

European Project Space on Intelligent Systems, Pattern Recognition and Biomedical Systems

Lisbon, Portugal
January, 2015

Ana Fred (Ed.)

Sponsored and Organized by INSTICC
Published by SCITEPRESS

Copyright © 2015 by SCITEPRESS – Science and Technology
Publications. Lda.
All rights reserved

Edited by Ana Fred

Printed in Portugal
ISBN: 978-989-758-095-6
Depósito Legal: 394072/15

Foreword

This book contains the revised and extended versions of papers describing a number of European projects that were presented at the two European Project Space (EPS) events organized in Lisbon, January 2015, associated with the set of conferences namely BIOSTEC (International Joint Conference on Biomedical Engineering Systems and Technologies), ICAART (International Conference on Agents and Artificial Intelligence); ICPRAM (International Conference on Pattern Recognition Applications and Methods) and ICORES (International Conference on Operations Research and Enterprise Systems).

All these events were sponsored by the Institute for Systems and Technologies of Information, Control and Communication (INSTICC) in collaboration with several international associations and other scientific partners.

The objective of the EPS is to provide a platform and meeting point for researchers and practitioners to discuss and exchange experiences on European Research and Development projects which are funded and developed in the European research area or in collaboration with European partners, thus providing insights into cutting edge research work originating from Academia or Industry, either in Europe or elsewhere.

We aim to establish the EPS as a regular event to create opportunities for knowledge and technology sharing, and establish the basis for future collaboration networks involving current project partners and interested conference delegates.

These two events included a panel discussion with representatives and experts from the European Community.

The panel entitled “The Road Ahead: New Challenges for Learning and Intelligent Systems” held on January, 11th had the presences of Marcello Pelillo, University of Venice, Italy (chair of the panel); Eunice Ribeiro, Horizon 2020 National Contact Point (NCP) – Energy, FET &

IV

ICT, Portugal, Mario Figueiredo, Technical University of Lisbon - IST, Portugal, Nello Cristianini, University of Bristol, United Kingdom, Carmen Gervet, Université de Savoie, France and Oswaldo Ludwig, KU Leuven, Belgium.

On the January 13th the panel entitled “European Project Space on Biomedical Engineering” included the presences of Ana Fred from the Instituto de Telecomunicações / IST, Portugal as chair of the panel and the panel participants: Alexandra Veiga, Horizon 2020 National Contact Point (NCP) –ERC, Health, Fusion, IMI, Portugal, Marta Bienkiewicz, TUM, Technische Universität München, Germany, Justyna Maculewicz, Aalborg University Copenhagen, Denmark and Mohamed Oussama Ben Salem, Tunisia Polytechnic School, Tunisia.

The EPS technical program included, in addition to an opening panel, the presentation of five projects which, after the event, have been invited to publish a short report in this EPS book.

We would like to thank the project representatives that decided to take their time and effort to respond to our invitation, whose reports correspond to the five chapters of this book.

Ana Fred

Editor

Organization

Panel Chairs

Ana Fred, Instituto de Telecomunicações / IST, Portugal
Marcello Pelillo, University of Venice, Italy

Panel Participants

Marta Bienkiewicz, TUM, Technische Universitat Munchen, Germany
Nello Cristianini, University of Bristol, United Kingdom
Mario Figueiredo, Technical University of Lisbon - IST, Portugal
Carmen Gervet, Université de Savoie, France
Oswaldo Ludwig, KU Leuven, Belgium
Justyna Maculewicz, Aalborg University Copenhagen, Denmark
Eunice Ribeiro, Horizon 2020 National Contact Point (NCP) – Energy,
FET & ICT, Portugal
Mohamed Oussama Ben Salem, Tunisia Polytechnic School, Tunisia
Alexandra Veiga, Horizon 2020 National Contact Point (NCP) –ERC,
Health, Fusion, IMI, Portugal

Presented Projects

Acronym: Cogwatch

Presenter: Marta Bienkiewicz, TUM, Technische Universitat Munchen, Germany

Acronym: NIW

Presenter: Justyna Maculewicz, Aalborg University Copenhagen, Denmark

Name: New Technologic Solutions for a Browser and Robot-based Orthopedic Surgery

Presenter: Mohamed Oussama Ben Salem, Tunisia Polytechnic School, Tunisia

Acronym: MUSE

Presenter: Thomas Provoost, KU Leuven, Belgium

Acronym: MINERVA

Presenter: Valery Naranjo, Universitat Politècnica de València, Spain

Name: Augmented Human Assistance

Presenter: Artur Arsenio, IST-ID / UBI / YDreams Robotics, Portugal

Acronym: El-MUNDO

Presenter: Carmen Gervet, Université de Savoie, France

Name: Self-organised Trusted Communities

Presenter: Jan Kantert, Leibniz University Hannover, Germany

Acronym: FI-CORE

Presenter: Hanno Hildmann, NEC Research Labs Europe, Germany

Acronym: SCAENERGY

Presenter: Madalina Drugan, Vrije Universiteit Brussel, Belgium

Acronym: MAVEN

Presenter: Luca Piras, University of Cagliari, Italy

Table of Contents

Papers

The Natural Interactive Walking Project and Emergence of Its Results in Research on Rhythmic Walking Interaction and the Role of Footsteps in Affecting Body Ownership.....	3
<i>Justyna Maculewicz, Erik Sikström and Stefania Serafin</i>	
New Robotic Platform for a Safer and More Optimal Treatment of the Supracondylar Humerus Fracture.....	25
<i>Mohamed Oussama Ben Salem, Zied Jlalia, Olfa Mosbahi, Mohamed Khalgui and Mahmoud Smida</i>	
Robotics Enabled Augmented Health.....	33
<i>Artur Arsenio</i>	
MINERVA Project, mid- To near Infrared Spectroscopy for Improved Medical Diagnostics.....	53
<i>Valery Naranjo, Francisco Peñaranda, Mariano Alcañiz, Bruce Napier, Mark Farries, Gary Stevens, John Ward, Cestmir Barta, Radek Hasal, Angela Seddon, Sławomir Sujecki, Samir Lamrini, Uffe Møller, Ole Bang, Peter M. Moselund, Munir Abdalla, Danny De Gaspari, Rosa M. Vinella, Hedda Malm, Gavin R. Lloyd, Nick Stone, Jayakrupakar Nallala, Juergen Schnekenburger, Lena Kastl and Björn Kemper</i>	
El MUNDO: Embedding Measurement Uncertainty in Decision Making and Optimization.....	70
<i>Carmen Gervet and Sylvie Galichet</i>	
Author Index.....	91

Papers

The Natural Interactive Walking Project and Emergence of Its Results in Research on Rhythmic Walking Interaction and the Role of Footsteps in Affecting Body Ownership

Justyna Maculewicz, Erik Sikström and Stefania Serafin

Architecture, Design and Media Technology Department, Aalborg University Copenhagen,
2450, Copenhagen, Denmark
{jma, es, sts}@create.aau.dk
<http://media.aau.dk/smc>

Abstract. In this chapter we describe how the results of the Natural Interactive Project, which was funded within the 7th Framework Programme and ended in 2011, started several research directions concerning the role of auditory and haptic feedback in footstep simulations. We chose elements of the project which are interesting in a broader context of interactive walking with audio and haptic feedback to present and discuss the developed systems for gait analysis and feedback presentation, but also, what is even more interesting to show how it influence humans behavior and perception. We hope also to open a discussion on why we actually can manipulate our behavior and show the importance of explaining it from the neurological perspective. We start with a general introduction, moving on to more specific parts of the project, that are followed by the results of the research which were conducted after project's termination but based on its results.

1 Introduction

Walking is an activity that plays an important part in our daily lives. In addition to being a natural means of transportation, walking is also characterized by the resulting sound, which can provide rich information about the surrounding and a walker. The study of the human perception of locomotion sounds has addressed several properties of the walking sound source. The sound of footsteps conveys information about walker's gender [1, 2], posture [3], emotions [2], the hardness and size of their shoe sole[2], and the ground material on which they are stepping [4]. It was proven that sounds of footsteps convey both temporal and spatial information about locomotion [5].

Auditory feedback has also strength to change our behavior. Studies show that interactive auditory feedback produced by walkers affects walking pace. In the studies of [6, 7] individuals were provided with footstep sounds simulating different surface materials, interactively generated using a sound synthesis engine [8]. Results show that subjects' walking speed changed as a function of the simulated ground material.

From the clinical perspective, sensory feedback and cueing in walking received an increased attention. It is well known that sensory feedback have a positive effect on gait in patients with the neurological disorders, among which is also Parkinson's disease

(PD) [9–14]. Rhythmic (metronome-like) auditory cues have been found to produce gait improvement in several studies [9–15]. External rhythms presented by auditory cues may improve gait characteristics [13–15], but also be used to identify deficits in gait adaptability [16].

Research on sensory feedback while walking is also important in the areas of virtual augmented realities.

The addition of auditory cues and their importance in enhancing the sense of immersion and presence is a recognized fact in virtual environment research and development. Studies on auditory feedback in VR are focused on sound delivery methods [17, 18], sound quantity and quality of auditory versus visual information [19], 3D sound [20, 21] and enhancement of self-motion and presence in virtual environments [22–24].

Within the study of human perception of walking sounds researchers have focused on topics such as gender identification [1], posture recognition [3], emotional experiences of different types of shoe sound (based on the material of the sole) on different floor types (carpet and ceramic tiles) [25], and walking pace depending on various types of synthesized audio feedback of steps on various ground textures [26].

2 The Objectives of NIW Project

The NIW project contributed to scientific knowledge in two key areas. First it reinforced the understanding of how our feet interact with surfaces on which we walk. Second, informed the design of such interactions, by forging links with recent advances in the haptics of direct manipulation and in locomotion in real-world environments. The created methods have potential to impact a wide range of future applications that have been prominent in recently funded research within Europe and North America. Examples include floor-based navigational aids for airports or railway stations, guidance systems for the visually impaired, augmented reality training systems for search and rescue, interactive entertainment, and physical rehabilitation.

The NIW project proceeded from the hypothesis that walking, by enabling rich interactions with floor surfaces, consistently conveys enactive information that manifests itself predominantly through haptic and auditory cues. Vision was regarded as playing an integrative role linking locomotion to obstacle avoidance, navigation, balance, and the understanding of details occurring at ground level. The ecological information was obtained from interaction with ground surfaces allows us to navigate and orient during everyday tasks in unfamiliar environments, by means of the invariant ecological meaning that we have learned through prior experience with walking tasks.

At the moment of the project execution, research indicated that the human haptic and auditory sensory channels are particularly sensitive to material properties explored during walking [4], and earlier studies have demonstrated strong links between the physical attributes of the relevant sounding objects and the auditory percepts they generate [27]. The project intention was to select, among these attributes, those which evoke most salient perceptual cues in subjects.

Physically based sound synthesis models are capable of representing sustained and transient interactions between objects of different forms and material types, and such methods were used in the NIW project in order to model and synthesize the sonic effects

of basic interactions between feet and ground materials, including impacts, friction, or the rolling of loose materials.

The two objectives which guided the project are:

1. The production of a set of foot-floor multimodal interaction methods, for the virtual rendering of ground attributes, whose perceptual saliency has been validated
2. The synthesis of an immersive floor installation displaying a scenario of ground attributes and floor events on which to perform walking tasks, designed in an effort to become of interest in areas such as rehabilitation and entertainment.

The results of the research within the NIW project can be enclosed in the three milestones:

- Design, engineering, and prototyping of floor interaction technologies
- A validated set of ecological foot-based interaction methods, paradigms and prototypes, and designs for interactive scenarios using these paradigms
- Integration and usability testing of floor interaction technologies in immersive scenarios.

The forthcoming sections will focus on research initiated at Aalborg University and continued after the termination of the project.

3 Synthesis of Footsteps Sounds

3.1 Microphone-based Model

A footprint 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).

Our goal in developing the footsteps sounds synthesis engine was to synthesize a footprint sound on different kinds of materials starting from a signal in the audio domain containing a generic footprint sound on a whatever material. Our approach to achieve this goal consisted of removing the contribution of the resonator, keeping the exciter and considering 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 were summed in order to have a complete footprint sound.

In order to simulate the footsteps sounds on different types of materials, the ground reaction force estimated with this technique was used to control various sound synthesis algorithms based on physical models, simulating both solid and aggregate surfaces [28, 29]. The proposed footsteps synthesizer was implemented in the Max/MSP sound synthesis and multimedia real-time platform.¹

Below we present an introduction to developed physically based sound synthesis engine that is able to simulate the sounds of walking on different surfaces. This introduction is an excision from [8]. Figure 2 presents the setup that was used for testing designed models of feedback delivery. We developed a physically based sound synthesis engine that is able to simulate the sounds of walking on different surfaces. Acoustic

¹www.cycling74.com



Fig. 1. A screenshot of the graphical user interface for the developed sound synthesis engine. [8].

and vibrational signatures of locomotion are the result of more elementary physical interactions, including impacts, friction, or fracture events, between objects with certain material properties (hardness, density, etc.) and shapes. The decomposition of complex everyday sound phenomena in terms of more elementary ones has been a recurring idea in auditory display research during recent decades [30]. In our simulations, we draw a primary distinction between solid and aggregate ground surfaces, the latter being assumed to possess a granular structure, such as that of gravel, snow, or sand. A comprehensive collection of footstep sounds was implemented. Metal and wood were implemented as solid surfaces. In these materials, the impact model was used to simulate the act of walking, while the friction model was used to simulate the sound of creaking wood. Gravel, sand, snow, forest underbrush, dry leaves, pebbles, and high grass are the materials which were implemented as aggregate sounds. The simulated metal, wood, and creaking wood surfaces were further enhanced by using some reverberation. Reverberation was implemented by convolving in real-time the footstep sounds with the impulse response recorded in different indoor environments. The sound synthesis algorithms were implemented in C++ as external libraries for the Max/MSP sound synthesis and multimedia real-time platform. A screenshot of the final graphical user interface can be seen in Figure 1. In our simulations, designers have access to a sonic palette making it possible to manipulate all such parameters, including material properties. One of the challenges in implementing the sounds of different surfaces was to find suitable combinations of parameters which provided a realistic simulation. In the synthesis of aggregate materials, parameters such as intensity, arrival times, and impact form a powerful set of independent parametric controls capable of rendering both the process dynamics, which is related to the temporal granularity of the interaction (and linked to the size of the foot, the walking speed, and the walker's weight), and the type of material the aggregate surface is made of. These controls enable the sound designer to choose foot-ground contact sounds from a particularly rich physically informed palette. For each simulated surface, recorded sounds were analyzed according to their combinations of events, and each subevent was simulated independently. As an example, the sound produced while walking on dry leaves is a combination of granular sounds with long duration both at low and high frequencies, and noticeable random sounds with not very high density that give to the whole sound a crunchy aspect. These different components were simulated with several aggregate models having the same density, duration, frequency, and number of colliding objects. The amplitude of the different components

