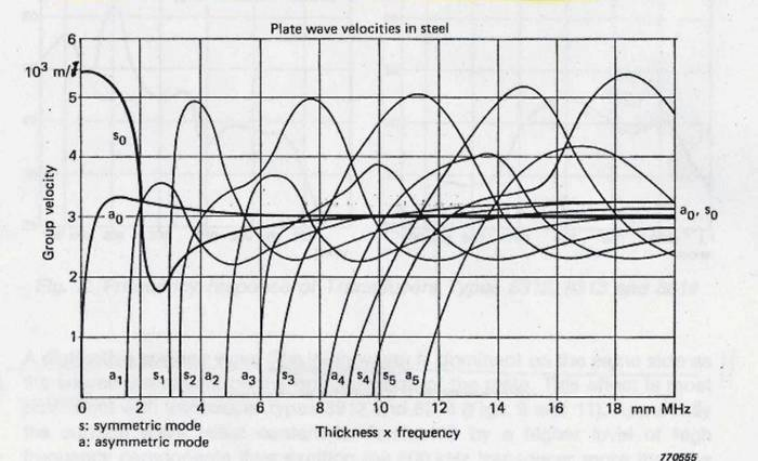


Figure 2.4: Dispersion phenomenon in Lamb wave modes.

Methods for acoustically-based in-solid source localization

As acoustic vibration propagates well in most materials, the information about the interaction is conveyed to a remote location, using the structure of the object itself as a transmission channel and suppressing the need for any overlay or any other intrusive device over the area one wishes to make sensitive. Moreover, only inexpensive acoustic contact sensors and a few audio inputs to a computer are sufficient.

Any physical contact with a solid object or a surface (wall, floor, table, etc.) modifies its state by the way mechanical-acoustic energy is distributed in the object and on its surfaces. As reported by the TAI-CHI project, such perturbation can be caused in two ways. Firstly, by the acoustic vibrations generated at the points of contact, for example, when tapping or moving a foot on the surface of a pavement (passive methods). Secondly, by the acoustic energy that is absorbed at the points of contact (proportional to contact pressure), when the object is activated with ultrasound (holography-based active methods). In [2], an overview of the methodologies studied in the TAI-CHI project is available. In this project, we only faced the study and the implementation of cases of passive methods such as:

- Black-box approach: when modelling the acoustic propagation is not possible, a black-box fashion is selected such as the TR method. This approach was implemented by UNIVR and is described in the following section.
- Model based approach: whereas the wave propagation in thin plates can be accurately modelled, then a model based approach can be employed such as TDOA. This approach was preliminary tested by AAU and will be described in future reports.

TR method: Footprints for step localization

TR in acoustics is a very straightforward and efficient solution for determining where a sound comes from, i.e. for localizing its source in a wide range of material including reverberating media. It is based on the principle that the impulse response in a chaotic cavity is unique for a given source location. In a sense, a wave propagating in the cavity contains the memory of its source location.

In the ninties, Draeger and Fink [3] experimented the time reversal invariance of acoustic waves in chaotic cavities. In the experiment reported in their paper, an ultrasonic pulse was sent from a transducer (the source) and propagated inside a silicon plate. Due to reflections and reverberation, the acoustic field, measured on a second transducer (the receiver) lasted hundreds of times the initial pulse duration. This long and complex field was then time reversed and reemitted from the receiver. The acoustic field propagated back to the source, recreating the short initial pulse.

One of the methods that were developed by the TAI-CHI project and that we will employ here is a particular case of TR technique [8]. It consists in detecting the in-solid acoustic waves generated by human steps on a walking surface. The detection is a two phase process. The first one is the acquisition of the impulse response for a number of locations that cover an area of the pavement accordingly to the requirements of the target application. More precisely, for each position, a short pulse is emitted by tapping on the surface of the object, which propagates through the solid cavity and reflects inside. The duration of the response depends on the absorption of the material and on the energy radiation property of the pavement "plate". Each recorded tap provides a sort of "footprint" of the corresponding location as "perceived" at a specific sensor. When all of the footprints are collected, the learning procedure is finished. In the second phase, as an event occurs, the information related to the source location is extracted from a virtual time reversal process run in the computer. In other words a cross-correlation test is performed. Obviously, when dealing with real steps on a pavement and a particular type of shoes, one has to record the step signal nearby one of the sensor and use that signal to perform a convolution with the characteristic impulse responses, in order to define the shoe-specific set of footprints.

The advantages of the TR methods are:

- It works on non-flat objects and 3-D objects.
- It has no special requirements for materials, shapes and homogeneity of the interactive object.
- It can even work with a single sensor.
- It can use the normal audio input channel of a PC.
- It is a low cost method.
- It is selective: in fact, it is possible to avoid false triggering due to ambient noise, or, similarly, to operate selective tracking: for example, a surface might be sensitive to the contact of a heal but not to a rubber sole.
- In principle, it will be possible to determine not only where but also how a step is done, providing a complete description of the interaction.

The drawbacks of the TR methods are:

- It requires long calibration time for a large number of impact points.
- The reliability can be affected by the change of the ambient temperature.
- Reliability of localization could be affected by the type of shoe.

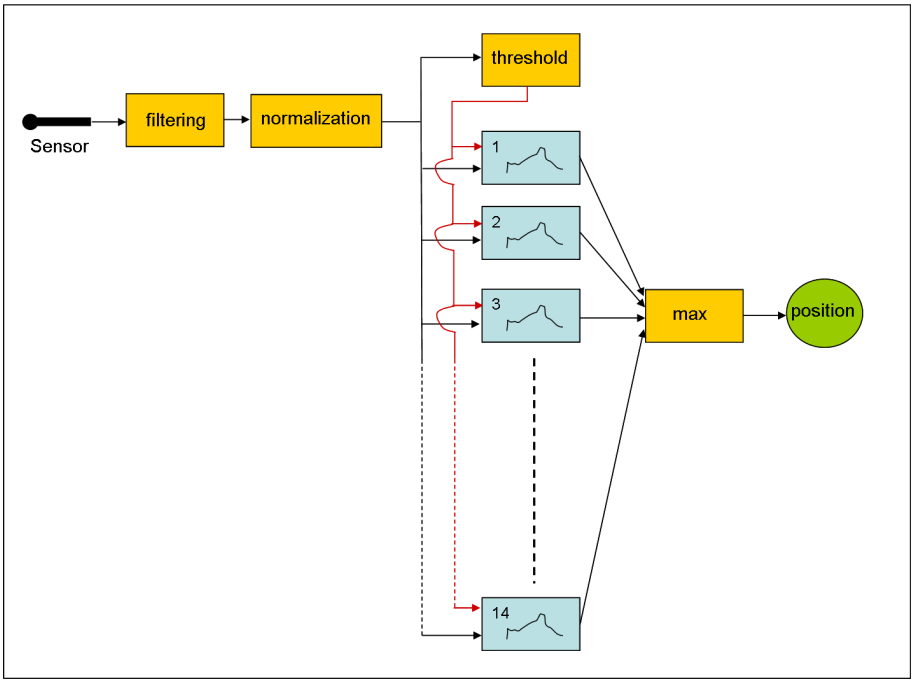


Figure 2.5: Scheme of the TR method for in-solid localization. Each one of the fourteen boxes represent a cross-correlation test.

Experimental outcomes

In this section, we describe the acquisition hardware that were implemented and assembled by AAU and UNIVR and the experimental results that were obtained by UNIVR for the case of the TR method.

Hardware setup

The base element of the acquisition hardware is given by the devices by Knowles Acoustics. These sensors are referred to as 1-axis accelerometers or vibration transducers by the company, and claim a flat frequency response in the range from 100 Hz to 10 KHz [6]. In NIW, they are used as contact microphones. In AAU a so-called interface was mounted, consisting of a hard plywood board with dimensions of 200 × 367 cm, and thickness of 5 mm, equipped with four devices, mounted as shown on Figure 2.6. In UNIVR, the were applied directly to the pavement as shown in Figure 2.11

The devices were soldered using thin enamel wire to a 3-pin molex connector on a small piece of perfboard PCB; the 3-pin molex connector allows subsequent attachment to a microphone cable. Thin enameled wire was used, as the connector pads on these devices proved extremely fragile, and tended to fall off when plain threaded copper wire was used (without a subsequent chance for repair). Stereo microphone cables are used, such that one of the channels carried the supply voltage for the respective accelerometer. The microphone cables from all of the devices are collected in a single perfboard PCB - acting, in a sense, as a hub - which simply implements the “Alternate 3-Wire Hookup” wiring for each individual sensor (as recommended in [6]). This setup is shown on the diagram on Figure 2.7(a):

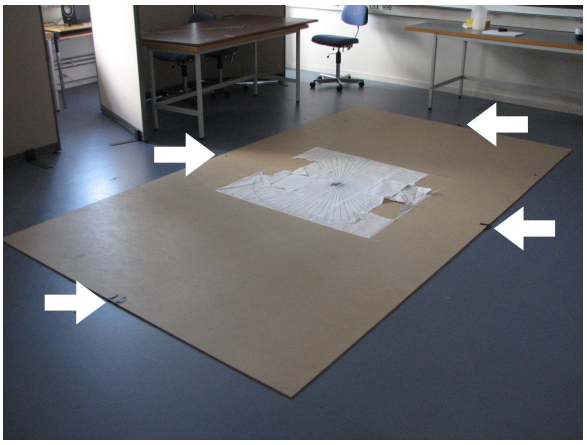


Figure 2.6: Image of the board used in AAU; arrows indicate the placement of sensors.

The devices were powered with 9V directly, as [6] notes 10V as upper limit for the power supply. The "hub" perfboard were short pieces of stereo audio cable, terminating with a TRS connector (where only a single channel carries the output from the sensor). The connectors were finally plugged in a multichannel soundcard, which was deemed a natural choice for a data acquisition hardware - as the sensor clearly produces output in the audible frequency range.

An adhesive product (known as a pressure-sensitive adhesive), branded "Bantex Tack-All", was used to affix the sensor devices to the board or the pavement. The choice of this material was imposed by the need to have a mobile setup, which would allow for easy mounting and unmounting of sensors to the board or to the pavement - while allowing for relatively firm contact between the sensor and the board/pavement while in operation. The perfboard pieces, which host the sensor, were deliberately made a little longer, so the board piece is the object in contact with the board through adhesive, rather than a sensor device. An early test of this concept is shown in Figure 2.8.

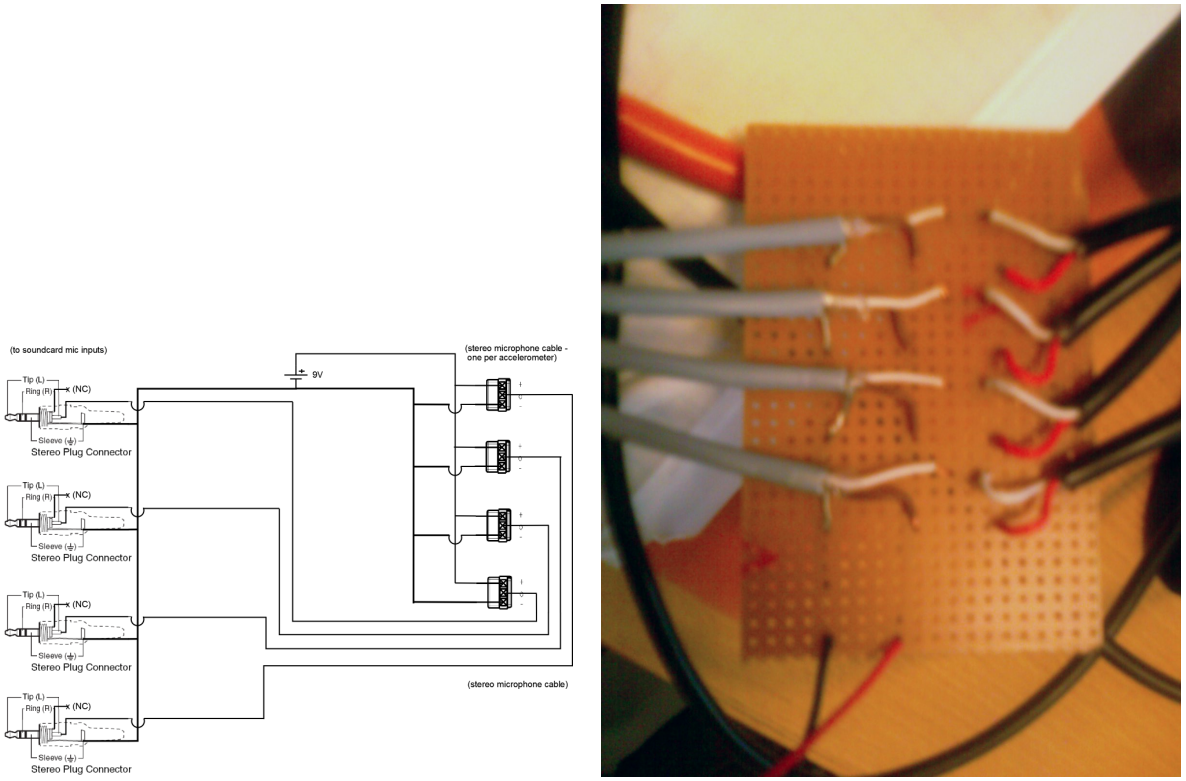
TR experimental results

In Verona, a basic scenario was chosen in order to test the footprint-based system. The scenario is that of defining a boundary, which divides a surface (the pavement) into two areas, one of which is forbidden. The idea is that, when approaching the boundary by walking, the system reacts by means of warning signals with a higher level of urgency as the walker gets closer to the boundary.