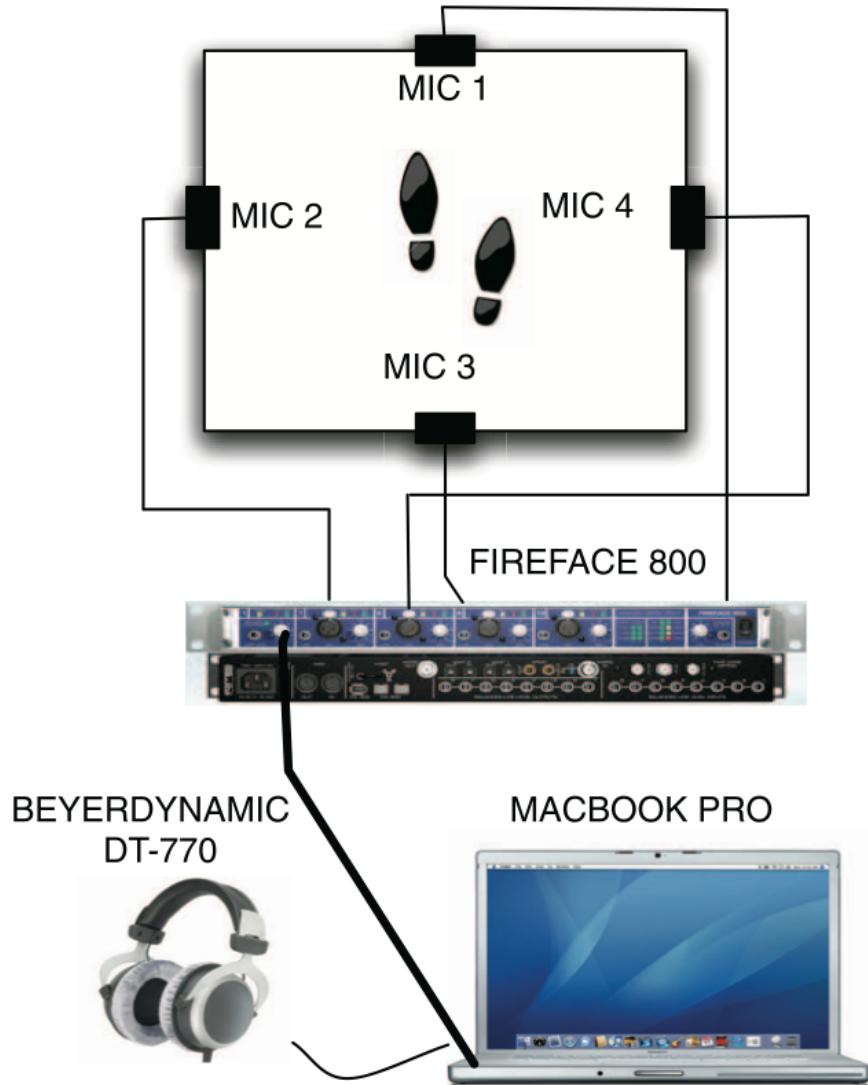


Fig. 2. Hardware components of the developed system: microphones, multichannel soundcard, laptop, and headphones [8].

were also weighted, according to the same contribution present in the corresponding real sounds. Finally, a scaling factor was applied to the volumes of the different components. This was done in order to recreate a sound level similar to the one happening during a real footprint on each particular material.

3.2 The MoCap-based System

In this section we address the problem of calculating the GRF from the data tracked by means of a Motion Capture System (MoCap) in order to provide a real-time control of the footsteps synthesizer. Our goal was to develop a system which could satisfy the requirements of shoe independence, fidelity in the accuracy of the feet movements, and free navigation.

Figure 3 shows a schematic representation of the overall architecture developed. This system is placed in an acoustically isolated laboratory and consists of a MoCap², two soundcards³, sixteen loudspeakers⁴, and two computers. The first computer runs the motion capture software⁵, while the second runs the audio synthesis engine. The two computers are connected through an ethernet cable and communicate by means of the UDP protocol. The data relative to the MoCap are sent from the first to the second computer which processes them in order to control the sound engine.

The MoCap is composed by 16 infrared cameras⁶ which are placed in a configuration optimized for the tracking of the feet. In order to achieve this goal, two sets of markers are placed on each shoe worn by the subjects, in correspondence to the heel and to the toe respectively.

Concerning the auditory feedback, the sounds are delivered through a set of sixteen loudspeakers or through headphones.

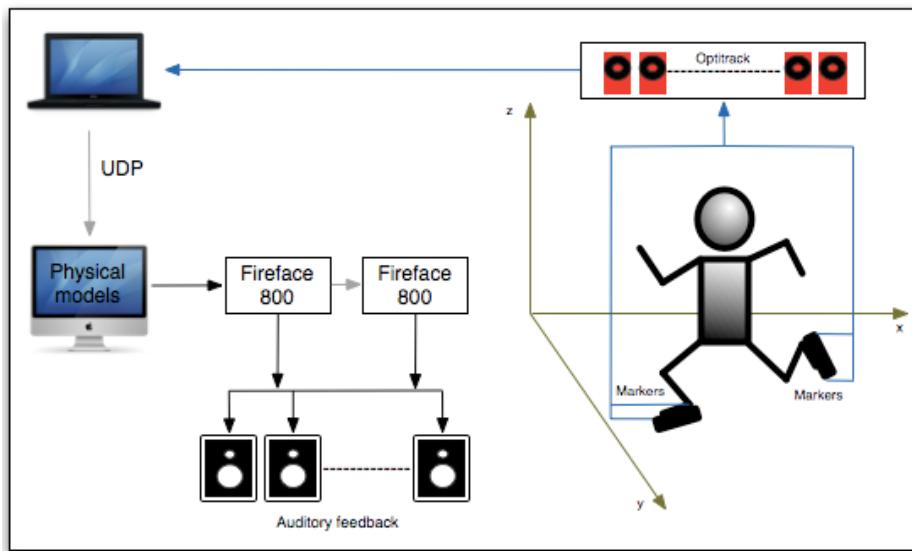


Fig. 3. A block diagram of the developed system and the used reference coordinates system.

²Optitrack: <http://naturalpoint.com/optitrack/>

³FireFace 800 soundcard: <http://www.rme-audio.com>

⁴Dynaudio BM5A: <http://www.dynaudioacoustics.com>

⁵Tracking Tools 2.0

⁶OptiTrack FLEX:V100R2

3.3 Wireless Shoes

In this section we address the problem of calculating the GRF from the data tracked by sensors placed directly on the walker shoes, with the goal of providing a real-time control of the footsteps synthesizer using in addition a wireless transmission.

The setup for the developed shoe-integrated sensors system is illustrated in Figure 4. Such system is composed by a laptop, a wireless data acquisition system (DAQ), and a pair of sandals each of which is equipped with two force sensing resistors⁷ and two 3-axes accelerometers⁸. More in detail the two FSR sensors were placed under the insole in correspondence to the heel and toe respectively. Their aim was to detect the pressure force of the feet during the locomotion of the walker. The two accelerometers instead were fixed inside the shoes. Two cavities were made in the thickness of the sole to accommodate them in correspondence to the heel and toe respectively. In order to better fix the accelerometers to the shoes the two cavities containing them were filled with glue.

The analog sensor values were transmitted to the laptop by means of a portable and wearable DAQ.

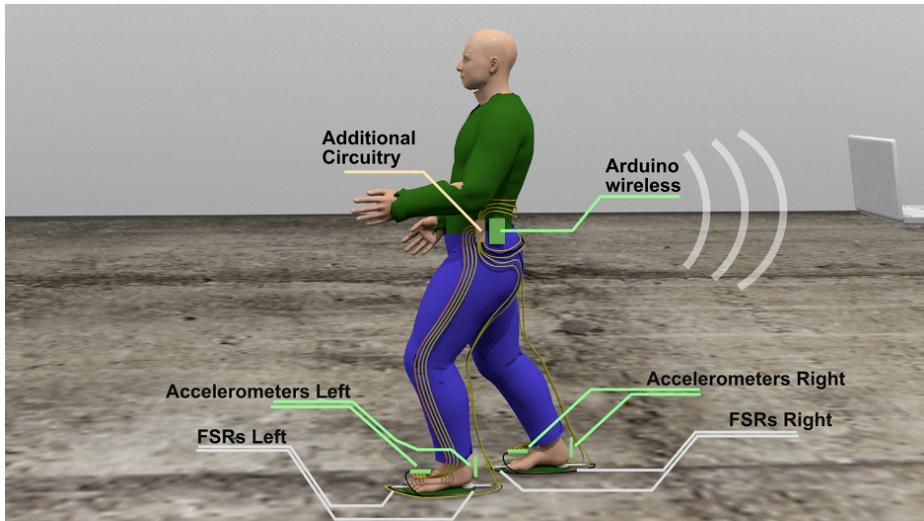


Fig. 4. Setup for the wireless shoes system: the user wears the sensor enhanced shoes and the wireless data acquisition system.

The wireless DAQ consists of three boards: an Arduino MEGA 2560 board⁹, a custom analog preamplification board, and a Watterott RedFly¹⁰ wireless shield. In the nomenclature of the Arduino community, a “shield” is a printed circuit board (PCB) that

⁷FSR: I.E.E. SS-U-N-S-00039

⁸ADXL325: <http://www.analog.com>

⁹<http://arduino.cc>

¹⁰<http://www.watterott.net/projects/redfly-shield>

matches the layout of the I/O pins on a given Arduino board, allowing that a shield can be “stacked onto” the Arduino board, with stable mechanical and electrical connections.

All three boards are stacked together. In this way the wireless DAQ system can be easily put together in a single box, to which a battery can be attached. This results in a standalone, portable device that can be attached to the user’s clothes, allowing greater freedom of movement for the user.

Since each foot carries two FSRs and two 3-axis accelerometers, which together provide 8 analog channels of data, the system demands capability to process 16 analog inputs in total. That is precisely the number of analog inputs offered by an Arduino MEGA 2560, whose role is to sample the input channels, and format and pass the data to the wireless RedFly shield. The analog preamplification board is a collection of four quad rail-to-rail operational amplifier chips (STmicroelectronics TS924), providing 16 voltage followers for input buffering of the 16 analog signals, as well as of four trimmers, to complete the voltage divider of the FSR sensors; and connectors. The Watterott RedFly shield is based on a Redpine Signals RS9110-N-11-22 WLAN interface chipset, and communicates with the Arduino through serial (UART) at 230400 baud. Preliminary measurements show that the entire wireless DAQ stack consumes about 200 mA with a 9V power supply, therefore we chose a power supply of 9V as the battery format.

3.4 Discussion

At software level, both the proposed systems are based on the triggering of the GRFs, and this solution was adopted rather than creating the signal in real-time. Indeed one could think to generate a signal approximating the shapes of the GRFs.

For instance the contributions of the heel strike, (as well as the one of the toe), could be approximated by means of a signal created by a rapid exponential function with a certain maximum peak followed by a negative exponential function with a certain decay time. Nevertheless it was not possible to create such signal in real time for two reasons. First the synthesis engine needs to be controlled by a signal in the auditory domain having a sample rate of 44100 Hz, but the data coming from both the systems arrive with a frequency much lower, 1000 Hz. Therefore mapping the data coming from the tracking devices to an auditory signal will result in a step function not usable for the purposes of creating a proper GRF to control the synthesis engine. Secondly the computational load to perform this operation added to the one of the algorithms for the sound synthesis would be too high, with the consequent decrease of the system performances mostly in terms of latency.

Hereinafter we discuss the advantages and disadvantages of the two developed systems in terms of portability, easiness of setup, wearability, navigation, sensing capabilities, sound quality, and integration in VR environments.

Portability. The MoCap based system is not portable as it requires to carry all the components of the architecture discussed in section 3.2. Conversely the wireless shoe system is easily portable.

Easiness of Setup. At hardware level, while the MoCap based system is not easy to setup since it consists of many components, instead the wireless shoe system does not show any difficulty in the set up process. Both the systems at software level require

an initial phase in which global parameters and thresholds of the proposed techniques have to be calibrated, but such calibration is however simple and quick.

Wearability. The MoCap based system allows users to wear their own footwear, and in addition no wires linked to the user are involved. The only technology required to be worn by the users consists of the four sets of markers which have to be attached to the shoes by means scotch tape. However, they are very light and therefore their presence is not noticeable, and in addition they do not constitute an obstacle for the user walk. Conversely the wireless shoe system is not shoe-independent since the users are required to wear the developed sandals. In addition walkers need to carry the box containing the Arduino board which is attached at the trousers, but its presence is not noticeable.

Navigation. In both the systems the user is free to navigate as no wires are involved. However the walking area in the case of the MoCap based system is delimited by the coverage angle of the infrared cameras, while the wireless shoes can be used in a wider area, also outdoor.

Sensing Capabilities. As regards the MoCap based system the functioning of the proposed technique is strictly dependent on the quality of the MoCap system utilized. The requirements for the optimal real-time work of the proposed method are a low latency and a good level of accuracy. The latency problem is the most relevant since the delivery of the sound to the user must be perfectly synchronized with the movements of his/her feet in order to result into a credible closed-loop interaction. For a realistic rendering the latency should be at maximum 15 milliseconds, therefore the current latency of 40 milliseconds is too much for the practical use of the proposed method. However this limit could be lowered by improving the MoCap technology since the latency is due in most part to it. Concerning the accuracy, a high precision MoCap system allows a better tracking of the user gestures which have to be mapped to GRFs for the subsequent generation of the sounds. In general, the overall computational load of this technique is high but if it is divided into two computers, one for the data acquisition and the other for the sound generation, it results acceptable.

As concerns the wireless shoes the total latency is acceptable for the rendering of a realistic interaction and the use of the accelerometers allows to achieve a mapping between the feet movements and the dynamics in footstep sounds similar to the one obtainable by using the microphones system described in [31].

Sound Quality. The sound quality of the system depends on the quality of the sound synthesis algorithms, on the sensing capabilities of the tracking devices, as well as on the audio delivery methods used. As concerns the quality of the synthesized sounds, good results in recognition tasks have been obtained in our previous studies [32, 33]. In addition, an highly accurate MoCap system, as well as good FSRs and accelerometers, allow to detect the footsteps dynamics with high precision therefore enhancing the degree of realism of the interaction. Concerning the audio delivery method, both headphones and loudspeakers can be used.

Integration in VR Environments. Both the system have been developed at software level as extension to the Max/MSP platform, which can be easily combined with several interfaces and different software packages. Both the system allow the simultaneous coexistence of interactively generated footsteps sounds and of soundscapes pro-

vided by means of the surround sound system.

The architecture of the MoCap based system can be integrated with visual feedback using for example a head mounted display to simulate different multimodal environments. However using the MoCap based system it is not possible to provide the haptic feedback by means of the haptic shoes developed in previous research [34] since such shoes are not involved and their use will result in a non wireless and not shoe-independent system.

It is possible to extend the wireless shoes system embedding some actuators in the sandals in order to provide the haptic feedback. For this purpose another wireless device receiving the haptic signals to must be also involved. Nevertheless, the latency for the round-trip wireless communication would be much higher.

4 The AAU Research after the NIW Project

4.1 Multi Sensory Research on Rhythmic Walking

Due to the successful implementation of physical models into design of ecological feedback and interest in rhythmic walking interaction with auditory and haptic feedback, we continued exploring these areas of research. We are specifically interested in the influence of auditory and haptic ecological feedback and cues on rhythmic walking stability, perceived naturalness of feedback and synchronization ease with presented cues. From our hitherto research emerged several effects, which are interesting in a context of gait rehabilitation, exercise and entertainment. Until now we have been testing gravel and wood sound as ecological feedback and a tone as a non-ecological sound. Results show that when we ask people to walk in their preferred pace, they have the slowest pace with gravel feedback, then wood and a tone motivates to the fastest walking [35]. To test this effect even further we added soundscape sounds which are congruent and incongruent with the sounds of the footsteps. The preliminary analysis shows that feedback sounds can manipulate participants pace even more than footsteps sounds alone [36]. When people are asked to synchronize with above-mentioned rhythmic sounds their results are similar with a slight worse performance with gravel cues [7]. Even though this feedback produces the highest synchronizing error it is perceived as the one, which is the easiest to follow [35]. In the same study [35] we also investigated the influence of feedback in the haptic modality, in rhythmic walking stimulation. We have seen that haptic stimulation is not efficient in rhythmic cueing, but it might help to improve the naturalness of the walking experience (Fig. 5, 6). In order to understand these results, we turn to neurological data in search of an explanation. The results of our preliminary exploratory encephalographic (EEG) experiment suggest that synchronization with non-ecological sounds requires more attention and in synchronization with ecological sounds is involved a social component of synchronizing with another person. Synchronizing with the pace, which is similar to the natural walking, also requires less attention. The analysis of the EEG data is ongoing [37].