The experimental setup was configured in the way depicted in Figure 2.9. Fourteen positions-footprints define an area with four levels of proximity to an ideal borderline. The monitoring system is provided by an array of eight accelerometers of the kind described in the previous section.

The first phase of the experiment consisted in recording a footprint for each position. The signals collected by means of a single recording for each position are sufficient to define a reliable set of footprints of the area. A set of 8X14 footprints was so obtained, which underwent a process of filtering and normalization. Figure 2.12 shows the footprints for 2 positions as sensed at one of the microphones.

After the learning phase, whenever the walker crosses the area covered by the position matrix, as in the example depicted in Figure 2.13, the system operates a TR calculation over the whole set of footprints. For each sensor, the quantized estimated position is identified by the footprint, providing the absolute maximum among all of the cross correlation coefficients coming out of the TR method. Figure 2.14 represents the outcomes related to the third step of Figure 2.13 as detected by the first microphone of the array. The estimation is considerably improved by taking the median value over all of the results of the eight microphones as illustrated in Figure 2.15. Finally, Figure 2.16 shows the overall result for the third step, considering the outcomes along the whole lines, corresponding to the warning levels.



(a) Diagram of the wiring in AAU.

(b) Perfboard image.

Figure 2.7: Diagram of the wiring adopted in AAU (2.7(a)), and image of the perfboard (2.7(a)).

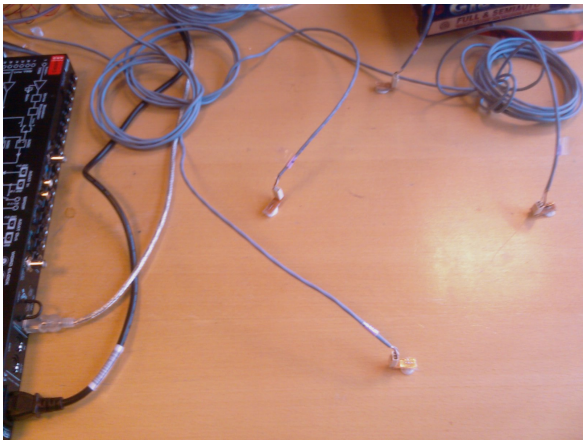


Figure 2.8: Test mounting of sensors used in the AAU board on a desk using tack adhesive.

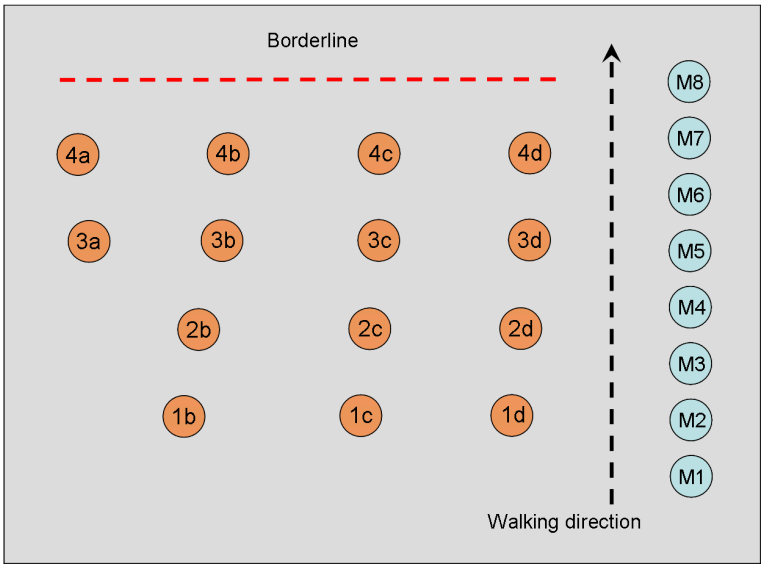


Figure 2.9: Measurement setup: 14 positions (footprints), monitored by an array of 8 microphones dislocated in parallel with respect to the motion direction. The numbers of each line of positions correspond to a progressively higher degree of warning.

Figure 2.10: The real position matrix.

Figure 2.11: The sensor array.

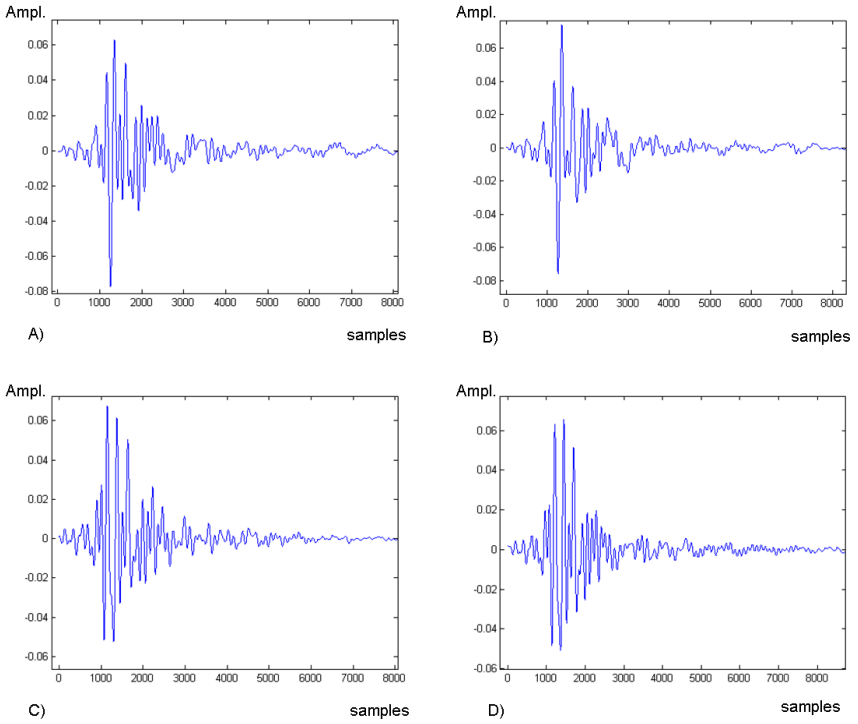


Figure 2.12: A) Step at position 2b in the learning walk as detected by microphone 4, B) Step at position 2b in a following walk as detected by microphone 4, C) Step at position 2d in the learning walk as detected by microphone 4, D) Step at position 2d in a following walk as detected by microphone 4.

Figure 2.13: Example of path through the sensitive pavement.

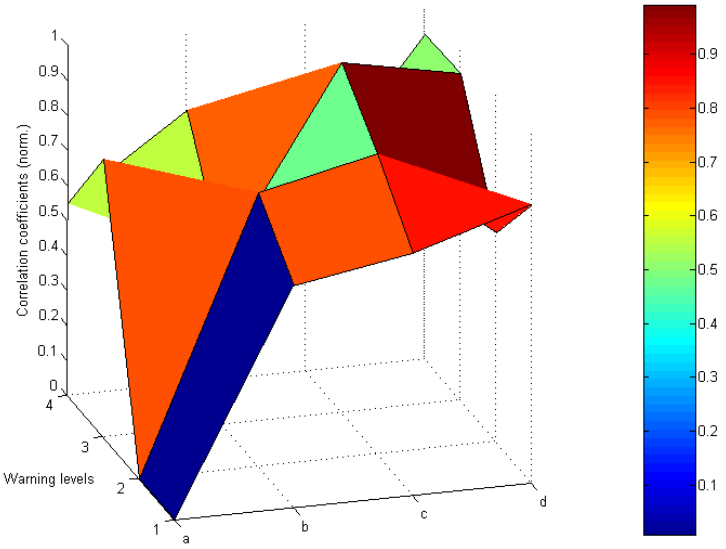


Figure 2.14: TR correlation coefficients for the third step of path of Figure 2.13 obtained by means of a single sensor.

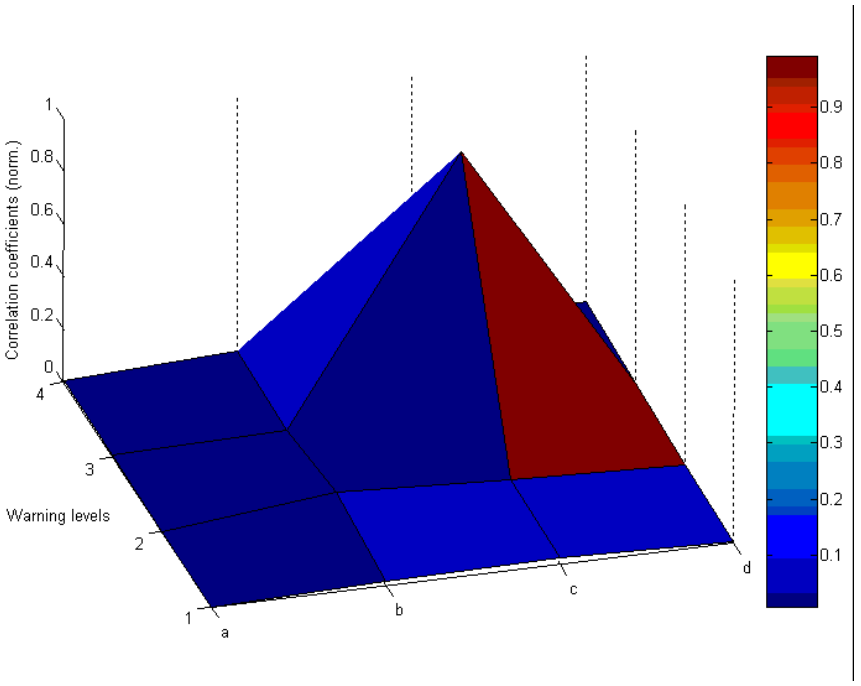


Figure 2.15: TR correlation coefficients for the third step of path of Figure 2.13 obtained by means of all of the eight sensors.

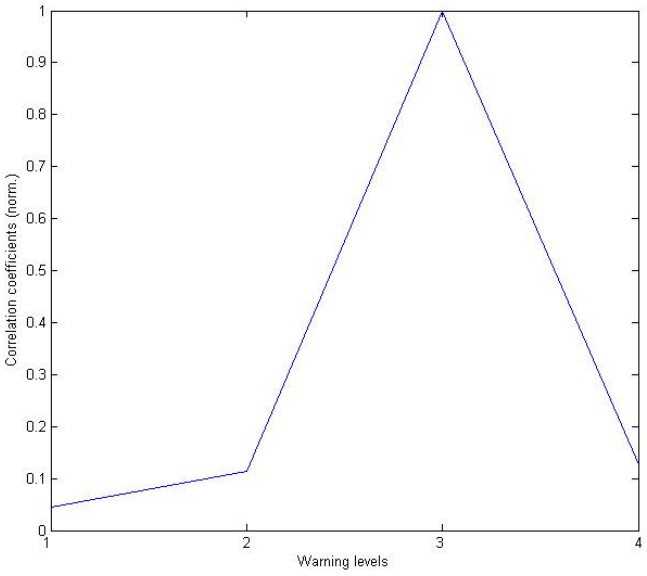


Figure 2.16: Same as in Figure , collapsed by warning level, i.e. along the four lines of the position matrix.

3 In-Shoe Contact Based Sensing (UNIVR)

3.1 Introduction

This part of the deliverable describes in detail the sensing interface embedded in the active shoes setup: main goal of that shoes-based system is to provide user with auditory cues of ground while walking. In order to generate ecological walking sounds, data supplied by sensors embedded in the shoes drive a set of physically-based synthesis algorithms running on a laptop that finds place inside a backpack. The laptop's audio output is sent to loudspeakers worn by the subject. The setup has been developed in form of two prototypes: although the ideas behind them are the same, they differ from a technological point of view. Making the early prototype in fact, had been very useful in order to identify the major problems that had to be solved: these were related in particular, to the reliability and wearability of the hardware, to the optimization of acquisition software and to the quality of acoustic display. Figure 3.1 shows a picture of the first prototype, while the second prototype is shown in Figure 3.2.

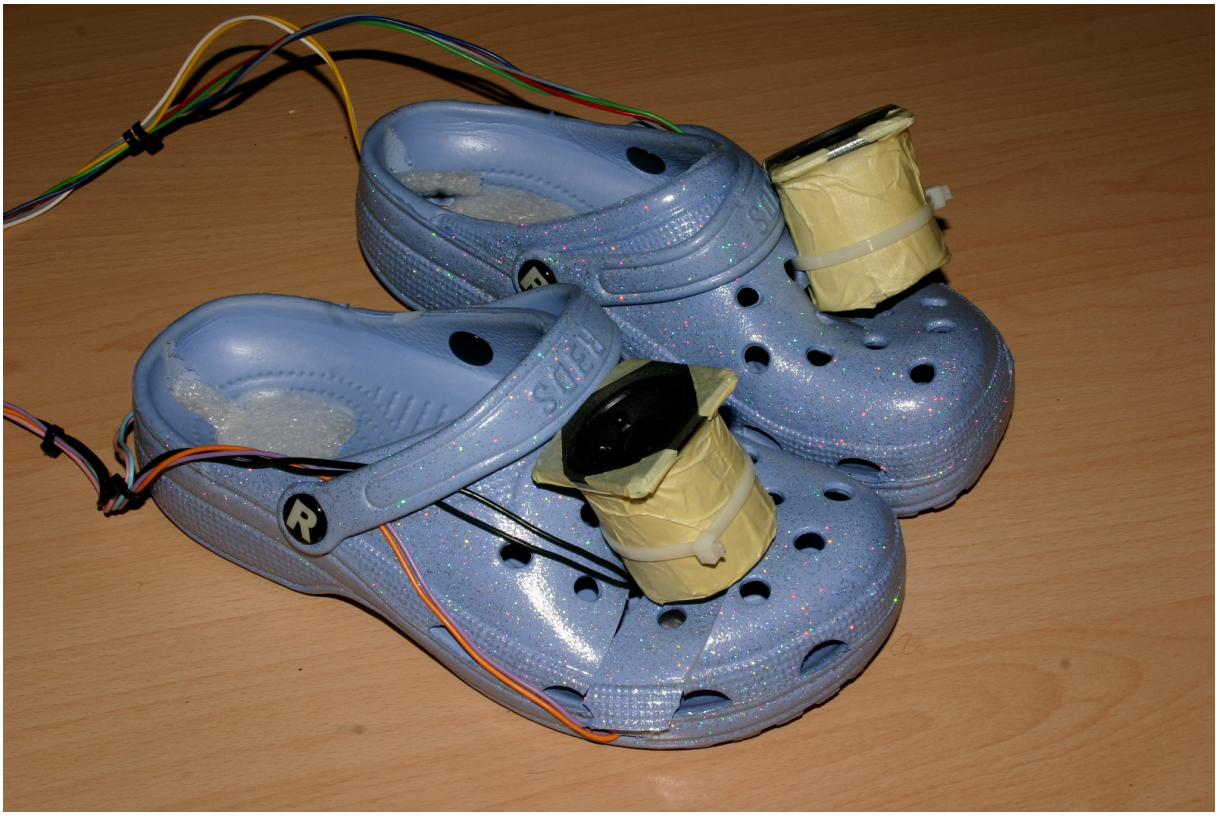


Figure 3.1: Shoes chosen for the first prototype: a pair of plastic sabots. The wires are connected to the sensors and to the shoe-mounted speakers.

Although the two prototypes differ each other from the implementation and the physical points of view, they follow quite the same basic ideas: in order to recreate walking experiences with auditory cues in fact, the following operations must be executed in exact order.

Figure 3.2: Shoes chosen for the second prototype: a pair of sneakers.

1. In-shoe measurement of physical quantities
2. Digitization and transmission of the acquired signal
3. Sound synthesis
4. Acoustic display

Points 3 and 4 refer to other parts of the deliverable, while points 1 and 2 are described in the present part (more precisely in sections 3.2.1 and 3.2.2) from both the theoretical and the implementation points of view; differences between the two prototypes are also pointed out.

3.2 Contact based interface

This chapter describes in details the hardware system made for measuring physical quantities with sensored shoes and the software written for sending the so acquired data to the computer which provides sound synthesis: section 3.2.1 reports details of the hardware part, while section 3.2.2 is devoted to the software.

3.2.1 In-shoe measurement of physical quantities

Since the setup must in the end, offer a simulation of the walking experience from an acoustic point of view, it has to provide the user with the most significant auditory cues related to the action of walking: these reside mainly in the sounds of footsteps caused by the interaction between shoes and ground. This interaction is well represented by the *variation* in the intensity of *force* due to the walker's weight and exerted by the foot/shoe on the ground itself: with this in mind, an effort was made from

the beginning in designing a *force-sensing* apparatus placed inside each shoe. Such a sensing system requires some indispensable features, mainly related to its peculiar application: first of all, it has to be *reliable* from a mechanical point of view. The action of human walking in fact, implies the repeated application of quite strong forces on walker's feet: sensors must then withstand mechanical stress as long as possible without breaking. This implies that possibly, quite strong sensors should be chosen and they should be placed in the shoe in a way that they could not be damaged by the walking action. The sensing devices should also be compliant with the force range involved: in particular they should not be *too sensitive* in order to avoid saturation; their typical response characteristics should also be kept in mind in order to provide a proper signal conditioning if needed. However, from a practical point of view and keeping in mind the project's aim, doesn't really matter if the force values typically exerted exceed the sensitivity range of the chosen sensors during normal walking, since the whole setup is not aimed to provide an exact measure of a person's weight, but (as previously said) to detect changes in the intensity of the exerted force. Moreover, the number and the position of the sensing devices in the shoe can affect the accuracy of force detection and measurement [4]. Another relevant point is related to the interaction purposes of the whole shoe-based system: as well known, a real-time interactive system must provide as *lower latency* as possible, thus the sensors should not introduce latency, i.e. their response must be as faster as possible. The last requirement for the sensing devices concerns the wearability of the complete setup: in particular, the walking action should not be affected by the presence of sensors and other hardware. This implies that the chosen devices should be *small* and *lightweight*. Price is also an issue.

Since we decided to avoid custom-made solutions, a thorough research for commercial ones satisfying all the aforementioned issues was performed at the beginning, but it soon pointed out that a product being able to completely afford all the above requirements did not exist: the best choice was represented by the *Force Sensing Resistors* (FSR from now on) by Interlink Electronics[1]. These are polymer thick film (PTF) devices which exhibit a decrease in resistance with an increase in the force applied to the active surface (see Figure 3.3). Although their force sensitivity is optimized for use in human touch control of electronic devices (see Figure 3.4), they are quite compatible with the above listed requirements. More details on FSR can be found on the Interlink website [2].

[1] `www.interlinkelectronics.com`

[2] `www.interlinkelectronics.com/library/media/papers/pdf/fsrguide.pdf`

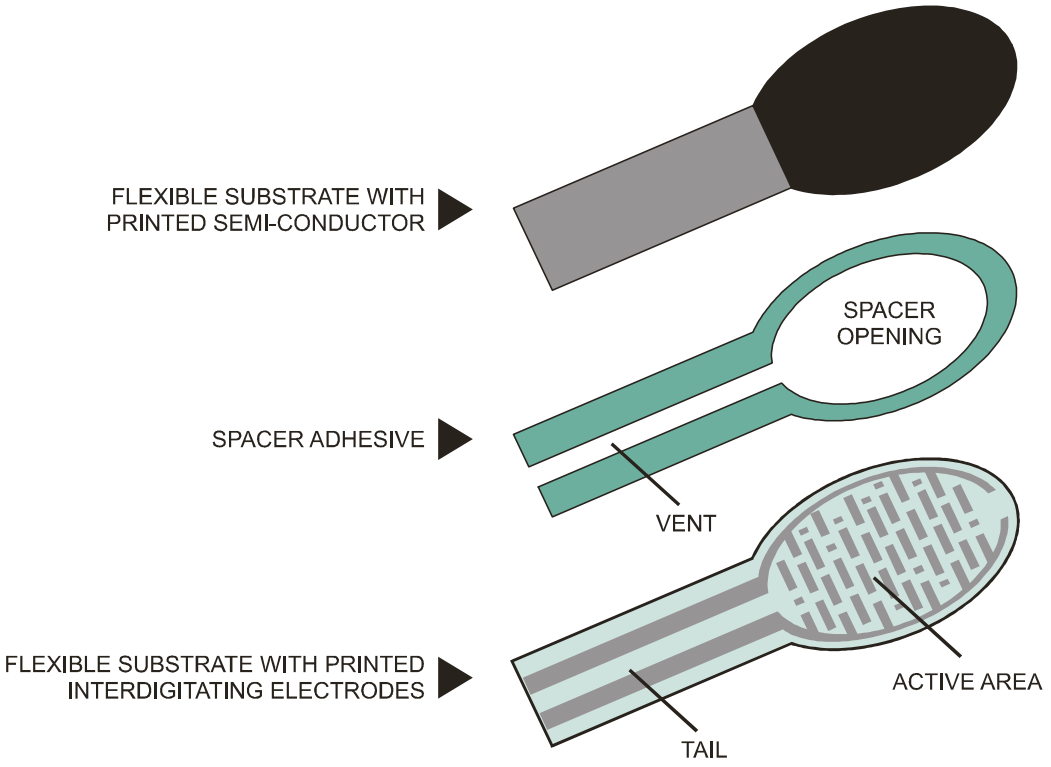


Figure 3.3: Construction layers of an FSR.

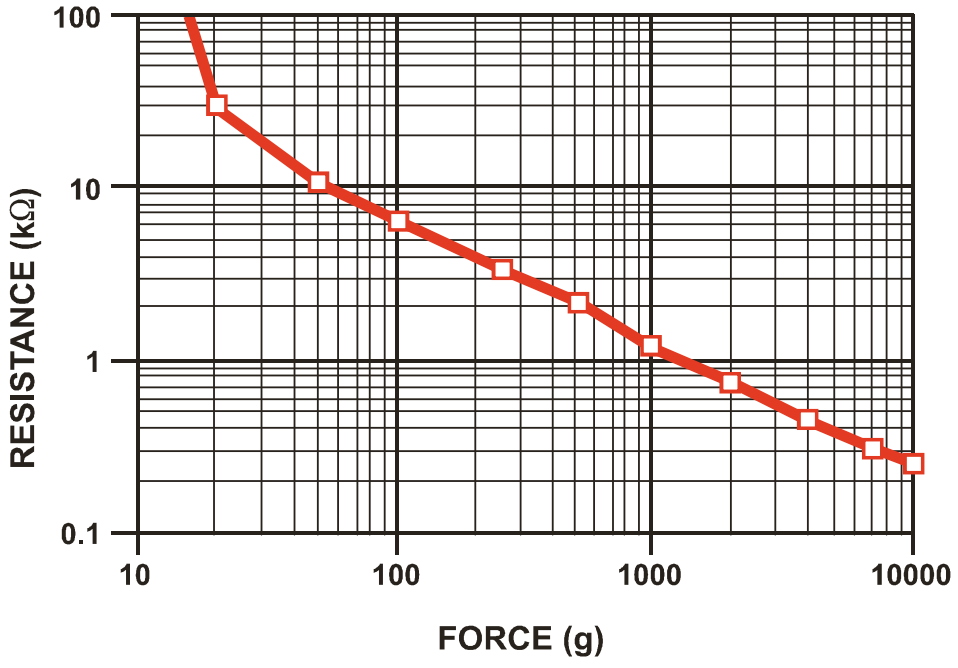


Figure 3.4: Force versus resistance characteristics.

We decided to detect force in two points under the foot: *tip* and *heel*. Two sensors were thus placed in each shoe: one FSR model 402 under the heel and one FSR model 400 under the tip. For what concerns the physical assembly, there are differences between the two prototypes: for the first one we decided to place the sensing devices directly into the sole. To this purpose, two holes were made in the sole using a small burr: the holes were shaped in order to easily host the sensors (see Figure 3.5). The sole was drilled to allow wires to pass thru.

Figure 3.5: Physical placement of sensors in the first prototype. Some pieces of rubber are placed on the sensors in order to protect them and to transmit force from the bottom of the foot to the active layer of each of them.

For the second prototype, something different was made: the sensing devices are fixed directly on a couple of rigid plastic plates attached to the bottom of the insole with some scotch tape (see Figure 3.6). The two plastic plates act as mechanical decoupling devices, in order to avoid force detection area to be limited by sensor's size and shape. For acquiring force data and sending them to the computer, the sensors are connected to an *Arduino Duemilanove*[3] USB board. Each one of the four sensors forms a voltage divider together with a 33kΩ resistor: each divider is connected between the 5V output of the Arduino and ground. Signal is taken between the FSR and the resistor and sent to one Arduino's analog input. Figure 3.7 schematizes the circuit here described. We decided to void the use of buffers (i.e. voltage followers) between the voltage divider and the analog inputs in order for keeping the circuitry as simple as possible: according to the Atmel ATmega168 user's manual in fact, the input impedance of the microcontroller's built in ADC is high enough to avoid the use of buffers. Moreover, no antialiasing low-pass filters are needed because the frequency spectrum of the involved signals is compatible with board's sampling frequency. The Arduino-based setup is used in both the first and the second prototypes.