Many different ways of feedback delivery to the user were presented. We can see that different types of auditory feedback can be crucial in all the aspects mentioned before. Many behavioral effects were observed while presenting feedback through mentioned applications. We believe that there is a need now to understand why feedback or

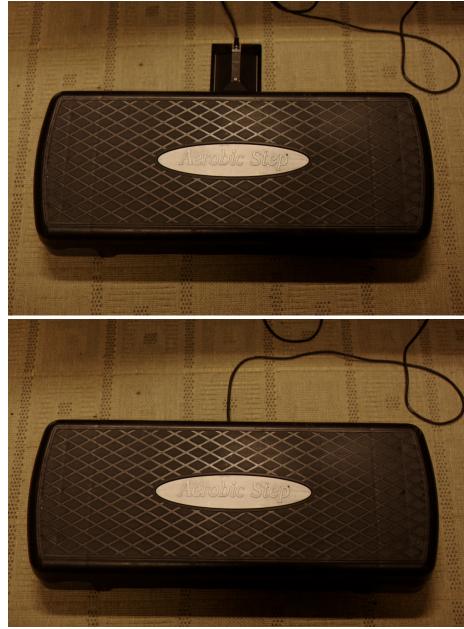


Fig. 5. Experimental setup used in the studies following the NIW project. A microphone was placed below a stepper to detect person's steps. An actuator was located under the top layer of a stepper [7].

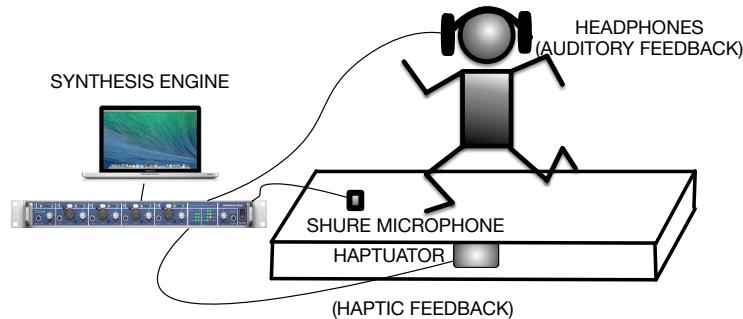


Fig. 6. A visualization of the setup used in the studies following the NIW project. It clarifies the placement of the actuator. We can see that feedback was presented through headphones connected to synthesis engine describe in the previous section via Fireface 800 sound card [35].

cues can manipulate our behavior. Deeper understanding is needed to gain basic knowledge about the neural bases of altered behavior, which will help to build more efficient and precise feedback systems and also to design feedback signal in a way they could be the most efficient for specific need.

4.2 Footstep Sounds and Virtual Body Awareness

An experiment was designed and conducted with the aim of investigating if and how the user's perception of the weight of a first person avatar could be manipulated by the means of altering the auditory feedback. Inspired by the approach used by Li et al. [1], using very basic audio filter configurations, a similar methodology was adopted. Instead of using one floor type, two floors with different acoustic properties were used in the experiment.

Experiment. In order to investigate whether it would be possible to manipulate a user's perception of how heavy their avatar could be by manipulating the audio feedback generated from interactive walking in place input, an experiment was set up involving an immersive virtual reality setup with full body motion tracking and real-time audio feedback. The environment featured a neutral starting area with a green floor and two corridors equipped with different floor materials from which the walking sounds would be evaluated; concrete tiles or wooden planks. The area with the green floor were used for training and transitions. In the experiment, subjects were asked to perform six walks (explained further under the experiment procedure sub-section) while orally giving estimates of the perceived weight of the virtual avatar, as well as rating the perceived suitability of the sounds as sound effects for the ground types that they just walked over. For each of the walks, the audio feedback from the interactive footsteps was manipulated with audio filters that came in three different configurations (A, B and C, explained further in the audio implementation section). Thus, the six walks were divided into three walks on concrete featuring filters A, B and C (one at a time), while three walks were on the wooden floor, also featuring filters A, B and C. As a part of the experiment the virtual avatar body was also available in two sizes (see Avatar and movement controls). The hypothesis was formulated so that the filters would bias the subjects into making different weight estimate depending on the filter type. It was also expected that the filters that had lower center frequencies would be estimated as being heavier.

Implementation. The virtual environment was implemented in Unity 3D¹¹ while audio feedback was implemented in Pure Data¹². Network communication between the two platforms were managed using UDP protocol.

Avatar and Movement Controls. The avatar used for the experiment consisted of a full male body. The body was animated using inverse kinematics animation and data acquired from a motion capture system. The avatar also had a walking in place ability that allowed the subjects to generate forward translation from the user performing stepping in place movements. The model of the avatar body was also available in two versions where one (a copy of the original model) had been modified to have an upper body with a larger body mass with thicker arms and a torso with a bigger belly and chest (see Fig. 7). A calibration procedure also allowed the avatar to be scaled to fit a subject's own height and length of limbs.

¹¹<http://www.unity3d.com>

¹²<http://puredata.info>

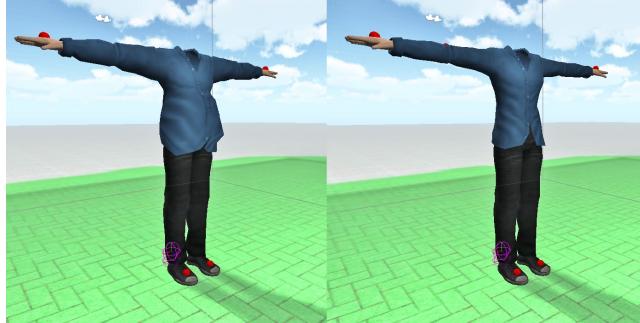


Fig. 7. The big and the small avatar bodies.

Audio Implementation. A skilled sound designer would probably utilize a combination of DSP effects and editing techniques, when asked to make a certain footstep sequence sound heavier or lighter. We have for the sake of this experiment chosen to resort to only using one method for changing the sounds for the sake of experimental control. The audio feedback was made up out of two sets of Foley recordings of footsteps, one for the concrete surface sounds and one for the wooden surface sounds. The recordings were made in an anechoic chamber using a Neumann u87ai microphone¹³ at a close distance and an RME Fireface 800 audio interface¹⁴. Each set contained seven variations of steps on either a concrete tile or wooden planks (a wooden pallet). The playback was triggered by the walking in place script in the Unity implementation and cues were sent via UDP to the Pure Data patch that would play the sounds. The footstep sounds were played back at a random order, including the avoiding of consecutive repetitions of individual samples. Three different filters were applied to the recordings using the Equalization plug-in in Audacity¹⁵ with the graphic EQ setting with the B-spline interpolation method selected. These filters were rendered into copies of the concrete and wood sets of the Foley recordings, generating a total of 6x7 audio files. The filters were set up in peak-shapes (similarly to Li et al. [1]), amplifying center frequencies of either 80Hz (hereafter labelled “filter”), 160Hz (labelled “filter B”), or 315Hz (labelled “filter C”) with +20 dB, as shown on the slider of the interface. The two adjacent sliders were set to be amplifying their respective frequency areas with +10 dB (see Fig. 8). According to the frequency response curve presented in the plug-in, the actual amplification of the center frequency was lower than +20 dB and a bit more spread out (for an overview of all three filters, see Fig. 9). Finally all of the audio files were normalized to the same output level. The footstep sounds were then presented to the subjects through headphones (see section Equipment and Facilities), at approximately 65 dB (measured from the headphone using an AZ instruments AZ8922 digital sound level meter). Which files would be played was determined by a ground surface detection script in Unity that identifies the texture that the avatar is positioned above. The filters would be selected manually using keyboard input.

¹³<https://www.neumann.com>

¹⁴<http://www.rme-audio.de>

¹⁵<http://audacity.sourceforge.net/>

For the green areas in the virtual environment, a neutral type of footstep sound was used consisting out of a short “blip” sound, a sinus tone following an exponential attack and decay envelope (the ead~ object in Pure Data) with a 5 ms attack and 40 ms decay.

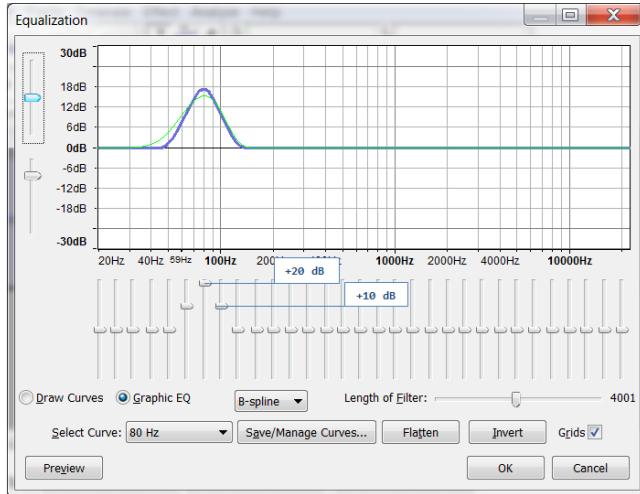


Fig. 8. Filter settings for filter A.



Fig. 9. Overview of the filters.

Equipment and Facilities. The virtual environment, audio and motion tracking software were running on one Windows 7 PC computer (Intel i7-4470K 3.5GHz CPU, 16 GB RAM and an Nvidia GTX 780 graphics card). The head mounted display was a nVisor SX with a FOV of 60 degrees with a screen resolution 1280x1024 pixel in each eye. The audio were delivered through an RME Fireface 800 with a pair of Sennheiser HD570 headphones. The motion-tracking was done with a Naturalpoint optitrack motion-tracking system with 11 cameras of the model V100:R2 and with 10 3-point trackables attached on the subjects feet, knees, hip, elbows hands and on the head mounted display.

Discussion. Changing the size of the avatar did not make the subjects provide significantly different evaluations in the weight estimates or suitability ratings. This could

possibly have to do with the subject did not pay much attention to the appearance of their virtual bodies, even though they were given an opportunity to see it during the calibration process and the training session. It may also have to do with technical limitations of the implementation and the hardware such as the head mounted display's rather narrow field of view.

In between the weight estimates of the filters, the results indicate a small effect that at least partially supports the experimental hypothesis. The effect is not entirely consistent, in that filter A for the concrete surface was given lighter estimates than filter B in the big body group and in the combined groups analysis. Similarly for the wood surface context, the estimates for filter B were not significantly different from filter A or C in any variant of the analysis and in the big body group there were no significant differences among the estimates for any of the filters. Perhaps further investigations with a greater number of filter configurations, including a greater variety of filter characteristics, could give more detailed information regarding weight estimates in this kind of context.

The suitability ratings were among the filters only significantly different in one case, where in the combined groups filter A given a significantly higher rating than filter C. This could be interpreted that filter C had a negative impact of the audio quality of the footstep samples, as sound effects for the concrete tiles. This information could be useful and should be taken into account for those who plan to use automated Foley sounds in virtual reality applications. Interestingly, the suitability ratings for the wooden floor were very similar for all the three filter settings, suggesting a higher tolerance associated with this material for the type of manipulations used here, than the concrete surface material.

There are also a couple of issues of technical nature since the implementation is still at a prototype stage (and likely not representable of the capacity of some current or future commercial systems) that could have had an impact on the subjects' performances during the experiment. The triggering of the footstep sounds were suffering from a latency that was noticeable during fast walking and occasionally footstep sounds would be triggered more than once per step. There were also an issue in that while the subjects were walking in one of the directions, many times the footsteps would only trigger for one of the two feet. When this happened, the subjects were asked to turn around and walk in the opposite direction in order to get sounds from the steps from both feet and to get the time they needed with a representative and working feedback. These limitations could have made the virtual reality experience less believable and it would be interesting to see if a system of higher quality would yield a different experience in terms of weight estimates and suitability judgements.

Since the experiment used a walking in place-type of method for generating translation in the virtual environment and for triggering the audio feedback, the resulting interaction cannot be considered entirely similar to real walking. This may of course also have had an impact on the subjects' experiences, but it was necessary to employ this type of technique since the motion tracking area in the facilities was quite small.

The audio methods for manipulating the feedback could be more elaborate and advanced (as mentioned in the audio implementation section) and further investigations should consider involving other approaches such as pitch shifting and layering of sev-

eral sound effect components (also including creaking, cracking and crunching sounds that may belong to the ground surface types).

4.3 Footsteps Sounds and Presentation Formats

When evaluating a sound design for virtual environment, the context where it is to be implemented might have an influence on how it may be perceived. We performed an experiment comparing three presentation formats (audio only, video with audio and an interactive immersive VR format) and their influences on a sound design evaluation task concerning footprint sounds. The evaluation involved estimating the perceived weight of a virtual avatar seen from a first person perspective, as well as the suitability of the sound effect relative to the context.

We investigated the possible influence of the presentation format, here as in immersive VR and two less (technologically) immersive, seated and less interactive desktop formats, on the evaluations of a footprint sound effect with a specific purpose. This purpose is to provide auditory feedback for a walking in place interaction in VR and to describe the avatar's weight. Previously we have seen that the perceived weight of a VR avatar [38], as well as of the actual user [39], might be manipulated by changing the sonic feedback from the walking interactions.

Our aim is to support the process of sound design evaluations, by bringing understanding to what role the context of the presentation may have. As previous research provides hints [40–50], that presentation formats may potentially bring about a slightly different experience of the sound being evaluated. More information on this topic would be important for sound designers when developing audio content for VR simulations and for researchers doing experiments on audio feedback in these contexts.

Experiment. As immersive VR is potentially able to influence subjective judgements of audio quality, we here perform a study complementing a previous study conducted in a full immersive virtual reality (IVR) setting (full body avatar, gesture controlled locomotion and audio feedback) with two less immersive presentation formats. The formats were:

- An immersive VR condition (VR), using a head mounted display (HMD), motion tracking of feet, knees, hip, hands, arms and head with gesture controlled audio feedback triggered by walking in place input (which is also used for locomotion)
- An audiovisual condition (Video), produced utilizing a motion capture and screen recordings from the VR setup, presented in full screen mode on a laptop screen
- An audio only condition (Audio), using only the audio track from the above mentioned screen recording

The VR condition was also fully interactive, requiring the user to use their whole body to control the avatar and navigate inside the environment, while the Video and Audio conditions were passive with the subject sitting down at a desk. For each of the walks, the audio feedback from the interactive footsteps were manipulated with audio filters in three different peak shaped configurations with center frequencies at either 80 Hz (filter A), 160 Hz (filter B) or 315 Hz (filter C). The filters were all applied to both of

the two floor materials. Using 9 point Likert scales, the subjects were asked to estimate the weight of the avatar and the suitability of the footstep sounds as sound effects for the two ground materials. Both a concrete and a wooden surface were evaluated by each subjects, but in one presentation format only per subject.