[3] `www.arduino.cc`

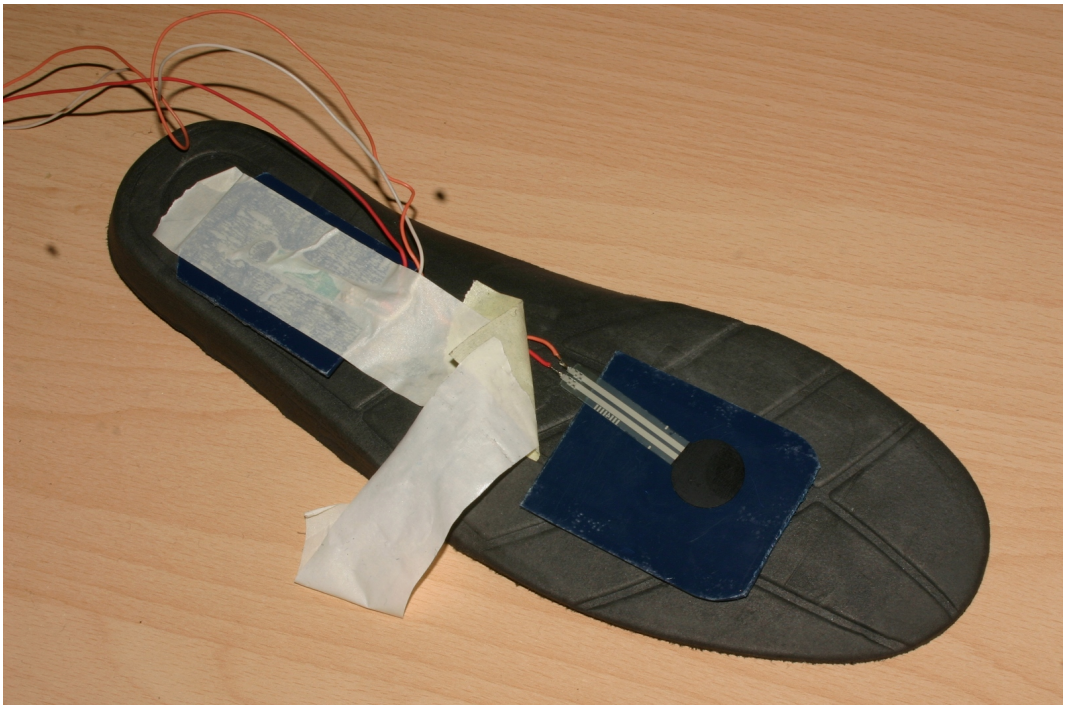


Figure 3.6: Sensors in the second prototype.

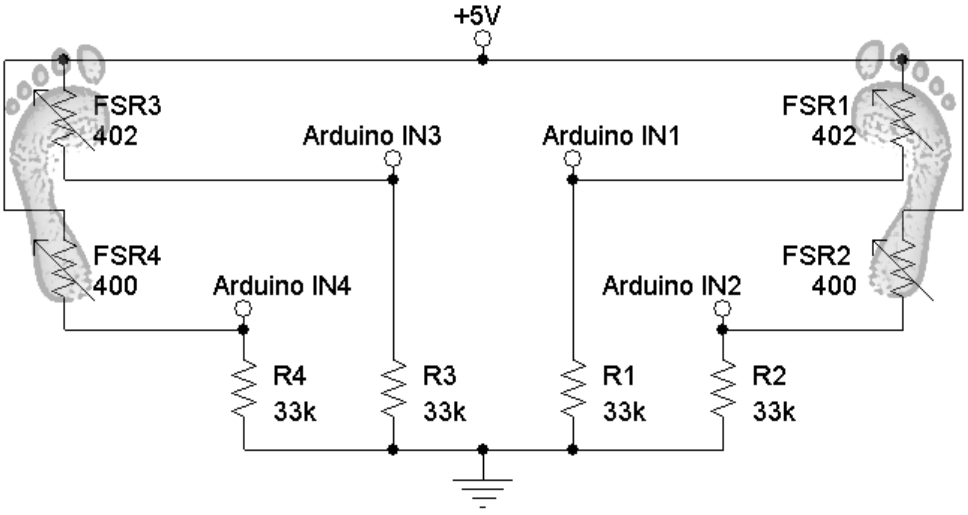


Figure 3.7: Schematic of the sensing circuit connected to the Arduino.

Although the force measurement system is present in the first as well as in the second prototype, the latter features also *slip detection*. The method chosen for detecting slip is based on mechanical contact between shoe and floor and requires the presence of *mechanical friction* between the two surfaces. A mini-trackball [4] is placed in a hole made in the sole of each shoe (see Figure3.8): when the shoe slips, the trackball rolls against floor closing periodically at least one of the four internal micro-switches. Each switch provides detection of movement along one of the four default directions

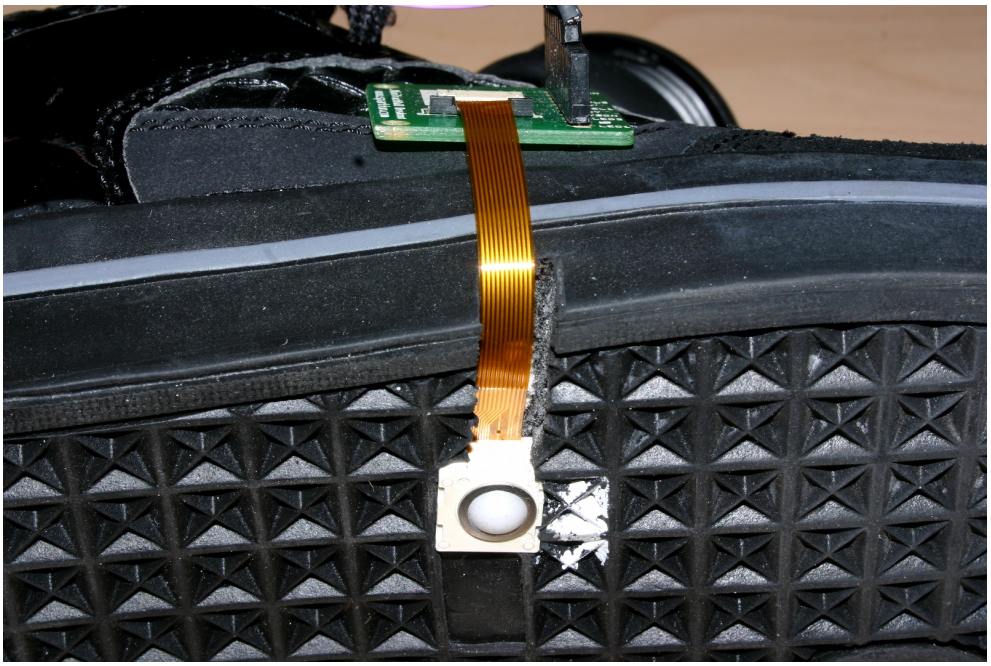


Figure 3.8: Second prototype: the trackball.

(see Figure 3.9). Although the trackballs allow detection along four directions, we decided to avoid discrimination between the positive and negative ones, thus keeping in account for slipping along the X and Y axis only (by connecting both terminals of the switches for detecting rolling along one axis to Vcc). Moreover, a special switch is closed when the trackball is pressed, i.e. when the shoe is pushed against floor. Since a trackball is basically a set of on-off switches, its outputs can be connected to some of the Arduino's digital input lines using for example a *pull-down* resistor configuration for each line: this is what we decided to do for interfacing the trackball with the acquisition board. Figure 3.10 shows the schematics of the circuit made to this purpose, while figure 3.11 depicts the PCB realized for interfacing the sensors in the two shoes (FSRs and trackballs) to the Arduino: the board has been made for the second prototype.

[4] `www.sparkfun.com/datasheets/Components/Trackball.pdf`

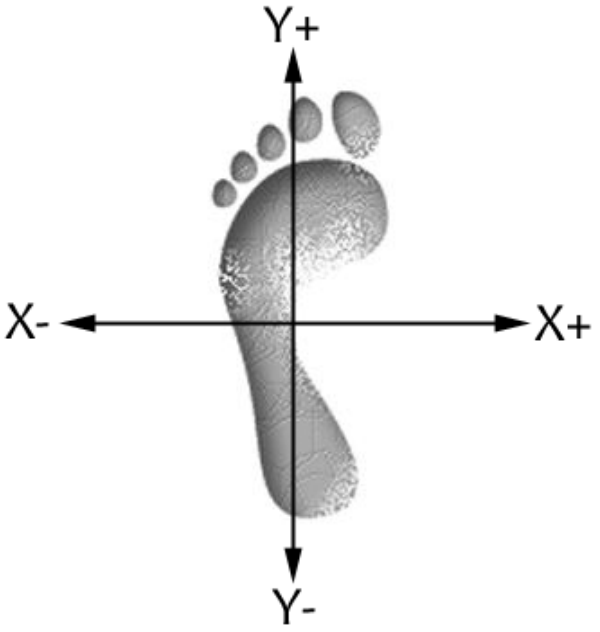


Figure 3.9: Axis orientation for slip detection by mini-trackball.

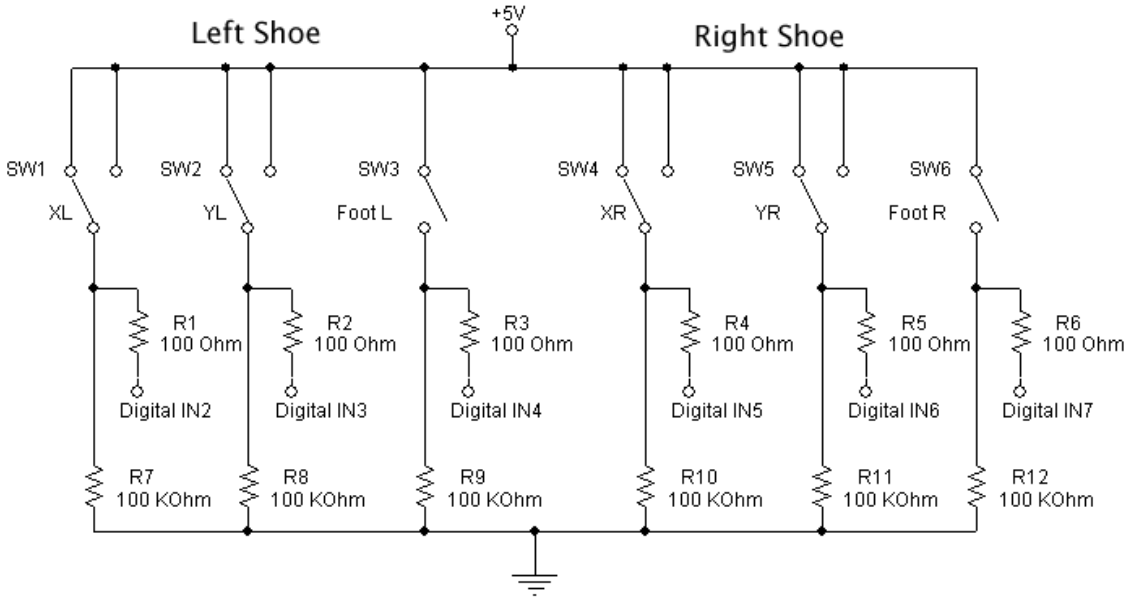


Figure 3.10: Schematics for interfacing the two trackballs to the Arduino.

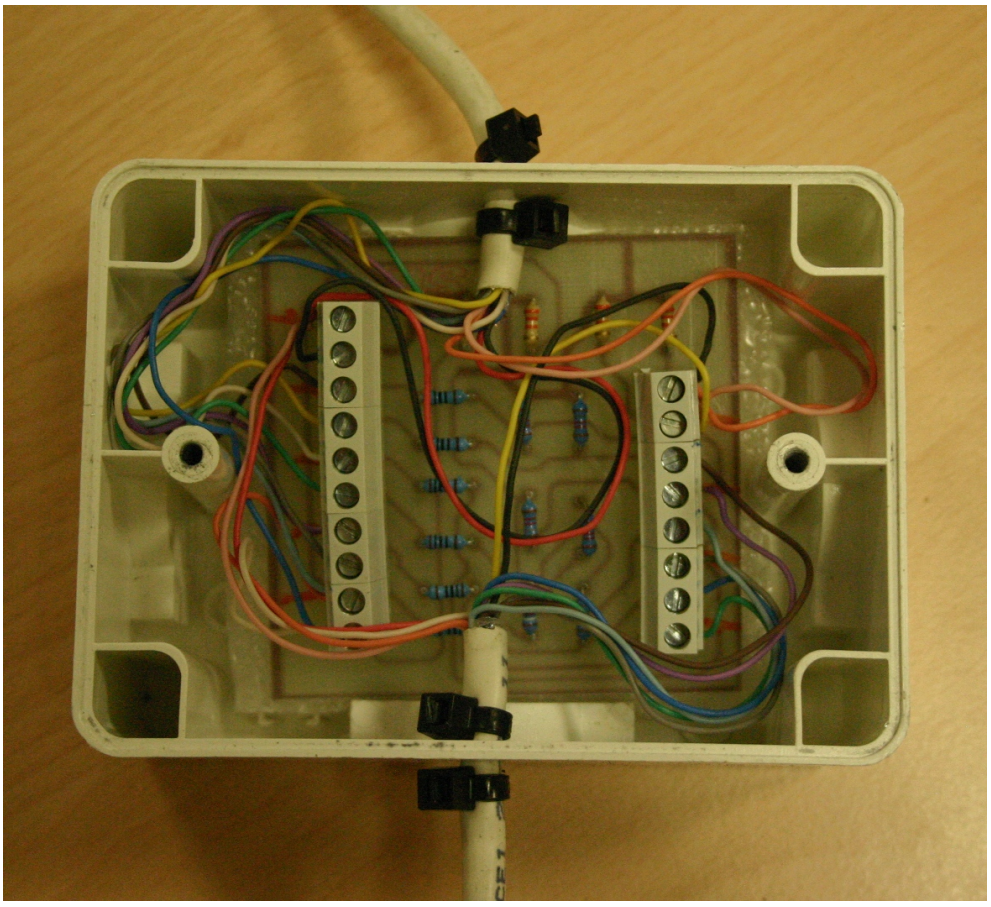


Figure 3.11: Board for interfacing the sensors to the Arduino (second prototype).

3.2.2 Digitization and transmission of the acquired signal

As previously said, force signals are digitized by the ADCs embedded in the Arduino's microcontroller (Atmel ATmega168): in order to apply some conditioning on the so acquired data and sending them to the synthesis engine, additional software has been developed. This is divided in two main blocks: the first one is the *firmware* written for the Arduino board, the second is constituted by a patch for the *Pure Data* environment. The main function of the patch is to listen to the port to which the Arduino board sends data. Two different versions of both firmware and patch have been developed for the first and the second prototype respectively.

3.2.3 Software for the first prototype

The firmware written for first prototype is constituted by a main loop in which the four force values are sequentially read with a polling-like approach: considering the i-th instant, each force value is considered valid only if the absolute value of the difference between the correspondent value at the $i-1$-th instant is equal or greater than 32. If at least one of the four values is valid, a data packet is built and written to the serial port. Each packet is made by four ASCII strings, each one representing the force value on the corresponding sensor at the i-th instant. Strings in the packet are ordered as follows:

```
right_tip_value right_heel_vale left_tip_value left_heel_value
```

Strings are separated by a space character and packet ends with an end-of-line character. The aforementioned threshold control method, although reducing in some way the conversion resolution, avoids band saturation on the port. The Pure Data patch listens to the serial port corresponding to the Arduino's driver: each packet is unpacked in four strings which are then converted to numbers. The so obtained values are considered valid if above a fixed threshold. Now, for each sensor, the first order derivative of the sequence of received numbers is computed: only positive variations are considered valid and sent to the corresponding synthesis engine (one engine for every sensor). Figure 3.12 shows the aforementioned Pure Data patch.

3.2.4 Software for the second prototype

In order to limit the required bandwidth for sending data, considering also that the second prototype sends sliding data too, a more optimized transmission protocol has been implemented. The idea is to transmit binary data instead of ASCII coded values: this implies changes in the firmware as well as in the Pure Data patch. Firmware is in this case based on the same polling approach with threshold control used in the first prototype: differences reside in the data packet assembly. Every packet is formed by *two bytes* and is structured as shown in Figure 3.13. Where the four bits signed with C identify the channel number (16 channels are allowed), while the ones signed with V identify the measure value (the ADCs have $10bits$ resolution). The protocol is structured as follows:

- 10000vvv 0vvvvvvv right tip force value;
- 10001vvv 0vvvvvvv right heel force value;
- 10010vvv 0vvvvvvv left tip force;
- 10011vvv 0vvvvvvv left heel force value;
- 10100000 00000001 left trackball X switch;
- 10100000 00000010 left trackball Y switch;
- 10100000 00000011 left trackball foot switch;

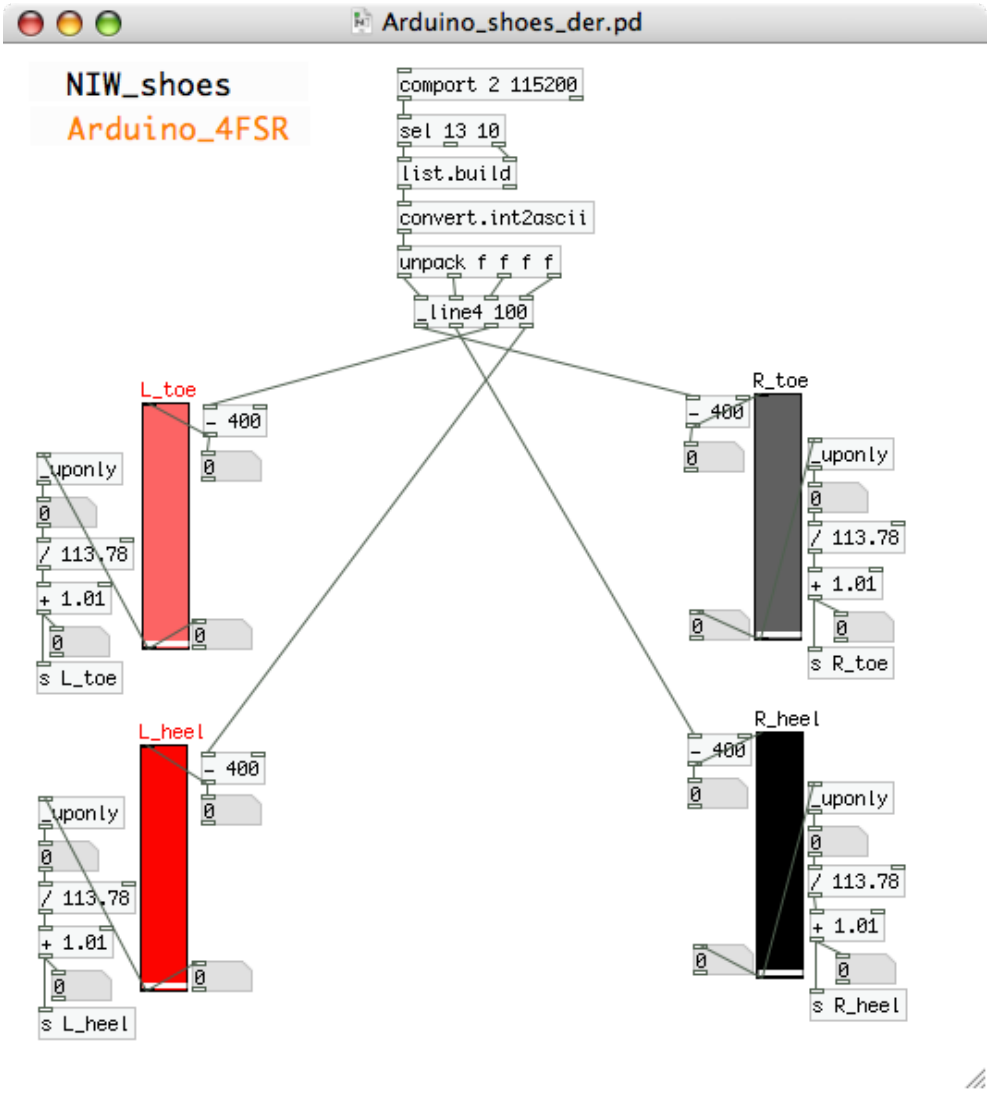


Figure 3.12: Pure Data patch for signal acquisition (first prototype).

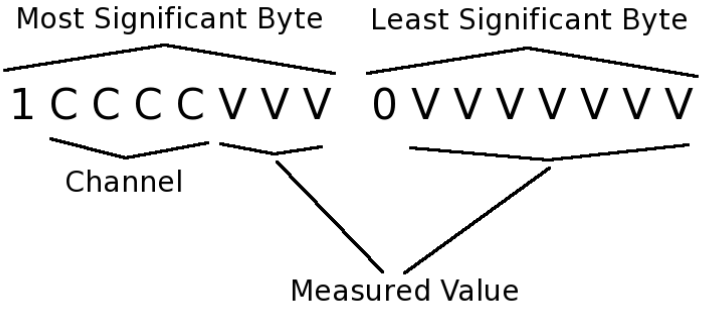


Figure 3.13: The two bytes structure of the protocol.

- 10101000 00000001 right trackball X switch;
- 10101000 00000010 right trackball Y switch
- 10101000 00000011 right trackball foot switch

For the second prototype, two different firmwares and corresponding patches had been developed: all of them are based on the above mentioned protocol but are aimed at different purposes. The first firmware and patch pair has been made for producing data required by the *crumpling* synthesis engine, the second fits the requirement of the *soft impact* engine [9]. For what concerns the crumpling, the same approach of the first prototype has been followed, of course the polling procedure is extended also to fetch data from trackballs. The Pure Data patch unpacks data packets and sends force values to the proper synthesis engine, while trackballs signals are displayed in a graphical way (see Figure 3.14). The firmware written for the soft impact model continuously reads the four force values: considering

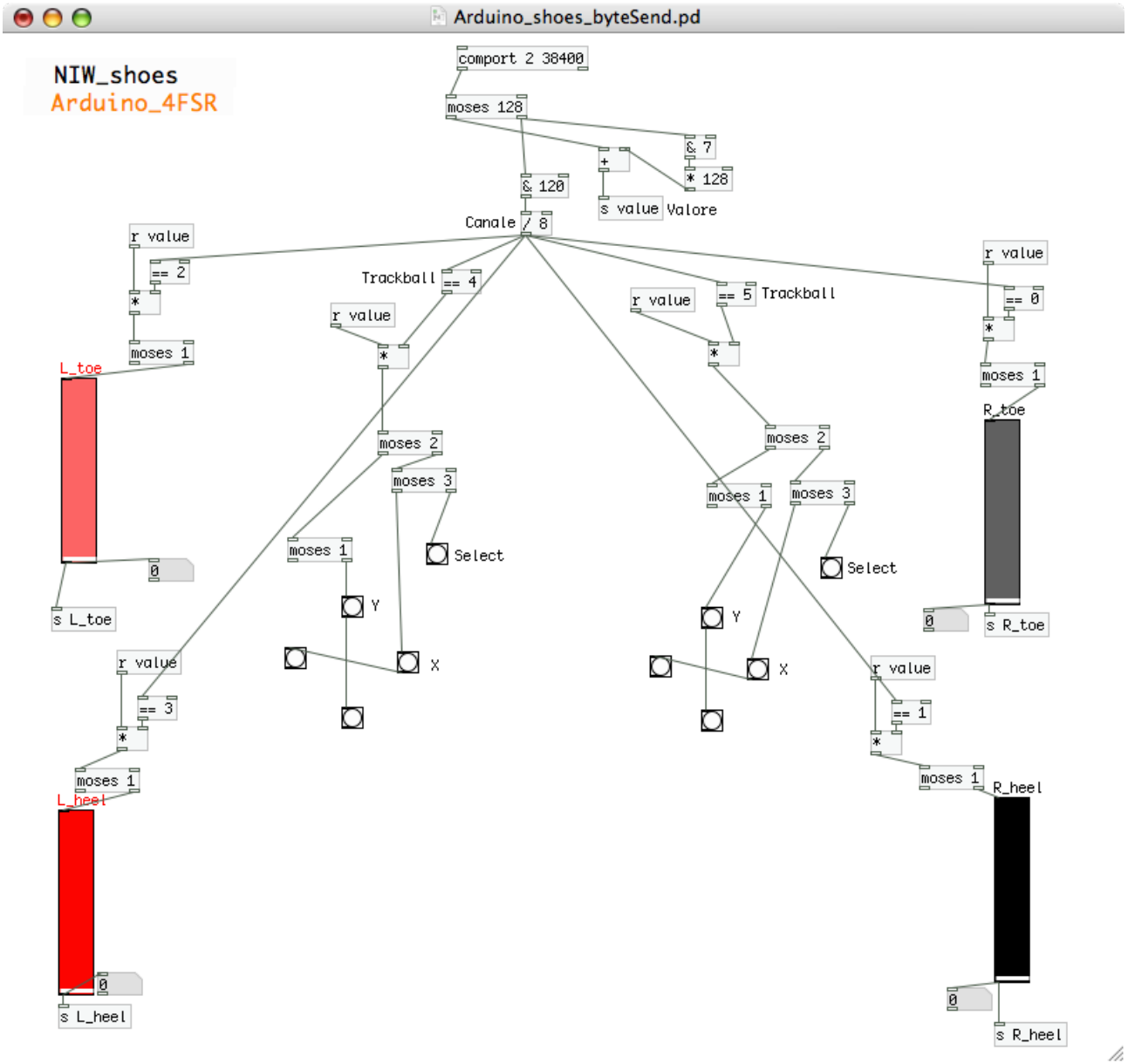


Figure 3.14: Pure Data patch for the crumpling model in the second prototype.

one sensor, when the read value is above a chosen threshold, after a fixed waiting interval, the system reads the value again. The last acquired number is packed and sent as force value to the soft impact

synthesis engine. The Pure Data patch in this case, only unpacks data and routes values to the proper engine.