Hypothesis:

- H₁: The degree of immersiveness and interaction in the presentation formats will have an influence on the participants ratings of the weight estimates
- H₂: The degree of immersiveness and interaction in the presentation formats will have an influence on the participants ratings of the suitability estimates
- H₃: The same patterns observed in the VR presentation format in regards to weight estimates will also be observed in the Audio and Video conditions

In the VR condition 26 subjects participated, while 19 subjects participated in each of the two other conditions. In the VR condition the test population had a mean age of 25 ($M = 25.36$, $SD = 5.2$) and of which 7 were female. The test population of the Audio and Video groups, consisted of University staff and students of which 10 were females, had an average age of 29 years ($M = 29.12$, $SD = 7.86$). One of these reported a slight hearing impairment, and one reported having tinnitus.

Implementation. The virtual environment was implemented using Unity3D¹⁶ while the audio feedback was implemented using Pure Data¹⁷. For the Video condition, a video player was implemented using vvvv¹⁸. Network communication between the different software platforms were managed using UDP protocol.

All software in the VR implementation were running in a Windows 7 PC computer (Intel i7-4479K 3.5GHz CPU, 16 GB ram, with an Nvidia GTX780 graphics card). In the Video and Audio conditions, all software were running on a Dell ProBook 6460b laptop (Intel i5-2410M 2.3 GHz, 4 GB ram, with Intel HD Graphics 3000), also with Windows 7.

The audio feedback implementation used for the VR condition in this experiment was identical to the one presented in the previously described study (see 4.2 Footstep Sounds and Virtual Body Awareness - implementation).

The video was captured using FRAPS with 60 FPS frame rate (although the frame rate of the IVR implementation was not the same as this) and the same screen resolution as used in the head mounted display (1280*1024 pixels per eye) but with one camera (the right) deactivated (since stereoscopic presentation was not used in this experiment), resulting in a resolution 1280*1024 pixels. The audio was captured at stereo 44.1 kHz, 1411 kbps. In order to make the video files smaller (the original size was over 6 Gb per file) further compression was applied using VLC¹⁹, but same resolution with a data rate of 6000 kbps at 30 frames/second and with the audio quality remaining the same as in the original video capture. The interface used for the test was developed using Pure

¹⁶www.unity3d.com

¹⁷puredata.info

¹⁸vvvv.org

¹⁹www.videolan.org/vlc/

Data and vvvv. The Pure Data patch contained the graphical user interface presented on an external monitor (Lenovo L200pwD), the data logging functions and networking components for communicating with the vvvv patch. The vvvv patch held the functions necessary for streaming the video files from disk with audio and presenting on a 14 inch laptop monitor (Dell Probook 6460b) in fullscreen mode. The audio was presented through a pair of Sennheiser HD570 headphones connected to an RME Fireface 800 audio interface at the same approximate 65 dB level as in the VR condition.

Experiment Procedure. The participants partaking in the VR condition were the same as in the in the experiment described in section 4.2 Footstep Sounds and Virtual Body Awareness.

For the Video and Audio conditions, the experiment procedure was shorter than the VR condition. The participants were seated in by a desk in a quiet office space and received written instructions and a graphical user interface consisting of buttons (for triggering the stimuli) and sliders (for the likert scales). Once the participants had read through the instructions they were allowed to begin the evaluations, listening and watching (depending on the condition) to the pre-recorded walk as many times as they needed and giving the weight and suitability ratings, before continuing on to the next. The question had the same formulation as in the VR experiment. They were not allowed to go back and change their ratings, or deviate from the presentation order which was randomized before each session.

Discussion. The degree of support for our first hypothesis is not very strong. There were three out of 18 comparisons with significant differences for both weight estimates and suitability ratings. In concrete there were the three significant cases and for the footsteps on the wooden material there were no differences at all between the different presentation formats for weight estimates and the suitability ratings. Despite this, the Audio condition rendered the most significant differences between the filters, as if the effect of the different filters were more pronounced in that context. That might seem obvious as the Audio only condition offers the least amount of distraction, but it might also have been that the Video and the VR conditions were not distractive enough. An action packed computer game presented in the VR condition might have yielded a different result.

The suitability for two of the filters (b and c) were also given significantly higher ratings in the VR condition than in Audio once (ConcB) and in Video twice (ConcB and ConcC). This may hint that the VR presentation format may make users more tolerant to poor sound effects.

As various interactive tasks may have different demands on the attentional resources of the user. In order to learn more about how demanding the employed tasks are, some kind of instrument should be applied, such as the NASA task load index (NASA-TLX) [51] for measuring efforts required. This information could then be taken into account when studying results from audio quality judgements provided by users of VR simulations with varying levels of interaction complexity (such as wandering around and exploring, performing simple tasks or complicated tasks that must be completed within a limited amount of time).

What we also can see from the analysis is that the effect of the audio filters follow similar patterns in all presentation formats, with lighter estimates for higher center frequencies. This is especially pronounced for the wood material and less prominent for the concrete.

The results hint that a presentation format with only audio would yield a more pronounced effect in a comparison of different sound designs than in a presentation accompanied by visual feedback or as in an immersive VR format. These findings are somewhat coherent with previous research that suggests that tasks such as computer games may change the user's experience of a sound design or even reduce the user's ability to detect impairments in sound quality.

5 Conclusions

In this chapter we introduced several research directions related to walking, specifically on simulating audio and haptic sensation of walking, simulating walking in virtual reality and walking as a rhythmic activity. We presented ways of audio and haptic feedback generation in a form of footsteps natural and unnatural sounds. It was shown how different types of feedback can influence our behavior and perception. Our plans for the near future is to broaden the explanation of how can we actually can manipulate these from the neurological perspective to build more efficient, precise, and goal-directed feedback systems.

References

1. Li, X., Logan, R.J., Pastore, R.E.: Perception of acoustic source characteristics: Walking sounds. *The Journal of the Acoustical Society of America* 90 (1991) 3036–3049
2. Giordano, B., Bresin, R.: Walking and playing: What's the origin of emotional expressiveness in music. In: Proc. Int. Conf. Music Perception and Cognition. (2006)
3. Pastore, R.E., Flint, J.D., Gaston, J.R., Solomon, M.J.: Auditory event perception: The source perception loop for posture in human gait. *Perception & psychophysics* 70 (2008) 13–29
4. Giordano, B. L., McAdams, S., Visell, Y., Cooperstock, J., Yao, H. Y., Hayward, V.: Non-visual identification of walking grounds. *The Journal of the Acoustical Society of America* 123 (2008) 3412–3412
5. Young, W., Rodger, M., Craig, C.M.: Perceiving and reenacting spatiotemporal characteristics of walking sounds. *Journal of Experimental Psychology: Human Perception and Performance* 39 (2013) 464
6. Turchet, L., Serafin, S., Cesari, P.: Walking pace affected by interactive sounds simulating stepping on different terrains. *ACM Transactions on Applied Perception (TAP)* 10 (2013) 23
7. Maculewicz, J., Jylha, A., Serafin, S., Erkut, C.: The effects of ecological auditory feedback on rhythmic walking interaction. *MultiMedia, IEEE* 22 (2015) 24–31
8. Nordahl, R., Turchet, L., Serafin, S.: Sound synthesis and evaluation of interactive footsteps and environmental sounds rendering for virtual reality applications. *Visualization and Computer Graphics, IEEE Transactions on* 17 (2011) 1234–1244
9. De Dreu, M., Van Der Wilk, A., Poppe, E., Kwakkel, G., Van Wegen, E.: Rehabilitation, exercise therapy and music in patients with parkinson's disease: a meta-analysis of the effects of music-based movement therapy on walking ability, balance and quality of life. *Parkinsonism & related disorders* 18 (2012) S114–S119

10. Thaut, M.H., Abiru, M.: Rhythmic auditory stimulation in rehabilitation of movement disorders: a review of current research. (2010)
11. McIntosh, G.C., Brown, S.H., Rice, R.R., Thaut, M.H.: Rhythmic auditory-motor facilitation of gait patterns in patients with parkinson's disease. *Journal of Neurology, Neurosurgery & Psychiatry* 62 (1997) 22–26
12. Suteerawattananon, M., Morris, G., Etnyre, B., Jankovic, J., Protas, E.: Effects of visual and auditory cues on gait in individuals with parkinson's disease. *Journal of the neurological sciences* 219 (2004) 63–69
13. Roerdink, M., Lamoth, C.J., Kwakkel, G., Van Wieringen, P.C., Beek, P.J.: Gait coordination after stroke: benefits of acoustically paced treadmill walking. *Physical Therapy* 87 (2007) 1009–1022
14. Nieuwboer, A., Kwakkel, G., Rochester, L., Jones, D., van Wegen, E., Willems, A.M., Chavret, F., Hetherington, V., Baker, K., Lim, I.: Cueing training in the home improves gait-related mobility in parkinsons disease: the rescue trial. *Journal of Neurology, Neurosurgery & Psychiatry* 78 (2007) 134–140
15. Thaut, M., Leins, A., Rice, R., Argstatter, H., Kenyon, G., McIntosh, G., Bolay, H., Fetter, M.: Rhythmic auditory stimulation improves gait more than ndt/bobath training in near-ambulatory patients early poststroke: a single-blind, randomized trial. *Neurorehabilitation and neural repair* 21 (2007) 455–459
16. Bank, P.J., Roerdink, M., Peper, C.: Comparing the efficacy of metronome beeps and stepping stones to adjust gait: steps to follow! *Experimental brain research* 209 (2011) 159–169
17. Storms, R., Zyda, M.: Interactions in perceived quality of auditory-visual displays. *Presence* 9 (2000) 557–580
18. Sanders Jr, R.D.: The effect of sound delivery methods on a users sense of presence in a virtual environment. PhD thesis, Naval Postgraduate School (2002)
19. Chueng, P., Marsden, P.: Designing auditory spaces to support sense of place: the role of expectation. In: CSCW Workshop: The Role of Place in Shaping Virtual Community. Citeseer, Citeseer (2002)
20. Freeman, J., Lessiter, J.: Hear there & everywhere: the effects of multi-channel audio on presence. In: Proceedings of ICAD. (2001) 231–234
21. Västfjäll, D.: The subjective sense of presence, emotion recognition, and experienced emotions in auditory virtual environments. *CyberPsychology & Behavior* 6 (2003) 181–188
22. Larsson, P., Västfjäll, D., Kleiner, M.: Perception of self-motion and presence in auditory virtual environments. In: Proceedings of Seventh Annual Workshop Presence 2004. (2004) 252–258
23. Kapralos, B., Zikovitz, D., Jenkin, M.R., Harris, L.R.: Auditory cues in the perception of self motion. In: Audio Engineering Society Convention 116, Audio Engineering Society (2004)
24. Väljamäe, A., Larsson, P., Västfjäll, D., Kleiner, M.: Travelling without moving: Auditory scene cues for translational self-motion. In: Proceedings of ICAD05. (2005)
25. Tonetto, P.L.M., Klanovicz, C.P., Spence, P.C.: Modifying action sounds influences peoples emotional responses and bodily sensations. *i-Perception* 5 (2014) 153–163
26. Bresin, R., de Witt, A., Papetti, S., Civilani, M., Fontana, F.: Expressive sonification of footstep sounds. In: Proceedings of ISon 2010: 3rd Interactive Sonification Workshop. (2010) 51–54
27. Rocchesso, D., Bresin, R., Fernstrom, M.: Sounding objects. *MultiMedia, IEEE* 10 (2003) 42–52
28. Avanzini, F., Rocchesso, D.: Modeling collision sounds: Non-linear contact force. In: Proc. COST-G6 Conf. Digital Audio Effects (DAFx-01). (2001) 61–66
29. Cook, P.: Physically Informed Sonic Modeling (PhISM): Synthesis of Percussive Sounds. *Computer Music Journal* 21 (1997) 38–49

30. Gaver, W.W.: What in the world do we hear?: An ecological approach to auditory event perception. *Ecological psychology* 5 (1993) 1–29
31. Turchet, L., Serafin, S., Dimitrov, S., Nordahl, R.: Physically based sound synthesis and control of footsteps sounds. In: Proceedings of Digital Audio Effects Conference. (2010) 161–168
32. Nordahl, R., Serafin, S., Turchet, L.: Sound synthesis and evaluation of interactive footsteps for virtual reality applications. In: Proceedings of the IEEE Virtual Reality Conference. (2010) 147–153
33. Turchet, L., Serafin, S., Nordahl, R.: Examining the role of context in the recognition of walking sounds. In: Proceedings of Sound and Music Computing Conference. (2010)
34. Turchet, L., Nordahl, R., Berrezag, A., Dimitrov, S., Hayward, V., Serafin, S.: Audio-haptic physically based simulation of walking on different grounds. In: Proceedings of IEEE International Workshop on Multimedia Signal Processing, IEEE Press (2010) 269–273
35. Maculewicz, J., Cumhur, E., Serafin, S.: An investigation on the impact of auditory and haptic feedback on rhythmic walking interactions. *International Journal of Human-Computer Studies* 2015submitted.
36. Maculewicz, J., Cumhur, E., Serafin, S.: The influence of soundscapes and footsteps sounds in affecting preferred walking pace. *International Conference on Auditory Display* 2015submitted.
37. Maculewicz, J., Nowik, A., Serafin, S., Lise, K., Króliczak, G.: The effects of ecological auditory cueing on rhythmic walking interaction: Eeg study. (2015)
38. Sikström, E., de Götzen, A., Serafin, S.: Self-characteristics and sound in immersive virtual reality - estimating avatar weight from footprint sounds. In: *Virtual Reality (VR), 2015 IEEE*, IEEE (2015)
39. Tajadura-Jiménez, A., Basia, M., Deroy, O., Fairhurst, M., Marquardt, N., Bianchi-Berthouze, N.: As light as your footsteps: altering walking sounds to change perceived body weight, emotional state and gait. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, ACM (2015) 2943–2952
40. Zielinski, S.K., Rumsey, F., Bech, S., De Bruyn, B., Kassier, R.: Computer games and multichannel audio quality-the effect of division of attention between auditory and visual modalities. In: *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, Audio Engineering Society (2003)
41. Kassier, R., Zielinski, S.K., Rumsey, F.: Computer games and multichannel audio quality part 2'evaluation of time-variant audio degradations under divided and undivided attention. In: *Audio Engineering Society Convention 115*, Audio Engineering Society (2003)
42. Reiter, U., Weitzel, M.: Influence of interaction on perceived quality in audiovisual applications: evaluation of cross-modal influence. In: *Proc. 13th International Conference on Auditory Displays (ICAD)*, Montreal, Canada. (2007)
43. Rumsey, F., Ward, P., Zielinski, S.K.: Can playing a computer game affect perception of audio-visual synchrony? In: *Audio Engineering Society Convention 117*, Audio Engineering Society (2004)
44. Reiter, U., Weitzel, M.: Influence of interaction on perceived quality in audio visual applications: subjective assessment with n-back working memory task, ii. In: *Audio Engineering Society Convention 122*, Audio Engineering Society (2007)
45. Reiter, U.: Toward a salience model for interactive audiovisual applications of moderate complexity. (audio mostly) 101
46. Larsson, P., Västfjäll, D., Kleiner, M.: Ecological acoustics and the multi-modal perception of rooms: real and unreal experiences of auditory-visual virtual environments. (2001)
47. Larsson, P., Västfjäll, D., Kleiner, M.: Auditory-visual interaction in real and virtual rooms. In: *Proceedings of the Forum Acusticum, 3rd EAA European Congress on Acoustics*, Sevilla, Spain. (2002)

48. Larsson, P., Västfjäll, D., Kleiner, M.: The actor-observer effect in virtual reality presentations. *CyberPsychology & Behavior* 4 (2001) 239–246
49. Harrison, W.J., Thompson, M.B., Sanderson, P.M.: Multisensory integration with a head-mounted display: background visual motion and sound motion. *Human Factors: The Journal of the Human Factors and Ergonomics Society* (2010)
50. Thompson, M.B., Sanderson, P.M.: Multisensory integration with a head-mounted display: Sound delivery and self-motion. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50 (2008) 789–800
51. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology* 52 (1988) 139–183

New Robotic Platform for a Safer and More Optimal Treatment of the Supracondylar Humerus Fracture

Mohamed Oussama Ben Salem^{1,3}, Zied Jlalia², Olfa Mosbah³,
Mohamed Khalgui³ and Mahmoud Smida²

¹Tunisia Polytechnic School, University of Carthage, 2078, Tunis, Tunisia

²Orthopedic Institute of Mohamed Kassab, University Tunis El Manar, 2009, Tunis, Tunisia

³LISI Laboratory, INSAT, University of Carthage, 1080, Tunis, Tunisia

bensalem.oussama@hotmail.com, {zied_j, mahmoud.smida}@yahoo.fr,
{olfamosbah, khalgui.mohamed}@gmail.com

Abstract. Treating the supracondylar humerus fracture, a very common elbow's injury, can be very challenging for pediatric orthopedic surgeons. Actually, using the pinning technique to treat it leads sometimes to many neurological and vascular complications. Furthermore, the medical staff faces a serious danger when performing such surgeries because of the recurrent exposure to harmful radiations emitted by the fluoroscopic C-arm. Considering these issues, a national project was launched to create a new robotic platform, baptized BROS, to automate the supracondylar humerus fracture's treatment and remedy the said issues.