4 Acoustic extraction of ground reaction forces for real-time synthesis of walking sounds (AAU)

The goal of this research is the development of a shoe-independent system able to produce synthesized footsteps sounds in real time during the walk of a user.

We are particularly interested in developing a solution which requires a minimum amount of sensing technology.

The system proposed renders footsteps sounds on floors of different types of material giving to the user the impression of walking on a floor different from that the user is on.

4.1 The setup

In this paper, we describe two different configurations of our sensing technology. Both configurations are based on the use of microphones. In the first configurations the microphones are attached to the shoes of the subjects, while on the second configuration they are placed on the floor. In Section 4.3 we discuss advantages and disadvantages of both configurations.

4.1.1 The first setup

In our preliminary experiments, we used a system composed by two small contact microphones, a sound card, a computer running Max/MSP[1] and a set of headphones.
We tried two different configurations: one with wired and one with wireless microphones. In both cases, the microphones were attached to the shoes of the test subjects. In the wired case, the two microphones (one for each shoe) were placed on the shoes by means of two adhesive gums (see Figure 4.1). Precisely, they were attached on the exterior part of each shoe, at about 4 cm from the sole and at the center between heel and toe. The microphones were connected to the sound card by means of wires, and the wires were attached to the trousers by means of velcro.

In detail, each microphone is a back-electret pre-polarised condenser mic capsule (pressure receiver) with integrated impedance transducer. Precisely, we used the model KE4-211-1 by Sennheiser.[2]
The captured sounds were conveyed to the computer by means of a Fireface 800 sound card.[3]

In the second configuration, we used wireless contact microphones (Sennheiser SK 500 (evolution wireless G2)), that had better sensitivity and frequency response compared to the previous situation, and the obvious advantage that the users could move without any restriction. Each contact microphone was attached to the shoe in the same way as in the previous configuration. The wire connecting the microphone to the transceiver was attached to the clothes of the user.

4.1.2 The second setup

In a second configuration we adopted a set of non-contact microphones placed on the floor. In our experiments we used the Shure BETA 91[4], a high performance condenser microphone with a tailored frequency response designed specifically for kick drums and other bass instruments. Its features made it a good candidate for our purpose of capturing the footsteps sounds. In our experiments we placed

[1] http://www.cycling74.com
[2] http://www.sennheiser.com/
[3] http://www.rme-audio.com/english/firewire/ff800.htm
[4] http://www.shure.com/

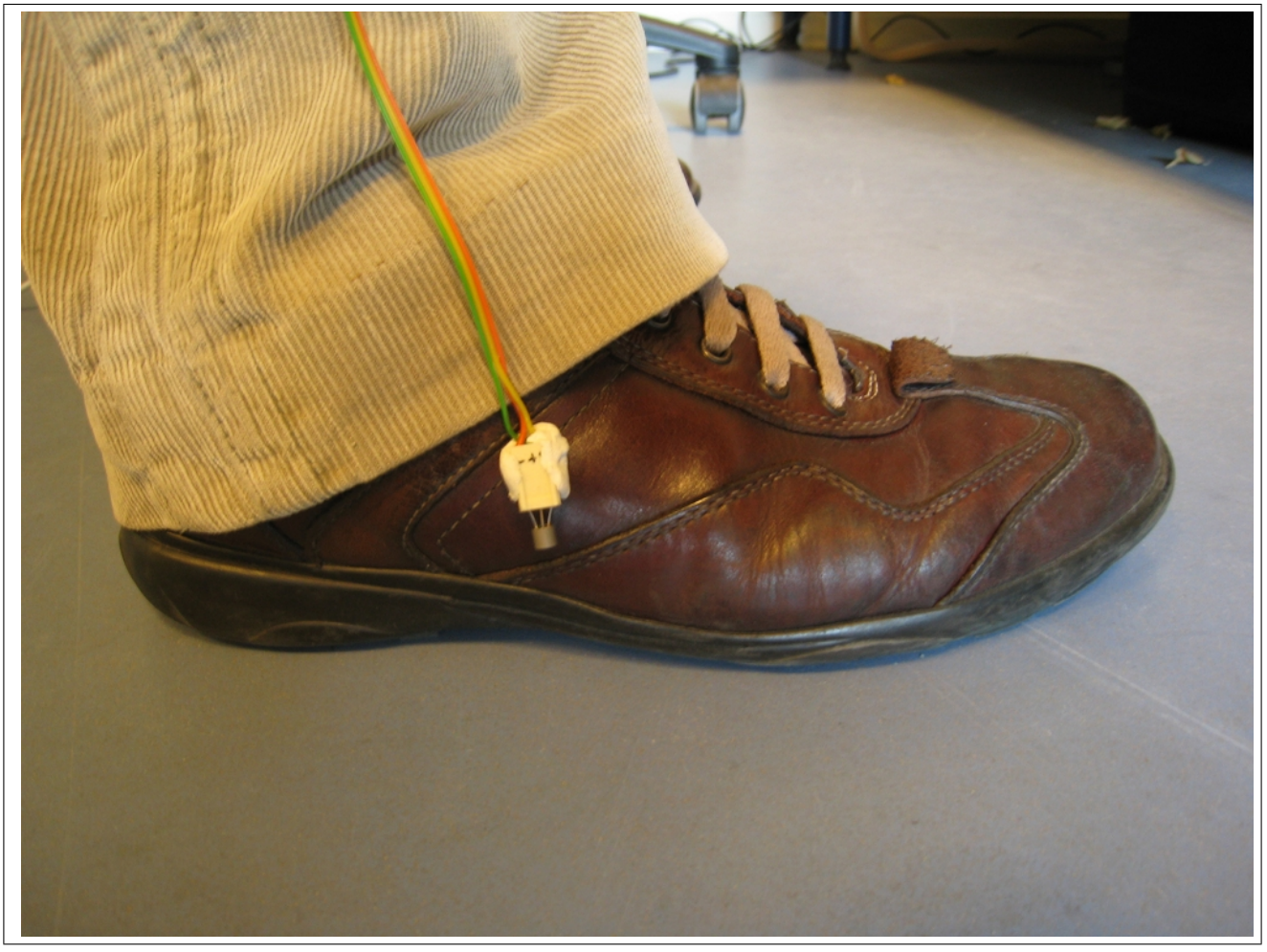


Figure 4.1: The contact microphone attached to the shoe.

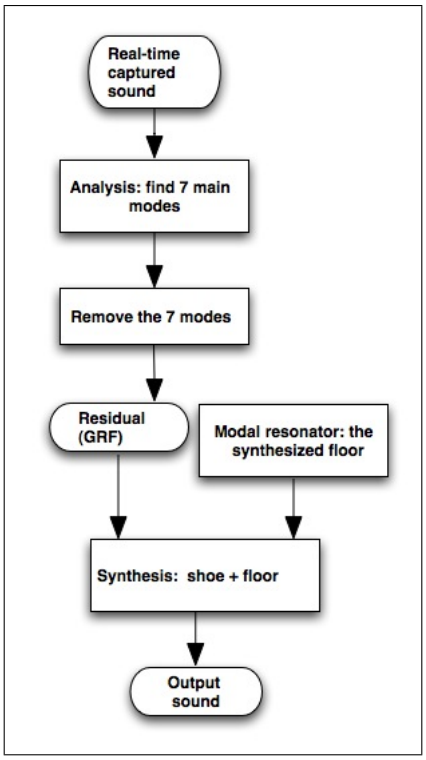


Figure 4.2: Block diagram of the analysis-synthesis algorithm used.

a couple of such microphones on the floor at 1.5 meters distance from each others. Such an approach still held the requisite of shoe independence, and made the requisite of wearability unnecessary.

4.2 Sound analysis: extraction of the ground reaction force

A footstep sound is the result of multiple micro-impact sounds between the shoe and the floor. The set of such micro-events can be thought as an high level model of impact between an *exciter* (the shoe) and a *resonator* (the floor). In such a vision the sound captured by the microphones can be considered as a composition of both these two components.

To extract the contribution of the shoe to the sound we used three techniques. The first was based on modal analysis and resynthesis (see Figure 4.2), the second on linear predictive coding (LPC), while the third on an amplitude follower. The three approaches are described in the following.

4.2.1 First approach

In order to achieve the final goal of producing the sensation of walking on floors made of different kinds of materials, we thought to remove the contribution of the resonator, keep the exciter and consider the latter as input for a new resonator that implements different kinds of floors. Subsequently the contribution of the shoe and of the new floor are summed in order to have a complete footstep sound. In the field of mechanics, such exciter is usually called ground reaction force (GRF), i.e., the reaction force supplied by the ground at every step. The aim of the phase of analysis consisted on finding some parameters that allowed us to extrapolate the exciter from the captured sound, i.e., finding the GRF from the acoustic waveform. Such an extrapolation consisted in removing, from the spectral representation of the sound, the main resonant frequencies, i.e., the modes. The first algorithm we used to remove such modes was by using a connection of notch filters. Such algorithm needed, as input parameters, the frequency and the bandwidth of the modes.

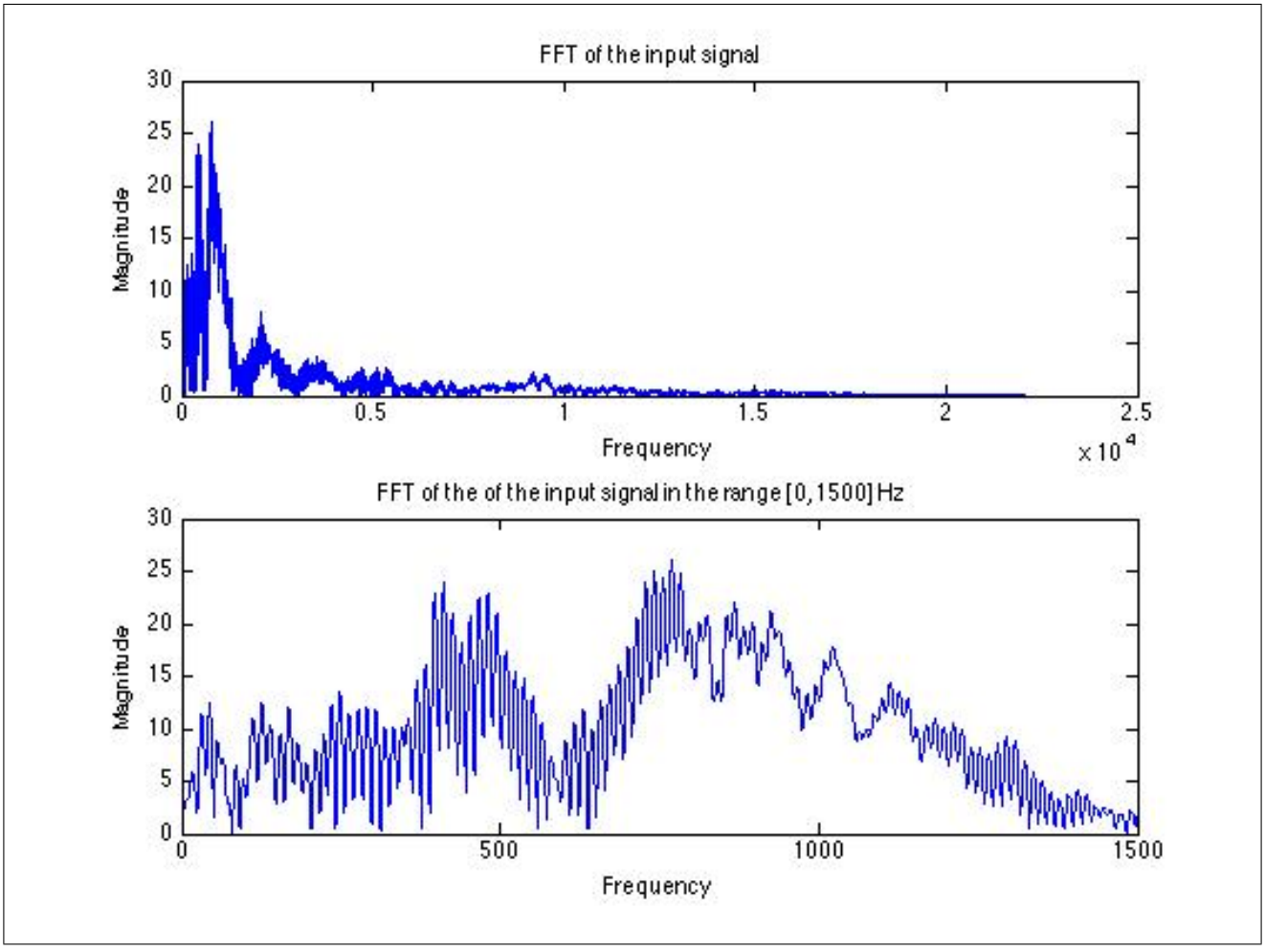


Figure 4.3: FFT of a sound of footstep on a concrete floor.