1 Introduction

When treating bone injuries, orthopedic surgeons often need precision, both in bone removal and in the placement of prosthetics, artificial devices that replace a missing body part [1]. This is due to the fact that, contrarily to soft tissues, bone is actually rigid and does not alter its shape once fully grown. Preoperative scans such as X-ray or CT (Computed Tomography) are common and procedures are planned in advance. These properties have made orthopedic surgery a privileged candidate for the implementation of medical robots. Also, as most procedures are not life threatening, there has been less skepticism over the implementation of these systems. Although most surgeons are satisfied with the outcome of conventional techniques [2], pressure to improve efficiency, implement less invasive procedures by reducing exposure of bony structures has enabled research into the area of Computer-Assisted Orthopedic Surgery (CAOS).

The supracondylar fracture of the humerus (or SCH) is one of the most common injuries faced by pediatric orthopedic surgery. It accounts for 18% of all pediatric fractures and 75% of all elbow fractures [7]. Occurring mainly during the first decade of life, it is more common among boys [8]. Completely displaced fracture can be one of the most difficult fractures to treat. The optimal aim of treatment is to obtain and maintain alignment of the fracture to allow full functional recovery of the elbow without residual deformity. This could be achieved through a reduction and stabilization of the fracture, which could be obtained using several approaches. But because of their best results and outcomes, closed reduction and lateral percutaneous pinning has become the

standard of care for most displaced supracondylar fracture [9]. This surgical technique requires an image intensifier and successive radioscopic images to control fracture reduction, and pin fixation. This technique fails in up to 25% of patients and some of them need re-operation because of inadequate reduction or wrong positioning of wires [10]. Inadequate reduction and/or insufficient stabilization can produce cubitus varus deformity, the most common complication. However, this "blind" surgical technique may also lead to neurovascular complications by pinning and damaging brachial artery or nerves [12, 13]. Another major inconvenient of the percutaneous pinning is the recurrent medical staff exposure to radiations when using the fluoroscopic C-arm [14]. These X-ray radiations are harmful, and fluoroscopic examinations usually involve higher radiation doses than simple radiographs. In fact, radiation exposures for spine surgeons may approach or exceed guidelines for cumulative exposure [15]. Another research showed that the fluoroscopically assisted placement of pedicle screws in adolescent idiopathic scoliosis, may expose surgeons to radiation levels that exceed established life-time dose equivalent limits [16]. The study in [11] shows that this exposure is responsible for the genesis of cancer, especially the thyroid one.

Considering these constraints and issues, a new national project, baptized BROS (Browser-based Reconfigurable Orthopedic Surgery), has been launched in Tunisia to remedy these problems. BROS is a multidisciplinary project reuniting the LISI Laboratory (INSAT), the Orthopedic Institute of Mohamed Kassab, ARDIA and eHTC. This work is carried out within a MOBIDOC PhD thesis of the PASRI program, EU-funded and administered by ANPR (Tunisia). BROS a new reconfigurable robotized platform dedicated to the treatment of supracondylar humeral fractures. It is capable of running under several operating modes to meet the surgeon's requirements and well-defined constraints. Thus, it can whether automatically perform the whole surgery or bequeath some tasks to the surgeon.

This chapter is organized as follows: the next section introduces the classification of supracondylar humeral fracture and its current treatment. The issues faced during the latter are also highlighted. Section 3 presents the architecture of the national project BROS and the reconfiguration modes under which it may run. We explain, then, how BROS will treat a SCH. Finally, we finish this work in Section 4 by a conclusion and an exposition of our future works.

2 Supracondylar Humerus Fracture

We present, in this section, the classification of supracondylar humeral fracture and how it is currently treated.

2.1 Classification of Supracondylar Humeral Fracture

Many classifications of the supracondylar humeral fractures were established. They are based on both the direction and the degree of displacement of the distal fragment [3]. The Lagrange classification system and the Gartland's are the most widely used. The first is the most widely used in the French literature. It divides these fractures into four types on the basis of antero-posterior and lateral radiographs [4]. In the English

literature, the second is the most commonly used: the Gartland's classification is based on the lateral radiograph and fractures are classified, as illustrated in Figure 1, according to a simple three-type system (Table 1) [5]. We adopt this classification in this paper.

Table 1. Gartland's classification of supracondylar fractures of the humerus.

Type	Radiologic characteristics
I	Undisplaced fractures
II	Displaced fracture with intact posterior hinge
III	Completely displaced fractures with no contact between the fragments

Fig. 1. Gartland's classification of supracondylar fractures of the humerus.

2.2 Supracondylar Humeral Fracture Treatment

In this section, we expose the treatment which was performed on a true case of a patient presenting a supracondylar humeral fracture who came to the Children Hospital of Béchir Hamza (Tunis). The patient who is a ten-year-old girl fell on her outstretched right hand on November 12th 2013. After clinical examination and radiological diagnosis, the patient's elbow was immobilized in a plaster splint and the patient was admitted in the pediatric orthopedics department and operated on the same day. Radiographs have showed a type III fracture according to Gartland's classification as show in Figure 2.

We were invited by Dr. Mahmoud SMIDA (Professor Medical Doctor, Head of Pediatric and Adolescent Orthopedics Department), our medical collaborator, to attend the surgical intervention. Closed reduction of fracture and lateral percutaneous pinning was performed under general anesthesia and fluoroscopic control. The injured elbow was, then, placed under the fluoroscopic image intensifier (Figure 3). The fracture was reduced by external maneuvers: pulling gentle, longitudinal traction and correcting frontal displacement, flexing the elbow and pushing anteriorly on the olecranon, hyperflexing the elbow and confirming maintenance of coronal alignment. Reduction was controlled by the image intensifier and a total of 9 radioscopic images were taken. The elbow was immobilized once a satisfying reduction was achieved (Figure 4). As illustrated in Figure 6, two lateral and parallel smooth pins were then percutaneously inserted from the lateral condyle through the opposite cortical bone to stabilize the fracture. After the placement of the two pins, the second pin had to be removed and reinserted since it did not straightaway follow the right trajectory. In this step, 15 fluoroscopic images were taken. After placement, the pins were bent over and cut off outside the skin. A long arm cast was then applied at the elbow in approximately 90° of flexion.

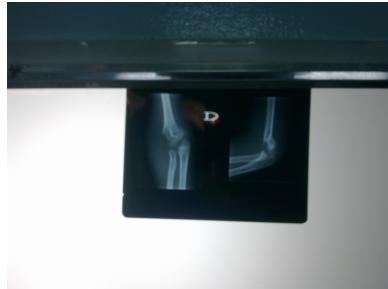


Fig. 2. The fracture's radiographs.



Fig. 3. The injured elbow installed under the fluoroscopic image intensifier.



Fig. 4. Elbow immobilization after obtaining fracture reduction.

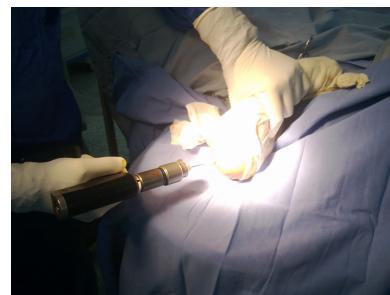


Fig. 5. Lateral percutaneous pinning.

During this total surgery, a total of 24 fluoroscopic images were taken, which involves high doses of radiation to the medical staff, especially since such interventions are performed 2 times per day on average.

3 Industrial National-European Project: BROS

We present in this section BROS's architecture and reconfiguration modes. We expose, thereafter, the constraints which have to be followed while implementing this robotized platform.

3.1 Architecture of BROS

BROS is a robotic platform dedicated to humeral supracondylar fracture treatment. It is able to reduce fractures, block the arm and fix the elbow bone's fragments by pinning. It also offers a navigation function to follow the pins' progression into the fractured elbow. BROS is, as shown in the class diagram hereafter, composed of a browser (BW), a control unit (UC), a middleware (MW), a pinning robotic arm (P-BROS) and 2 blocking and reducing arms (B-BROS1 and B-BROS2). The said components are detailed hereafter.

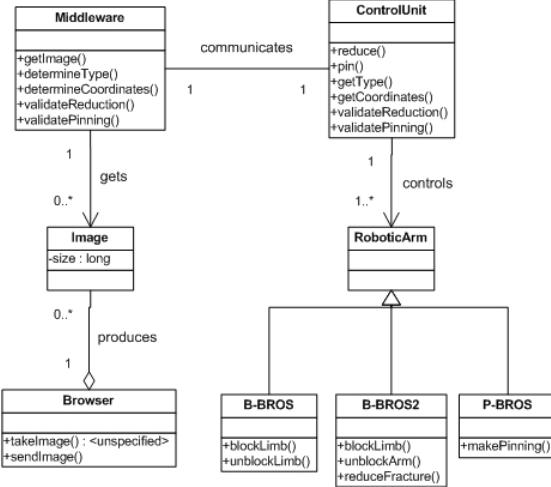


Fig. 6. BROS's class diagram.

Browser. The browser, which is a Medtronics's product and called FluoroNav, is a combination of specialized surgical hardware and image guidance software designed for use with a StealthStation Treatment Guidance System. Together, these products enable a surgeon to track the position of a surgical instrument in the operating room and continuously update this position within one or more still-frame fluoroscopic images acquired from a C-Arm. The advantages of this virtual navigation over conventional fluoroscopic navigation include: (i) the ability to navigate using multiple fluoroscopic views simultaneously, (ii) the ability to remove the C-Arm from the operative field during navigation, (iii) significant reduction in radiation exposure to the patient and staff. In addition, the FluoroNav System allows the surgeon to: (i) simulate and measure instrument progression or regression along a surgical trajectory, (ii) save instrument trajectories, and display the angle between two saved trajectories or between a saved trajectory and the current instrument trajectory, (iii) measure the distance between any two points in the cameras field of view, (iv) measure the angle and distance between a surgical instrument and a plane passing through the surgical field (such as the patient midplane). Primary hardware components in the FluoroNav System include the FluoroNav Software, a C-Arm Calibration Target, a reference frame, connection cables, and specialized surgical instruments.

Control Unit. The CU ensures the smooth running of the surgery and its functional safety. It asks the supracondylar fracture's type to the middleware, and then computes, according to it, the different coordinates necessary to specify the robotic arms' behaviors concerning the fracture's reduction, blocking the arm and performing pinning. The surgeon monitors the intervention progress thanks to a dashboard installed on the CU.

Middleware. The middleware is a software installed on the browser and which acts as a mediator between the CU and the BW. It is an intelligent component that provides

several features of real-time monitoring and decision making. The middleware contains several modules: (i) an image processing module, (ii) a controller, (iii) a communication module with the CU.

Pinning Robotic Arm. The pinning robotic arm, P-BROS, inserts two parallel Kirschner wires according to Judet technique [6] to fix the fractured elbow's fragments. To insure an optimal postoperative stability, BROS respects the formula:

$$S = B/D > 0.22 \quad (1)$$

where S is the stability threshold, B the distance separating the two wires and D the humeral palette's width [17].

Blocking and Reducing Robotic Arms. B-BROS1 blocks the arm at the humerus to prepare it to the fracture reduction. B-BROS2 performs then a closed reduction to the fractured elbow before blocking it once the reduction is properly completed.

3.2 Reconfiguration Modes

Reconfiguration is an important feature of BROS. It is designed to be able to operate in different modes. The surgeon can actually decide to manually perform a task if BROS does not succeed to automatically perform it, whether it is fracture reduction, blocking the arm or pinning the elbow. Thus, five different operating modes are designed and detailed hereafter: (i) Automatic Mode (AM): The whole surgery is performed by BROS. The surgeon oversees the operation running, (ii) Semi-Automatic Mode (SAM): The surgeon reduces the fracture. BROS performs the remaining tasks, (iii) Degraded Mode for Pinning (DMP): BROS only realizes the pinning. It is to the surgeon to insure the rest of the intervention, (iv) Degraded Mode for Blocking (DMB): BROS only blocks the fractured limb. The remaining tasks are manually done by the surgeon, (v) Basic Mode (BM): The whole intervention is manually performed. BROS provides navigation function using the middleware that checks in real time the smooth running of the operation.

3.3 Constraints Definition

To treat a humeral supracondylar fracture using BROS, the following steps are performed in the automatic mode:

- i) the surgeon launches the system and chooses one of the five operating modes;
- ii) CU asks MW about the fracture coordinates;
- iii) MW requests an image from BW and the latter sends it;
- iv) MW determines the different coordinates by image processing and sends them to CU;
- v) based on the received coordinates, CU orders B-BROS1 to block the arm at the humerus;

- vi) B-BROS1 blocks the limb;
- vii) CU asks B-BROS2 to reduce the fracture based on the latter's line;
- viii) B-BROS2 reduces the fracture;
- ix) CU asks MW to ensure that the reduction was successful;
- x) MW requests a new image from BW and checks the fracture reduction result. If it is satisfactory, BROS moves to step xi. Steps from vii to ix are repeated otherwise;
- xi) CU orders B-BROS2 to block the arm;
- xii) under the request of UC, P-BROS performs the first and the second pinning;
- xiii) once the pinning is successful, CU asks B-BROS1 and B-BROS2 to unblock the limb.

4 Conclusion

The work presented in this chapter consists in introducing a new robotic platform dedicated to the treatment of supracondylar humerus fracture, and its contributions. BROS is a flexible system since it may run under different operating modes to meet the surgeon requirements and the environment constraints: it is reconfigurable. Recent works proved the usefulness of this robotic platform to avoid complications that may be generated because of the blind pinning and prevent the danger posed by the recurrent exposition to radiations [18, 19]. We can, now, certify that BROS is an innovating project which will be of a great help to pediatric orthopedic surgeons. The next step is to proceed to the real implementation of BROS using the ABB robotic arms.

References

1. Christensen, R. W. (1970). U.S. Patent No. 3,488,779. Washington, DC: U.S. Patent and Trademark Office.
2. Jaramaz, B., Hafez, M. A., & DiGioia, A. M. (2006). Computer-assisted orthopaedic surgery. Proceedings of the IEEE, 94(9), 1689-1695.
3. Barton, Kelly L., et al. "Reliability of a modified Gartland classification of supracondylar humerus fractures." Journal of Pediatric Orthopaedics 21.1 (2001): 27-30.
4. Lagrange, J., and P. Rigault. "Fractures supra-condyliennes." Rev Chir Orthop 48 (1962): 337-414.
5. Pirone, A. M., H. K. Graham, and J. I. Krajbich. "Management of displaced extension-type supracondylar fractures of the humerus in children." J Bone Joint Surg Am 70.5 (1988): 641-50.
6. Judet, J. E. A. N. "Traitement des fractures sus-condyliennes transversales de l'humrus chez l'enfant." Rev Chir Orthop 39 (1953): 199-212.
7. Landin, Lennart A., and Lars G. Danielsson. "Elbow fractures in children: an epidemiological analysis of 589 cases." Acta Orthopaedica 57.4 (1986): 309-312.
8. Landin, Lennart A. "Fracture Patterns in Children: Analysis of 8,682 Fractures with Special Reference to Incidence, Etiology and Secular Changes in a Swedish Urban Population 1950-1979." Acta Orthopaedica 54.S202 (1983): 3-109.
9. Zhao, Jia-Guo, Jia Wang, and Peng Zhang. "Is lateral pin fixation for displaced supracondylar fractures of the humerus better than crossed pins in children?" Clinical Orthopaedics and Related Research 471.9 (2013): 2942-2953.

10. Aronson, D. C., E. Van Vollenhoven, and J. D. Meeuwis. "K-wire fixation of supracondylar humeral fractures in children: results of open reduction via a ventral approach in comparison with closed treatment." *Injury* 24.3 (1993): 179-181.
11. Schonfeld, S. J., C. Lee, and A. Berrington de Gonzlez. "Medical exposure to radiation and thyroid cancer." *Clinical Oncology* 23.4 (2011): 244-250.
12. Gosens, Taco, and Karst J. Bongers. "Neurovascular complications and functional outcome in displaced supracondylar fractures of the humerus in children." *Injury* 34.4 (2003): 267-273.
13. Flynn, JOSEPH C., JOSEPH G. Matthews, and ROGER L. Benoit. "Blind pinning of displaced supracondylar fractures of the humerus in children." *J Bone Joint Surg Am* 56.2 (1974): 263-72.
14. Clein, Norman W. "How safe is X-ray and fluoroscopy for the patient and the doctor??" *The Journal of pediatrics* 45.3 (1954): 310-315.
15. Rampersaud, Y. Raja, et al. "Radiation exposure to the spine surgeon during fluoroscopically assisted pedicle screw insertion." *Spine* 25.20 (2000): 2637-2645.
16. Haque, Maahir Ul, et al. "Radiation exposure during pedicle screw placement in adolescent idiopathic scoliosis: is fluoroscopy safe??" *Spine* 31.21 (2006): 2516-2520.
17. Smida, M., et al. "Un index de stabilité pour l'embrochage percutané latéral parallèle des fractures supracondylériennes du coude chez l'enfant." *Revue de Chirurgie Orthopédique et Réparatrice de l'Appareil Moteur* 93.4 (2007): 404.
18. Ben Salem, Mohamed O., et al. "ZiZo: Modeling, Simulation and Verification of Reconfigurable Real-Time Control Tasks Sharing Adaptive Resources. Application to the Medical Project BROS." *Proceedings of the 8th Int. Conference on Health Informatics* (2015): 20-31.
19. Ben Salem, Mohamed O., et al. "BROS: A New Robotic Platform for the Treatment of Supracondylar Humerus Fracture." *Proceedings of the 8th Int. Conference on Health Informatics* (2015): 151-163.

Robotics Enabled Augmented Health

Artur Arsenio

IST-ID, Universidade da Beira Interior, YdreamsRobotics
Edifício A Moagem - Cidade do Engenho e das Artes, Largo da Estação 6230-311 Fundão
{artur.arsenio}@ydreamsrobotics.com

Abstract. Nowadays it is increasingly important, for social and economic reasons, to provide augmented health assistance for people at home. This will mostly benefit some specific user groups, such as elderly, patients recovering from physical injury, or athletes. This chapter describes the application of robotics, under the scope of the Augmented Human Assistance (AHA) project, for assisting people health. Two complementary and interacting approaches are described. The first consists on a mobile robot nurse, that assists patients and elderly on their daily lives, such as for advising on medicine intake, or providing complementary biomedical sensing equipment for health monitoring. The other approach consists of multimodal feedback assistance through augmented reality, haptic sensing and audio, in order to guide and assist people on the execution of exercises at home.

1 Introduction

The world population is aging rapidly. There is an increasing need for health assistance personnel, such as nurses and physiotherapeutic experts, in developed countries. On the other hand, there is a need to improve health care assistance to the population, and especially to elderly people.

The World Health Organization estimates that sedentarism is the 4th main factor in worldwide mortality, and is associated with several diseases, such as breast and colon cancer, obesity, diabetes, and ischemic strokes. For instance, childhood obesity originates other health problems such as hypertension related to left ventricular hypertrophy, atherosclerosis and diastolic dysfunction. It is therefore important to identify children in risk, through continuous health monitoring (e.g. body temperature, blood pressure, electrocardiogram). Concerning elderly, estimates indicate that by 2030, ~4% of the USA population will have experienced a stroke, with related costs expected to rise from \$71.55 billion to \$183.13 billion between 2012 and 2030.

Hence, sedentarism is not only a social problem, but also economical, threatening the sustainability of current health systems. New technologies are required for enabling treatment at home, reducing pressure for health care personnel at hospitals. Furthermore, automated systems at hospitals will enable significant cost reductions and improved efficiency. Information technology has also been playing an increasing role in the health care area throughout the years, aiding it to be more accurate, faster to respond, and less susceptible to human errors. There is therefore the need not only to develop technological solutions that promote active aging and prevent sedentary behaviors, but also to find new technologies for assisting a growing, aging population.

In this context, the advances in information, robotic and assistive technologies have the potential to increase quality of life and change health care delivery models, reducing costs, and improving monitorization. The “AHA: Augmented Human Assistance” project is a novel, integrative and cross-disciplinary approach combining innovation and fundamental research in the areas of human computer interaction, robotics, serious games and physiological computing. AHA’s goal is to develop a new generation of ICT based solutions that have the potential to transform healthcare by optimizing resource allocation, reducing costs, improving diagnoses and enabling novel therapies, thus increasing quality of life. The project proposes the development and deployment of a novel Robotic Assistance Platform designed to support healthy lifestyle, sustain active aging, and support those with motor deficits.

The AHA project will develop a novel and modular set of ICT based solutions that in isolation or integrated as a Robotic Assistance Platform will address:

- Physical (re)training: Building on the existing expertise on Augmented Reality (AR) and serious games, we propose to develop adaptive AR physical training tools that deliver online feedback on performance to prevent sedentarism, support active aging and provide personalized tools for function re-training in motor impaired patients.
- Increasing self-awareness: Monitoring of user state by means of biosensors, computer vision systems and exercise performance data. User state will be assessed in a transparent manner and data will be visualized through friendly user interfaces, and shared with patients, clinicians and/or relatives.
- Augmented assistance: The above systems will be integrated on a mobile robotic platform with indoor navigation capabilities (in environments such as senior houses and hospitals) that will interact through a virtual coach system to assist patients, provide reminders on tasks, guide patients through exercises, and support them in daily routines.

These technologies may be very useful in other scenarios, such as those of MOnarCH project (Multi-Robot Cognitive Systems Operating in Hospitals). This CE FP7 project is focused on introducing a fleet of social robots that will interact with sick children and collaborate with medical staff on the pediatric ward of a hospital, the Portuguese Oncology Institute of Lisbon (IPOL).

1.1 The Need for Robots

There are several scenarios in which robots and augmented reality are very useful in the two aforementioned environments corresponding to these two abovementioned projects.

- Teaching, for robots to supporting human teachers by projecting augmented reality content on a wall or on objects.
- Patient rehabilitation exercises support, by projecting augmented reality content during physiotherapeutic activities in which the patient receives in real-time visual feedback and corrective postures
- Information providers, such as projection of AR content informing people that a new activity is about to start, or calling them with visual signs, or even moving along with people to places where action is going to take place (e.g. using projected directional arrows), or informing someone to stop performing an

exercise.

- People protection, such as projecting augmented reality content (for instance a stop sign) if a person moves into a forbidden zone or door, or performs a forbidden rehabilitation exercise or movement
- People entertainment: robots can play games with children or elderly people, according to content projected into the floor. In another scenario, a patient rehabilitation can involve game playing (serious games).

Furthermore, robots can perform several supporting activities at home, such as:

- support safe medicine delivery and intake. A robot may carry the medicine, and at the appropriate time, take it to the patient.
- In addition, the robot may also carry on board medical equipment, such as thermometer, arterial pressure measurement device, or electrocardiogram (ECG) equipment, making it available to patients at home.

1.2 Remote Health Assistance

Ageing population has an enormous economic and social impact in a various areas, especially healthcare systems. Elderly people are more vulnerable to physical or mental impairments, disabilities and chronic illnesses. Falls and problems with muscle bone can also limit the daily routines. Hence, such problems increase the potential need for assistance of elderly people.

Patient health monitoring was traditionally done through periodical visits to the doctor in order to undergo on-site tests for blood pressure, pulse, temperature, or sugar level. The alternative is stationary monitoring, upon internship at a health care provider. Currently, patients have the possibility to take home sensors attached on a belt, usually for a specific time period, to collect biosignal data for that period. However, the patient still has to move to the health care provider to get the sensor, and afterwards, he/she has to return to remove the sensor and deliver the data. Nowadays, we are reaching seamless health monitoring by placing a mobile sensor on the patient, executing biosignal collection, and transmitting data through a wireless access technology interface to a storage facility for further medical analysis [1].

On the other hand, medical assistance personnel play a very important role on patient recovery. In physical therapy, for instance, the therapist helps patients fighting their pain and recovering from injury. His role is fundamental on therapy planning. In addition, he not only demonstrates the correct execution procedure for the exercises, but he also makes sure afterwards the exercises are executed correctly. With this goal in mind, the therapist can intercede during the session and adapt the exercise schedule according to the patient's feedback [2]. However, the patient might perform incorrect movements at home, without the therapist presence, to avoid pain [3]. Hence, there is a need to provide assistive solutions that do monitor remotely the patients' execution of exercises, but also support, motivate and advise the patient to correctly perform the exercises. The former is accomplished through biosensors, augmented reality and haptic technology, that should monitor continuously the patient, transmitting wirelessly information concerning its health state and providing feedback. An interacting mobile robot should provide the patient motivation and guidance.

1.3 Chapter Structure

This chapter will start by presenting a review of previous relevant literature for augmented health assistance. Afterwards, section 3 overviews the Augmented Human Assistance project, namely its structure and challenges. Section 4 addresses the design of robot assistants targeting the AHA project user groups, which will act as nurses for elderly people and patients in recovery of physical injury. It will be also shown that robotics and augmented reality can provide further functionalities for providing visual or haptic feedback to users, as described on section 5. Finally, section 6 will draw the main conclusions, together with directions for future work.

2 Background Review

Various studies show evidence physically active elderly having lower rates of chronic conditions such as cardiovascular diseases, diabetes, cancer, hypertension and obesity [4]. There is the need, therefore, to develop solutions that promote healthy habits and prevent sedentarism. Since chronic patients experience loss of autonomy and low self-esteem, it is also important to provide assistance to patients with age related chronic conditions. With these goals in mind, hereafter we overview previous research works and discuss the most relevant strategies.

2.1 Service Robotics

Research interest in service robotics for active aging and health care has grown in the last few decades with potential applications on healthy people, elderly, children or patients. Robotic devices in elderly care [5][6], rehabilitation [7], autism diagnosis and therapy [8] and weight loss applications [9] have been empirically demonstrated to be effective. Hence, robotics raised great expectations on the use of robots as personal assistants. On such robot is the Nursebot platform, able to interact residents, remind or accompany them of an appointment, as well as provide information of interest to that person [10].

Intouch Health deployed their robot in a Neurology Intensive Care Unit. A study suggested improvement in critical care nursing team satisfaction [11]. The robot Paro in Japan (a robot resembling a baby seal, with expressive eyes) was reportedly able to improve the mood of elderly people, and simultaneously reduced stress not only to patients but also to their caregivers [12]. This has been demonstrated more recently to treat some cases of depression suffered by the survivors of the devastating earthquake and tsunami in the northeast coast of Japan in March of 2011.

Several solutions employing devices exist to support self-administration of medicine and manage personal medicine administration [13]. The Kompai robot has been tested for elderly assistance using a diary application for monitoring the medication and give information about daily events [11]. The autom robot is a weight-loss social robot [9] that asks questions concerning what a person ate, or how it exercised. It also provides personalized, helpful suggestions and feedback using facial expressions and a simple touchscreen interface. In Europe, the EMOTE (Robotic Tutors for Empathy based Learning) research project explores the usage of social robots for teaching chil-

dren, using the NAO robot. This project has shown that the robot appearance, as well as its motion, and functionality, plays a very important role on engaging the attention of the learner [14]. Other authors incorporated the strict functional constraints imposed by complex environments (such as hospitals) and specific groups of users, into the robot design process [15].

2.2 Biomedical Signal Analysis and Human State Estimation

Biomedical signal analysis is nowadays of greatest importance for data interpretation in medicine and biology, providing vital information about the condition and affective/emotional states of subjects. In patients with neuromuscular diseases, a constant monitoring of the patient's condition is necessary [16]. Heart rate variability, respiration, muscular and electrodermal activity signals are extremely important, since they indicate when a muscular crisis is occurring.

Tracking devices are important to infer the human state, such as the posture and motion of the human body during an activity (such as playing a sport or a therapeutic exercise). Usually, several tracking points are used to represent human body joints (with respect to both position and velocity). Cameras have been used to detect and estimate the pose of human subjects [17] and body parts [18], detect faces [19] and their expressions [20] and, at a close range, detect eye movement and gaze direction [19]. Hidden Markov Models have been successfully used in gait recognition [21].

One approach consists of using anthropometric proportions of human limbs and the characteristics of gait, to achieve view-independent markerless gait analysis [22]. Microsoft's Kinect is a markerless tracking device with an acceptable accuracy in comparison with other motion tracking alternatives [23][24]. It provides full skeleton tracking at low price. It is also easily portable, especially when compared to other solutions requiring special equipment (e.g. markers) on the human body. A large number of applications based on such sensors are addressing some of the difficulties in unsupervised rehabilitation [25][26]. For instance, Gama et al. [27] proposed a Kinect based rehabilitation system that tracks user position. The user sees himself on the screen with overlaying targets representing the desired position. Real-time feedback using visual messages is provided in case of incorrect postures.

2.3 Rehabilitation Systems

Various rehabilitation systems have been proposed to improve patient recovery. Many focus on specific injuries, e.g., stroke [25], or limbs rehabilitation [28][29]. Such systems may have an important impact on patient's rehabilitation on an ambulatory scenario (e.g. at home). Enabling the patient to comfortably exercise at home improves his motivation [25].