The algorithm has been implemented as a Max/MSP external. For computational efficiency, the external implements both the analysis and synthesis steps. The algorithm works as follows: the sound of a person walking is detected in real-time by the microphones described above. From this sound, the resonances corresponding to the impact of the shoe on the floor are removed, in order to extract the GRF. We assumed that the user utilizes our system walking on solid and homogeneous floors (as the concrete one for example). After some off-line analysis conducted on recorded sounds of footsteps on such types of floor, we concluded that, in average, the main modes (i.e., the main resonances present in the recordings) appear at low frequencies, and precisely they are included in the range of frequencies between 0 and 1500 Hz (see Figure 4.3). We established to consider seven main modes, two of which were chosen in the range (0,300) Hz, the other five in the range (300,1500) Hz. This choice is motivated by the fact that in average seven main resonances are detected in the recordings. Such resonances were detected in real-time, and removed by means of seven corresponding notch filters. We considered two distinct parts of the spectrum because we became aware of the great contribute of the modes at very low frequencies.

Figure 4.4 shows the time domain representation of an original single footstep sound (top) and or the same sound where the main resonances have been extracted. Such sound was used to test the algorithm offline, and obtained from the Hollywood Edge sound effects database.[5] The main two temporal components of the footstep sounds, i.e., the heel/toe event can be easily noticeable.

Figure 4.5 shows the spectrum of the residual sound after the removal of the seven main modes from the sound of the footstep on the concrete floor analyzed in Figure 4.3.

The obtained residual sound was used as the exciter part of the real-time modal synthesizer described in the following section.

4.2.2 Second approach

The second approach we followed to remove the main resonances of the sound and produce a residual, was by using linear predictive coding (LPC). We implemented an external for Max/MSP able to produce in real-time the residual (i.e., the error signal) resulting from a LPC analysis conducted with order up to 31. Our idea was to give such a residual in input of a LPC resynthesis filter (of the same order of the analysis filter), modeling the resonance structure of different kinds of floor. We also implemented a LPC resynthesis filter in another external, whose coefficients were found thanks to an off line analysis conducted with MATLAB on recorded footstep sounds. For our analysis and resynthesis we chose an order equal to 10. Unfortunately we realized that the LPC analysis did not work as desired on the most part of the walking sounds we analyzed, since the estimated signal was very far from the original one, and consequently the signal error was very high. We noticed that the analysis performed better on some solid surfaces (in particular wood and concrete, see Figure 4.6). This was not a sufficient condition, since the LPC analysis on a metal floor produced bad results as 4.7 shows. The error signal was much higher on non homogeneous floors like gravel, dirt, leaves or snow.

4.2.3 Third approach

We decided to consider the amplitude envelope of the input sound as the overall GRF. An envelope extraction therefore corresponded to the calculation of the GRF. To perform envelope extraction we used a simple non-linear low-pass filter proposed in [7]:

$$e(n) = (1 - b(n))|x(n)| + b(n)e(n-1)$$

where

$$b = \begin{cases} b_{up} & \text{if } |x(n)| > e(n-1) \\ b_{down} & \text{otherwise} \end{cases}$$

[5] www.hollywoodedge.com/

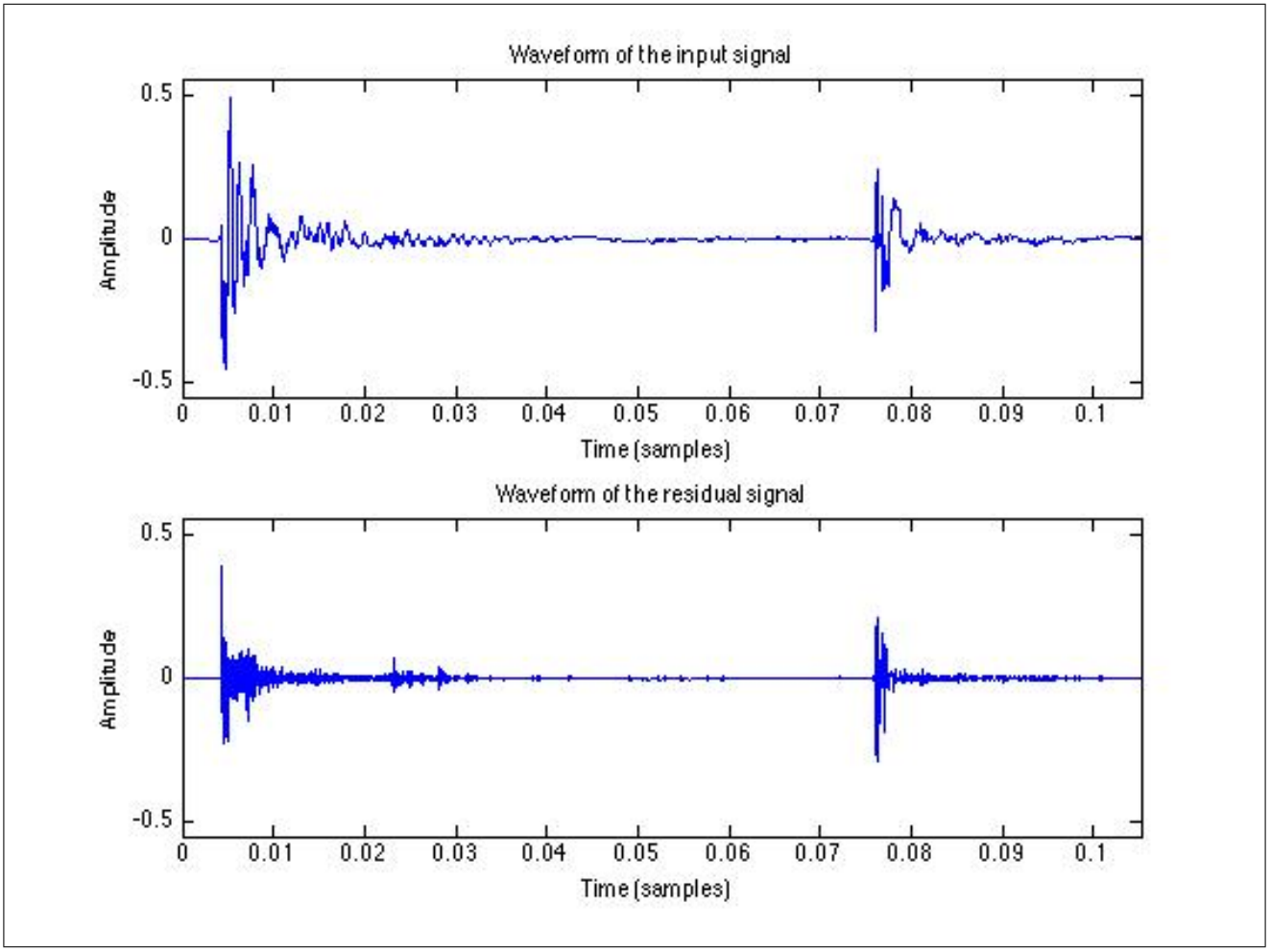


Figure 4.4: Time domain representation of a recorded footstep sound (top) and the corresponding residual (bottom). Notice how the two main temporal components of the footstep sound, i.e., heel and toe events, can be easily noticeable.

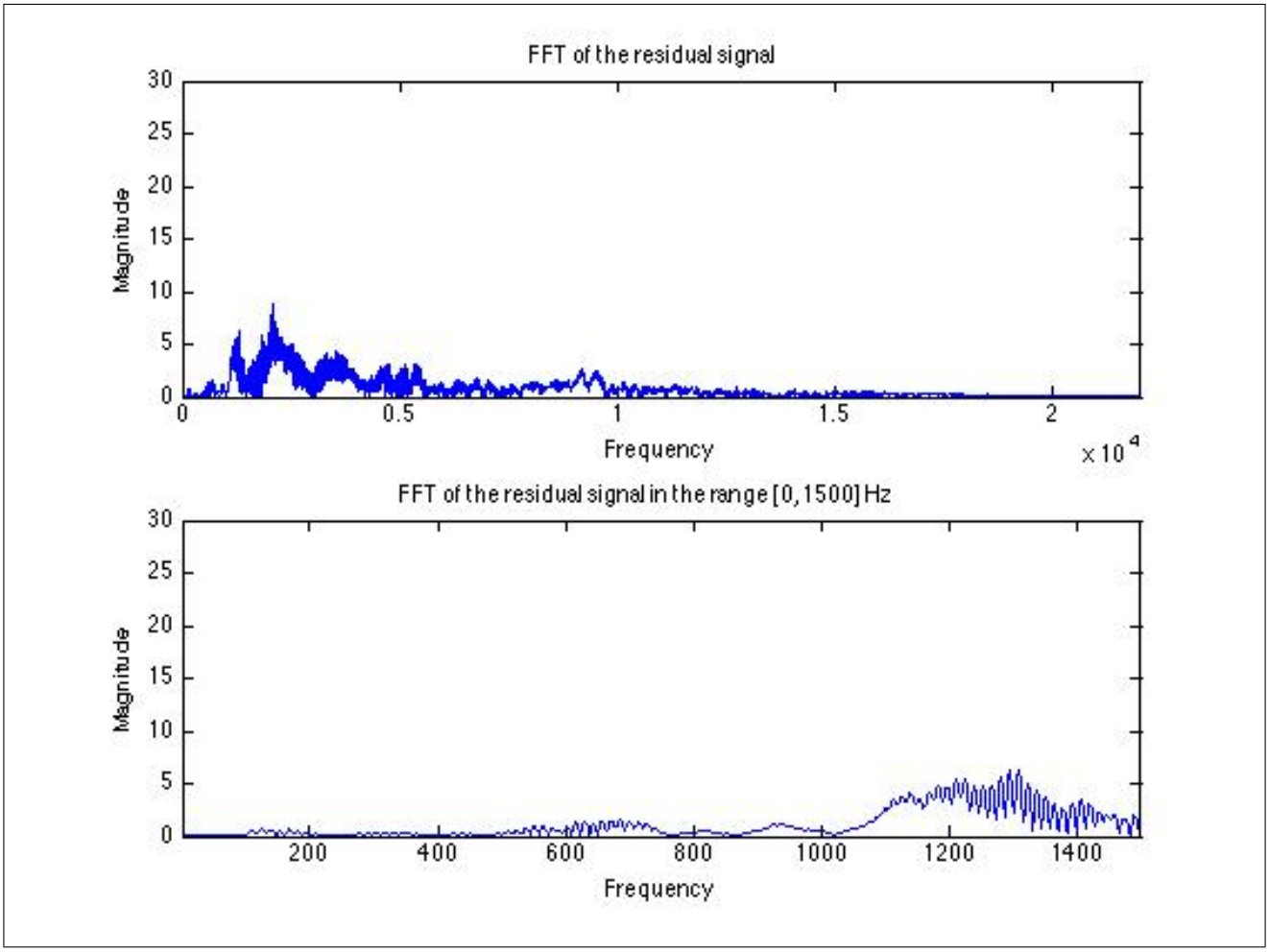


Figure 4.5: FFT of the residual sound analyzed in Figure 4.3.

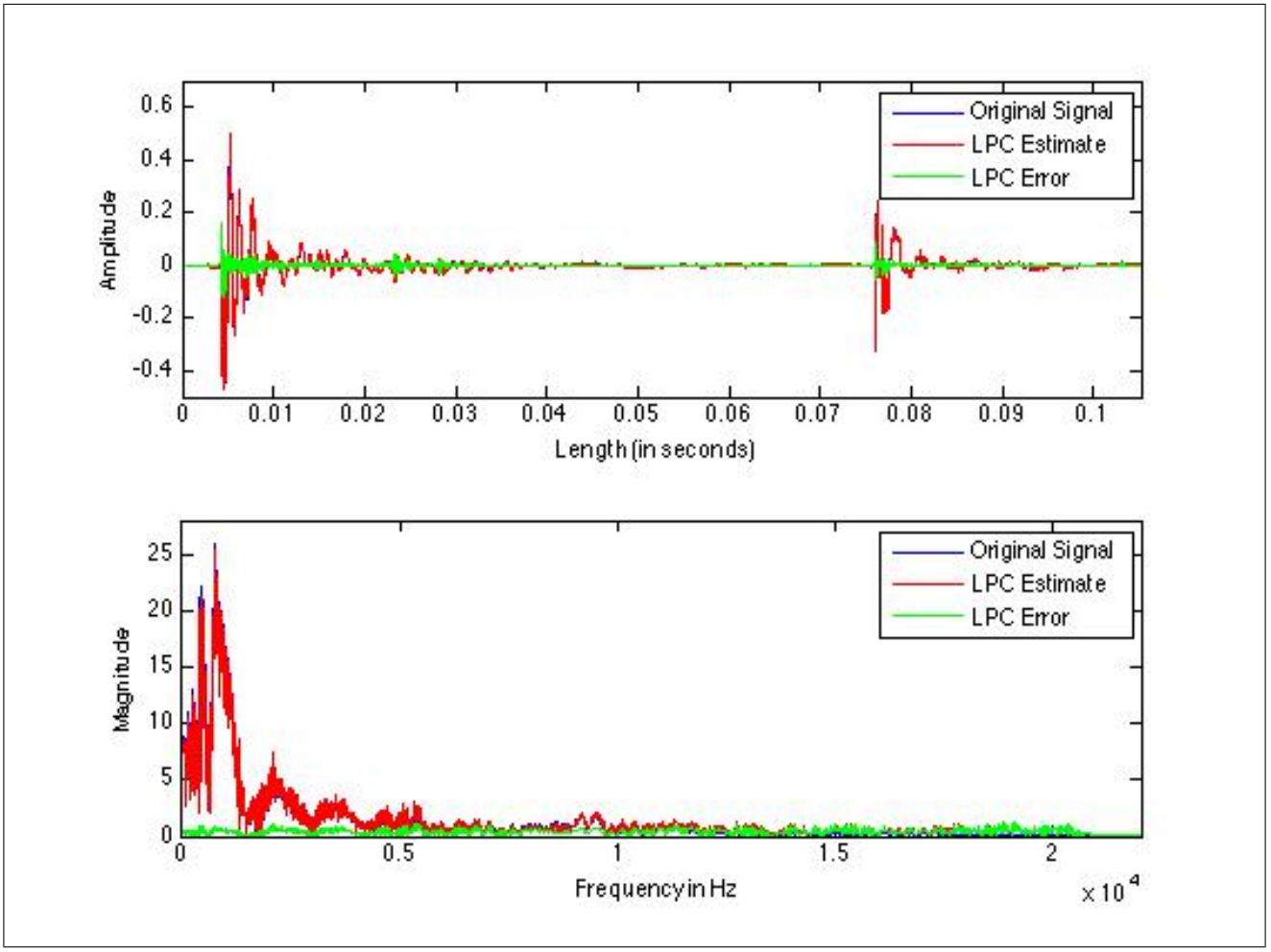


Figure 4.6: Original footstep sound on a concrete floor, its LPC estimate and the relative error (residual).