A patient's rehabilitation is affected by exercise repetition, expert feedback, and patient's motivation [30]. The repetitive nature of rehabilitation exercises can quickly become boring for a patient [28][31], therefore, there is a need of turning these exercises into something less tedious. Indeed, successful patient recovery depends on adherence to the scheduled planning [32].

Repetitive exercises should be divided into several sub-goals, so the patient achieves incremental success through each repetition. This improves motivation compared to the approach where success is only achieved after finishing the whole task [30]. Feedback can be given in two different ways, during the execution (concurrent feedback) and at the end of exercise execution (terminal feedback) [2]. Concurrent feedback is given in real-time for offering guidance or corrections in exercise execution. It allows the patient to have Knowledge of Performance (KP). Terminal feedback gives patients only Knowledge of Results (KR), since the patient receives feedback after fully executing the task [33][30]. Sigrist [2] suggests a temporal evolution along the recovery phases. It proposes to gradually reduce KP, giving more emphasis to KR, to stimulate patient's autonomy.

2.4 Motor Training on Rehabilitation

The term *exergaming* is used for gaming approaches that motivate players to engage in physical activity. Previous research [34] showed evidence that commercial tools can produce physical, social, and cognitive benefits [35].

Unfortunately, these current tools are not suited to elderly or motor (re)training. Motor rehabilitation, or motor re-learning, is an extensive and demanding process for a patient, requiring discipline. Moving injured body parts may produce discomfort or even significant pain [36]. Physical therapy sessions may be performed several times both at a clinical and at home, or on either one of them. Later on, the patient might have to continue therapy exercises at home [37] to avoid suffering a setback on rehabilitation [25] (or to decrease recovery time). This requires patients to learn the appropriate recovery exercises. Furthermore, these exercises should be executed correctly, to prevent an aggravation of the injury [3].

Repetition of specific movements is important for rehabilitation, whether at a clinic or at home [30]. However, this is one of the main causes of deteriorated rehabilitation at home, since patients tend to get bored and lose focus, due to the repetitive nature of the task [36]. A mobile robot nurse should play a very important role here in order to replace the therapist presence to guide and motivate the patient.

2.5 Augmented Reality

Augmented Reality (AR) is considered a promising technology in the rehabilitation field. Several studies presented evidence for the benefits of employing Augmented Reality (AR) techniques for supporting functional motor recovery [38], enabling effective motor training [39]. AR based approaches potentiate the combination of interesting features such as training customization, extended multimodal feedback, quantifiable therapy measures, extra motivation, among others [40].

Interactive AR applications have been proposed for different applications. Shader Lamps [41] is a seminal work on AR. Its purpose was to augment an existing blank model with computer-generated graphics to make the model exhibit realistic lighting, shadow and even animation effects, thus providing it with characteristics previously nonexistent. Dynamic Shader Lamps [42] allows users to interactively change the generated graphics, and to digitally paint onto objects with a stylus.

There are various spatial augmented reality (SAR) projects with different applications. iLamp [43] supports SAR by casting the pre-distorted image of virtual objects from the projector, which can capture the 3D environment based on structured patterns. An interactive SAR system is described in [44] with multiple depth cameras and a projector. The system detects the information of the surrounding environment along with the users motion through depth cameras, and displays images on the surface objects in a planar, real world table for the user to interact with. Pixelflex [45] adopted an array of multiple projectors to make a huge screen space. Automatically self-configured projectors are able to create a huge projection area.

A projection-based information display system was proposed in [46], which displays virtual images in a room with a series of sensors. The robot tracks the user viewpoint with the sensors, and then the robot can generate anamorphic (i.e., projection format in which a distorted image is stretched by an anamorphic projection lens to recreate the original aspect on the viewing screen), properly distorted images for users on flat-surfaced surfaces. WARP [47] allows designers to preview materials and finished products by projecting them onto rapid prototype models. The system uses a standard graphical interface, with a keyboard and mouse used for user interface. The projection is not restricted to a fixed area; all feedback is projected onto the tools. Surface Drawing [48] allows a user to sculpt 3D shapes using their hands. Spray modeling [49] uses a mock-up of a physical airbrush to allow the user to sculpt 3D models by spraying matter into a base mesh. The system enhances physical tools by projecting status information onto them allowing the overload of a single tool with several functions.

LightGuide [50] proposed projection mapping onto the user, using his body as a projection screen. Different type of real-time visual cues are projected onto the user's hand to guide him for performing 3D movements. This way, the user is less susceptible to be distracted by external factors.

3 The Augmented Human Assistance Project

This section described the Augmented Human Assistance project. It is presented the project structure of partners, its focus areas, as well as the main challenges and expected impact.

3.1 Project Partners

A consortium of key partners, addressing a large scope of technology fields, forms the AHA project. IST-ID is the project coordinator. The Computer and Robot Vision Lab (Vislab) research group at IST/Institute of Systems and Robotics addresses research on Robotics, Computer Vision, and Cognitive Systems. The Carnegie Mellon University (CMU) team is part of the Quality of Life Technology Centre, at the Robotics Institute, working on Artificial Intelligence and Human Computer Interaction. The M-ITI NeuroRehabLab at University of Madeira addresses Serious Games, Interactive Technology, and Rehabilitation. Two groups are specialized on biosignals data acquisition and processing. The team at FCT/UNL works on Signal Processing, Machine Learning, and Electrophysiology. And PLUX Biosignals, a company that creates innovative solutions for Sports, Healthcare and Research, by integrating biosignals

processing and miniaturized wireless sensor devices. YDreamsRobotics is another company specialized in robotics, mechatronics and the internet of things, which is addressing Augmented Reality, Robotics Design, and Haptic Feedback, for building a therapeutic robot and for providing patient feedback through actuation. Finally, the Interdisciplinary Centre for the Study of Human Performance (CIPER) at Faculdade de Motricidade Humana (FMH) is specialized in the main areas addressed by the project, namely Sports Science, Therapeutic exercises, and Human Function and Performance.

The project targets three main environments (clinical, sports, and home assistance), as described in the next subsection. Supporting external partners collaborate by providing solution requirements, on-going feedback, as well as facilities for empirical evaluation, as described in Table 1.

Table 1. External partners and their focus areas on the project.

Area	External partners
Rehabilitation	<i>HealthSouth of Sewickley</i> <i>Clinical Physiology Translational Unit, IMM</i>
Sports	<i>Associação da Madeira de Desporto Para Todos</i> <i>Via Activa Animação Turística</i>
Elderly	<i>Comfort Keepers</i>
Public Administration	<i>Camara Municipal da Ponta do Sol</i>

3.2 Project Focus Areas

The AHA project addresses three main target groups, as shown in Fig. 1. The Elderly People group, addressed both on a clinical setting as well as at home. The goal is to enable simultaneously sustainable care and active aging. New technologies are proposed, such as medicine and medical equipment delivery by a mobile robot, and remote assistance on therapeutic exercises using multimodal feedback and augmented reality. Remote assistance on therapeutic exercises is also proposed for a second target group, those with motor deficits. Serious games, emotional and haptic feedback, are among the technologies proposed to support patients on their recovery. Finally, the project addresses sports for supporting a healthy lifestyle. It aims to improve athletes' performance, and to facilitate their learning, training and performance evaluation using new technologies.

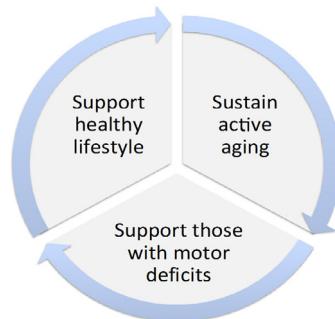


Fig. 1. User groups targeted by AHA: patients, athletes, and elderly people.

The three main target environments, namely clinical, sports facility, and home assistance at a senior house, have different characteristics.

The **clinical environment** is more accessible but it is often a more complex environment. A virtual coach module is proposed to interact with the patient, guiding him throughout the exercises and providing positive and corrective feedback.

Sports facility is characterized by large indoor open spaces where robot navigation is facilitated. The Robotic Assistance Platform will be able to identify an adequate surface for the projection of the Augmented Reality Training Games, and provide its users with the gaming training experience. Users will be able to challenge remote users to a competitive game, while playing the training games.

At a **senior house**, the Robotic Assistance Platform will visit some of the residents at specific times. The robot reminds the elder about the pills he needs to take, and also asks him to take heart rate measurements with the on board sensors of the robot. Furthermore, it will help the elder through his/ her gaming exercises, and it will provide a feedback summary on performance, making this information also remotely available to their clinicians.

3.3 Innovations and Impact

The AHA project proposes the development of novel robotics assistive platforms for health exercise program management. Several functionalities will be provided, such as user engagement, execution monitoring, and supporting therapists in patients' rehabilitation tasks. These functionalities will be developed considering the target user groups, with the aim to promote a healthy lifestyle, a sustainable active aging of the population, and supporting those with motor deficits. This should be accomplished using several interdisciplinary technologies, and combining these in innovative ways. For instance, it is proposed advanced "*exergaming*" and *assistive technologies* based on human-computer interaction, robotics, serious games, multimodal sensing, physiological computing, among other fields. The goal is to employ technology to help to prevent sedentarism related diseases, to facilitate therapy of chronic conditions related to aging and rehabilitation of motor deficits. Therefore, the project will tackle users with special needs, focusing on elderly and patients with motor deficits, as well as on athletes. It will provide customization and personalized tools for increasing motivation and engagement, using augmented reality, online feedback, gamification, social networking, and interactive interfaces. It is also proposed a better measurement and monitoring of user condition through multi-modal sensing (biosensors, computer vision) and multi-modal feedback (haptic devices, sound feedback, augmented reality). Augmented assistance will be provided by a nurse mobile robotic platform, together with virtual coaching, for guiding patients in their exercises and daily routines. The project also proposes to support clinicians through personalized user profiles and advanced display of information.

The project outcomes aim to impact significantly different entities:

- **Society:** Reduce effects of sedentarism related diseases and aging conditions.
- **Elderly:** Increase physical fitness, independence, autonomy, self-esteem.
- **Patients:** Can exercise independently at home, adherence to schedule.
- **Clinicians:** Follow patient progress remotely, customize therapies.

- **Science & Technology:** New algorithms and systems for eHealth technologies.
- **Education:** Formation of highly skilled human resources in close cooperation with technology integrators and end-users.
- **Companies:** Close to market technologies; large exploitation opportunities.

3.4 WorkPlan

AHA project work plan is shown in Fig. 2. It represents the platforms, the modules to be integrated on them, as well as the interactions among modules.

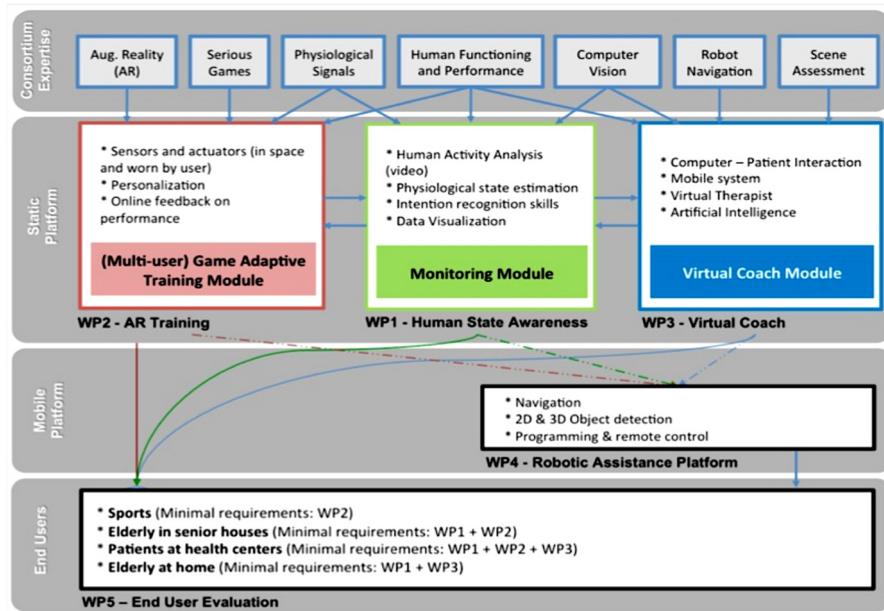


Fig. 2. Project workplan according to the different workpackages.

Human State Estimation (WP1) employs state of the art wearable technology and training, for the integration of vision based activity monitoring and emotion detection modules into a core technological and scientific component of the Robotic Assistant Platform - the human monitoring module. Augmented Reality Training (WP2) integrates AR technology that will serve as the basis for the development of a set of novel serious games for physical training. Integration of Virtual Coaches (WP3) on a physical context is aimed for the improvement of the quality of life of those with special needs. It includes (see fig. 3) user interaction aspects of the system, as well as interactions of the virtual coach and the human monitoring modules. Robotic Assistance Platform (WP4) consists of the sensing and software robot architecture. It includes integration of the monitoring, AR training, and virtual coach modules into the final Robotic Assistive Platform. User evaluation (WP5) will consist of the solution's experimental validation with real data on a clinical context.

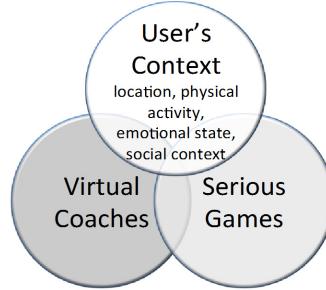


Fig. 3. Integration of interdisciplinary technologies on the AHA project.

4 Robot Design for Human Augmented Assistance

Recently, there is a trend for bringing robots into operations among people, helping people. Elderly people, athletes and patients constitute challenging groups for robot interaction. Addressing athletes' needs is challenging due to their energy and high performance requirements. Elderly people and patients pose strong safety concerns. These challenges need to be tackled on the design process. This is essential for the robots to be able to operate safely, while interacting socially with people.

Gonçalves and Arsenio [15] proposed a formal design process for building socially interacting robots within a group of children with special needs in a hospital environment. The design takes into consideration the input of the project partners. This is incorporated on the robot specifications during the development process. Robots design needs to account with several security factors. Under the scope of the AHA project, we have been applying this process to the design of mobile robots targeting new user groups. This "nurse" mobile robot, although posing similar challenges as the ones described by Gonçalves and Arsenio [15], also adds new requirements. Elderly people, athletes and patients poses specific requirements that often differ from those for children in hospitals. Concerning security factors and human-robot interaction features (see Fig. 4):

- Sick children often carry medical equipment carried, like wheeled structure to carry serum bags. This still applies to some groups of elderly people. It may also apply in less extent to patients on recovery from physical injury.
- No sharp surfaces that may cause injury to children. Although a less stringent requirement for elderly people and patients, it is still an important issue.
- Avoid geometries that may invite children to step up the vehicle or insert fingers on holes. This is not relevant for the group of users under consideration.
- Emotional expressions and engagement behaviors. A nurse robot targeting adults should pose a professional stance, compared to a more character like concept for children.
- Avoid inspiring fear and disgust: should instead induce comfort and trust, which applies for all user groups (except in specific niches, such as military)
- An appealing aesthetics to children often has a different meaning than for an adult.

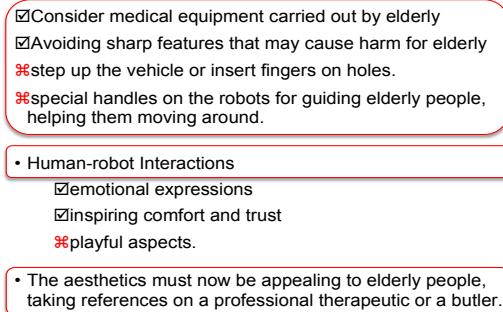


Fig. 4. Comparative analysis of security factors and human-robot interaction features, for robots to interact with children versus elderly people.