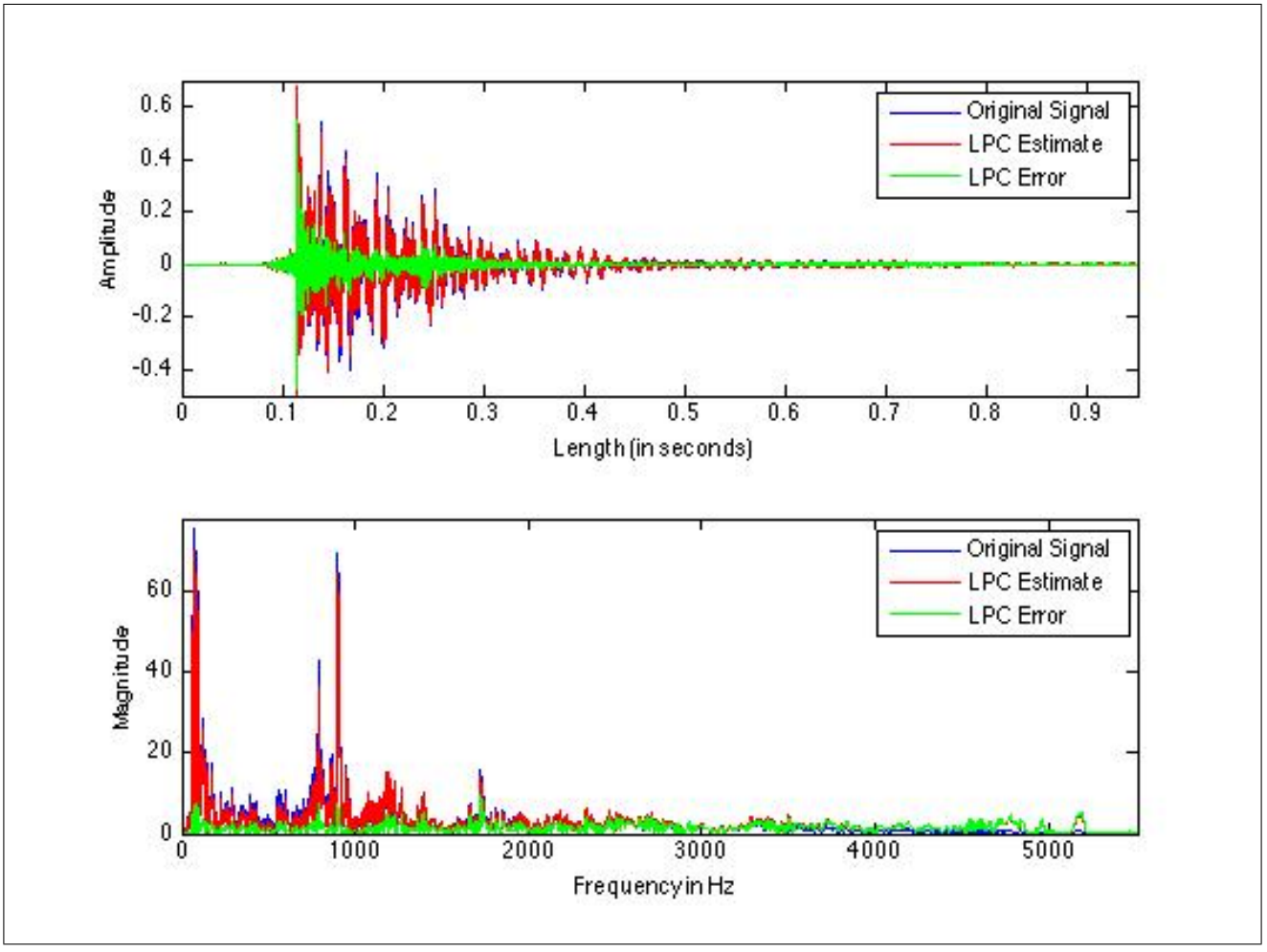


Figure 4.7: Original footstep sound on a metal floor, its LPC estimate and the relative error (residual).

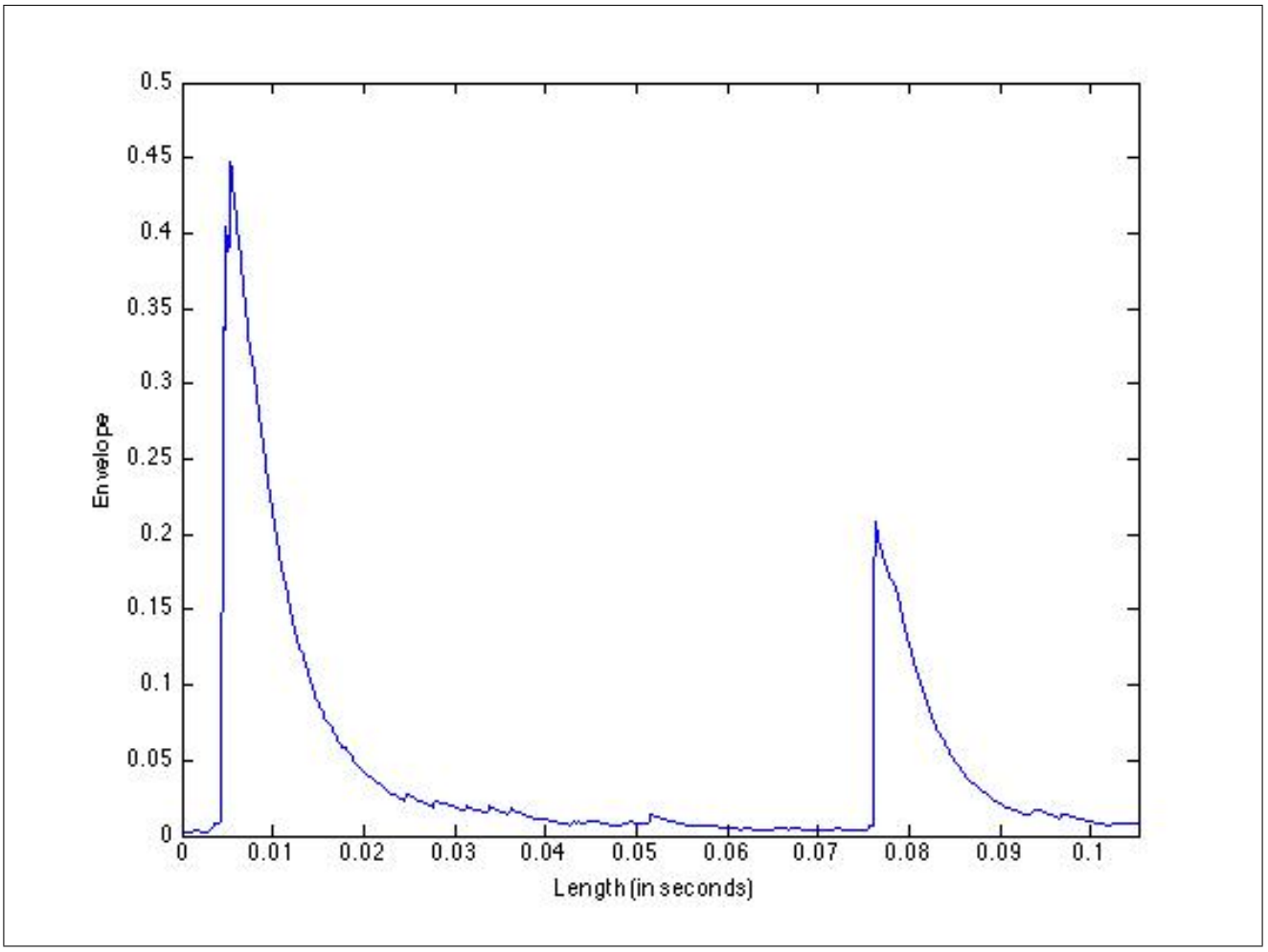


Figure 4.8: Envelope (GRF) extracted from the sound of Fig. 4.4.

In such an "envelope follower", the input signal is first rectified (absolute value) then the rectified signal is passed through a non-linear one-pole filter. If the rectified input is greater than the current output of the filter, a rapid "attack" tracking coefficient b_{up} is used. If the rectified input signal is less than the current filter output, a slower "release" coefficient b_{down} is used. The filter gain coefficient g is always set to $(1.0 - b)$, to keep the total dc gain of the filter equal to 1.0. Typical values for a 22,050 Hz sample rate clapping/walking file are $b_{up} = 0.8$ and $b_{down} = 0.995$. We used such values for our purposes. Fig. 4.8 shows the envelope extracted form the footstep of Figure 4.4.

4.3 Testing the systems

Both approaches with the contact microphones attached to the shoes showed the same problems. First of all, during the act of walking the sound of the air against the microphones was also captured. Such an unwanted sound constituted a not negligible input error for our system. Moreover, other extra noises were clearly captured, as the noise of the trousers on the shoes, the rubbing of the two sleeves of trousers on each others, or the noise of the wires against the trousers. Finally, the system turned out to be also not so wearable or comfortable as we had imagined. Indeed, it was not so simple to attach the microphones in a right way to the shoes, as well as the wires to the trousers in order for the wires not to cause any extra sound. In addition, setting up the microphones on the users could take a long time.
Conversely, the use of the microphones on the ground did not show any particular problem, and the GRF was extracted in a clear way.

References

[1] J. Benesty, J. Chen, and Y. Huang. *Microphone Array Signal Processing*. Springer, 2008.

[2] A. Crevoisier and P. Polotti. Tangible acoustic interfaces and their applications for the design of new musical instruments. In *Proc. of the Int. Conf. on New Interfaces for Musical Expression*, Vancouver, Canada, May 2005. NIME.

[3] C. Draeger and M. Fink. One-channel time reversal of elastic waves in a chaotic 2d-silicon cavity. *Physical Review Letters*, 79(3):407–410, 1997.

[4] A. Faivre, M. Dahan, B. Parratte, and G. Monnier. Instrumentated shoes for pathological gait assessment. *Mechanics Research Communications*, 31:627–632, 2004.

[5] Independent Nondestructive Testing and Evaluation journal ($NDT\&E$-International). www.elsevier.com/locate/ndteint.

[6] Knowles Acoustics. Bu series accelerometers. World Wide Web electronic publication, http://sigma.octopart.com/54401/datasheet/Knowles-Acoustics-BU-1771.pdf.

[7] L. Peltola, C. Erkut, P. Cook, and V. Välimäki. Synthesis of hand clapping sounds. *IEEE Trans. on Audio, Speech and Language Processing*, 15(3):1021–1029, 2007.

[8] N. Quieffin, S. Catheline, R. Kiri Ing, and M. Fink. Acoustic source localization model using in-skull reverberation and time reversal. *Applied Physics Letters*, 90(6), 2007.

[9] D. Rocchesso and F. Fontana, editors. *The Sounding Object*. Edizioni di Mondo Estremo, Florence, Italy, 2003.

[10] TAI-CHI. http://www.taichi.cf.ac.uk/. EU FP6-IST-2002 project, 2003-2005.

[11] I. A. Viktorov. *Rayleigh and Lamb Waves*. Plenum Press, 1967.