4.1 Shell Development Process

The shell development process is divided into several stages, and here briefly resumed. It started by an analysis of the environment and the way elderly people, athletes and patients should interact with the robots. At this stage we also performed research related to existing robots, and “soft” and “clean” materials prone to be used at clinical, hospital and sport facilities. With this information we concluded that rounded shapes, soft materials, and neutral, light colors should be used, together with active colors (red) on very small information symbols.

Different options were analyzed taking into account aesthetic properties, emotional feelings provoked by the robot, its the actual expression, feasibility constraints and functionality inherent to the robot, so that it could perform the envisioned tasks.

Afterwards, the material developed so far became the basis for the CAD model development. The CAD model and photo-real imagery was also developed taking into account the different production methods available, the assembly procedure, the maintenance operations, the cost involved, and of course the aesthetics and functional requirements defined at the early stages.

4.2 Operational Environment Analysis

The shell development process started by performing an analysis of the environments where the robots will operate, including the deployment space, targeted groups of people, and tasks to be executed by the robot, in order to define a set of features taking references on other existing and relevant systems.

The features to apply include rounded shapes and soft touch feeling materials, due to the danger of fall by elderly people or patients recovering from physical injury. This issue is also present in sport facilities due to high energy placed into activities. The robots were given a dynamic stance, although this requirement has much more relevance on sport facilities than on the senior houses or hospital/clinical environments.

The features to avoid include mechanical type shapes, which are more aggressive, and not appealing for the target user groups. Exposed components, salient from the robot body, are potentially dangerous and should be avoided as well. Robots outfit

should inspire users for interacting with them.

4.3 Ergonomics and Human Factors

Lets now consider the ergonomics and user factors, since problems can arise from bad positioning and dimensioning of the robot components.

According to Fig. 5, we opted for: i) a non-threatening overall stance; ii) no arms; iii) touch screen for better conveying user feedback, as well as to provide a simpler user interface for elderly people.

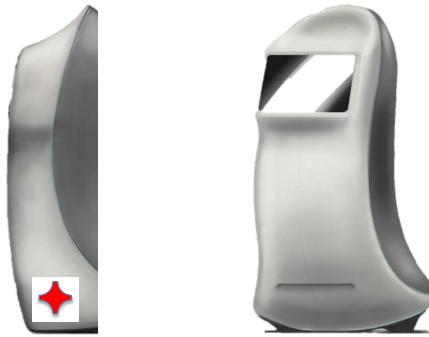


Fig. 5. Mobile robot platform – the nurse.

Similarly to the Monarch robots (see Fig. 6), these should have a modern and cool appearance. And since they operate in a hospital setting, there is a preference for lighter colors (such as white color, common in hospitals), since it is easier to identify dirty surfaces on robot and to clean them. Contrary to the Monarch robots, there is no multifunctional face on the nurse robots.



Fig. 6. Mobile robot platform for children at hospitals, developed under the Monarch project [15]. In contrast to the professional look and feel for the nurse robot, the monarch robots are character-like, but not resembling any particular known artificial character.

5 Assistance Robotics for Augmented Health

Novel approaches have taken Augmented Reality (AR) beyond traditional body-worn or hand-held displays, pushing such capabilities into mobile robots [51]. Spatial Augmented Reality (SAR) uses digital projectors to render virtual objects onto 3D

objects located in the robot's navigation environment. When mounting digital projectors on robots, this collaboration paves the way for unique Human-Robot Interactions (HRI) that otherwise would not be possible. This is especially useful for supporting patients performing recovery exercises, or for providing additional information to elderly people (e.g. time for medicine intake).

5.1 Augmented Reality for Robotic Assistants

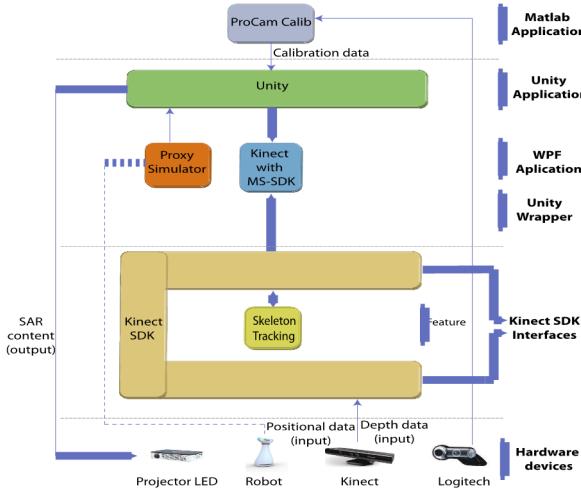


Fig. 7. Architecture of the proposed solution. The line in dash represents the connection to the mobile robot localization system.

The architecture for SAR enabled robots is represented in Fig. 7. It is comprised of six main categories:

- The hardware includes a projector, a camera, and a Kinect depth camera for human-robot interactions;
- The Kinect SDK interface, which allows direct access to Kinects RGB and depth sensors. One of Kinects' key features we used was skeleton tracking;
- Unity wrapper, which enables the use of Kinect SDK internal functions within Unity;
- A Windows Presentation Foundation application (WPF) was developed in order to simulate positional information retrieved from a robot localization system;
- Unity, the game engine on top of which the main applications were developed;
- The camera-projector calibration application, which makes use of ProCamCalib, developed on Matlab.

A collection of modules (given by a series of scripts in Unity) were designed to perform specific tasks:

- Homography matrix update: updates the matrix associated with each projection scene and apply that transformation in OpenGL's vertex pipeline;
- Tracking: Controlling each projection surface position in the virtual scene;
- Intrinsic Setting: Alters Unity's default projection matrix so that the values obtained from the projection calibration process can be properly applied;

- Update Unity's camera position: based on localization (position/pose) information provided by an exterior application (on the robot or elsewhere);
- Save and load: of projection surface positional data from an XML file;
- Human-Robot Virtual Interface (HRVI): Updates the game logic of an interactive AR game based on input received from Kinect's skeleton tracking information.

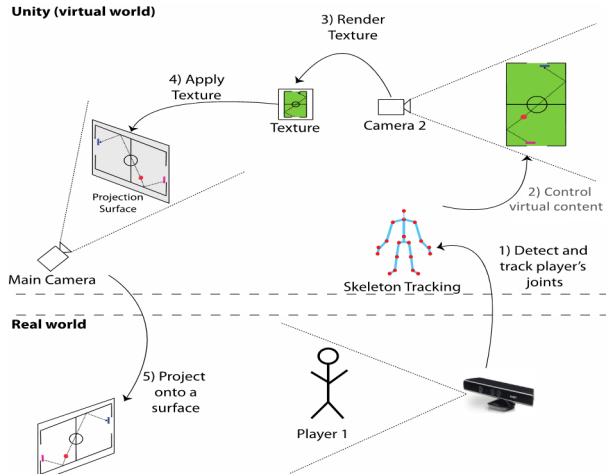


Fig. 8. Human interactions. One or two humans are tracked simultaneously using Kinect's skeleton tracking (step 1). The corresponding joint, for each player, controls the respective virtual object (step 2). Camera 2 renders the texture containing the view of the camera (step 3) that is then applied to one of the projection surfaces and updated in runtime (step 4). The final step can now take place: project onto one real world surface (step 5).

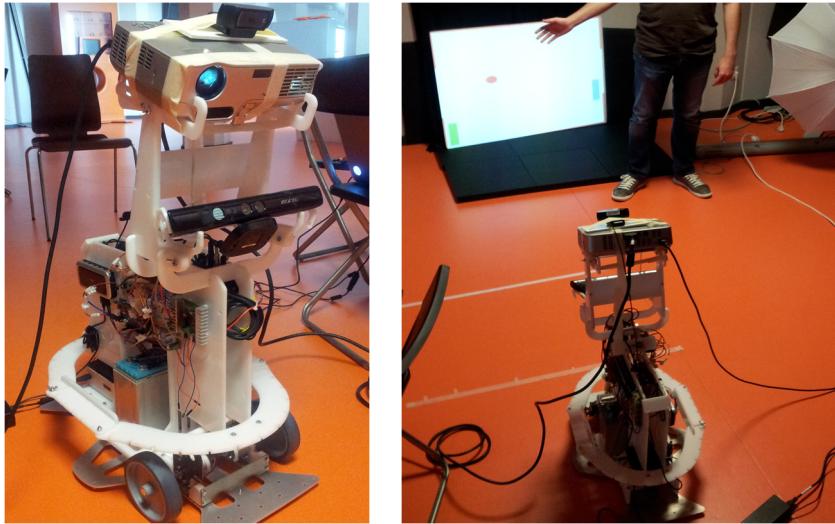


Fig. 9. Augmented reality on Wheels – Content is projected onto multiple surfaces in order to help user perform actions, or to play games with users (such as the Air Hockey game, as shown in the right image).

The goal is to project augmented reality content according to objects physics, or other entities. As such, we introduce the human interactions feedback in order to adapt the augmented reality content, and to achieve virtual-real objects integrated behavior, as represented in Fig. 8.

Hence, the application integrates augmented reality information, with projection distortion compensation, and human gestures recognition to enable interactions between people and undistorted augmented reality content. Virtual content is projected onto the real world where humans interact with such content through movements (see Fig. 9), as detected by Microsoft Kinect's skeleton tracking feature.

5.2 Haptic Feedback

The AHA project is also investigating the usage of robotic elements, behind the nurse mobile robot. Besides biosignal sensors on the human body (see Fig. 10a), actuators are being investigated to provide real-time feedback concerning movements to be executed, or corrections to imperfections on the execution of the exercises. Such actuators include an elastic sleeve (see Fig. 10b), enabled with a sensor to measure joint rotation and an array of motors to provide haptic feedback.

Currently two sensors have been tested: a Flex Sensor (see Fig. 11), a flexible strip that senses rotation, as well inertial sensor based on gyroscope and accelerometer. Other devices are planned to be tested, such as mobile phone gadgets (like Pebble), that are able to generate controlled vibrations.

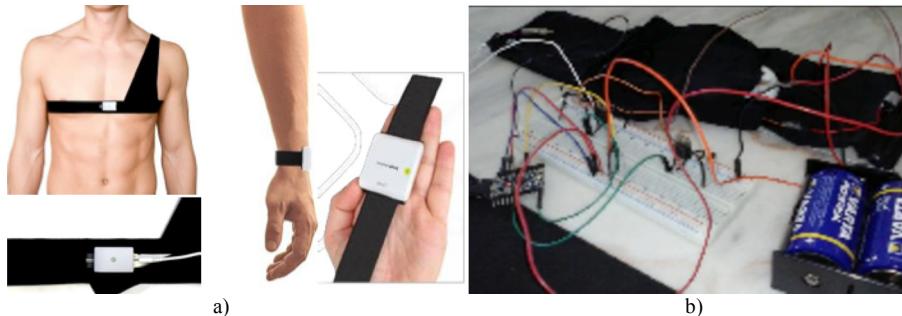


Fig. 10. a) Biosignal sensing devices, to acquire data concerning biological variables on the human body; b) Elastic sleeve prototype with a SensorFlex and two vibration motors for haptic feedback.

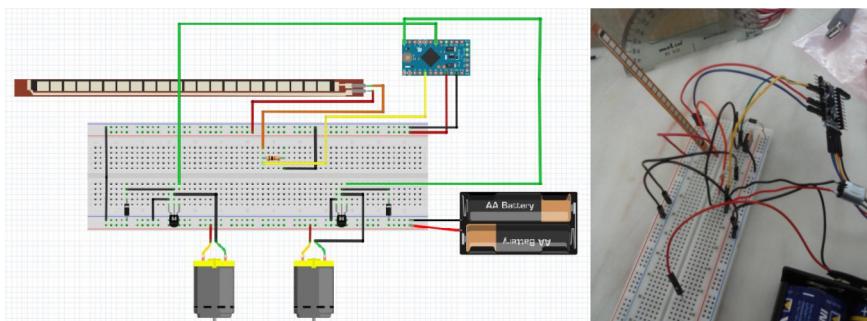


Fig. 11. Prototype schematics and testing, using a Flex sensor and two motors.

6 Conclusions and Discussion

This chapter presented some of the robotics technologies being developed for Augmented Health Assistance, under the scope of the AHA project.

It has been shown that one of the problems with remote rehabilitation is the correct execution of recovery exercises by the patient. Furthermore, due to the repetitive nature of such movements, there is consequently a motivation problem for keeping the patient engaged with the exercise program. The first issue can be addressed by the usage of augmented reality and haptic feedback for providing corrective feedback to the patient. The later issue can be addressed employing nurse robots. These replace the role of the human therapeutic on providing advice and motivation. The nurse robot also plays a very important role on supporting elderly people on their daily activities. For instance, these robots can take medicine to elderly at intake times, or make available medical equipment.

The AHA project proposed to develop new algorithms, and virtual as well as robotic systems for eHealth technologies. It expects to have a significant impact on the health and sport sectors, by employing interdisciplinary technologies at the service of elderly, athletes and patients with physical injury (both at home or on a clinical/hospital environment). Innovative approaches will bring benefits to society, by reducing negative effects of sedentarism, such as related diseases and aging conditions. It proposes as well to improve the quality of life of elderly people. Besides developing a mobile robot acting as a companion nurse to elderly (reducing the problem of elderly isolation and loneliness), it also aims to increase elderly physical fitness, independence from human nurses and family, and consequently improving their autonomy and self-esteem.

Concerning patients in recovery of physical injuries, the AHA project proposes to improve these patients quality of life. This is accomplished by enabling them to perform independently exercises at home, while receiving corrective feedback and motivational incentive to adhere to scheduled exercises.

The project promotes however to keep clinicians in the loop (even remotely) to follow patient progress. The clinician is also able to customize patients therapies, as well as to provide of-line, or real-time, corrections to therapies or exercise execution.

The combination of a large set of interdisciplinary technologies is expected to enable a new level of health assistance to different user groups, improving their quality of life and reducing government economic burden.

Acknowledgements. This work has been funded by CMU-Portuguese program through Fundação para a Ciência e Tecnologia, project AHA-Augmented Human Assistance, AHA, CMUP-ERI/HCI/0046/2013.

References

1. Arsenio, A., Andrade, J., Duarte, A.: Non-invasive Wireless Bio Sensing. Accepted for publication at Lecture Notes in Computer Science, series Communications in Computer and Information Science (CCIS), Springer-Verlag (2015)
2. Sigrist, R., Rauter, G., Riener, R., Wolf, P.: Augmented Visual, Auditory, Haptic, and multimodal feedback in motor learning: A review. *Psychonomic bulletin & review* 20, 1

- (Feb. 2013), 21–53
3. Tang, R., Alizadeh, H., Tang, A., Bateman, S., Jorge, J. Physio@Home: design explorations to support movement guidance. CHI '14 Extended Abstracts on Human Factors in Computing Systems (2014)
 4. World health statistics 2012. World Health Organization, 2012
 5. Bemelmans, R., Gelderblom, G., Jonker, P., de Witte, L.: Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *Journal of the American Medical Directors Association* 13.2 (2012), 114-120
 6. Broekens, J., Heerink, M., Rosendal, H.: Assistive social robots in elderly care: a review. *Gerontechnology* 8.2 (2009), 94-103
 7. Matarić M., Eriksson J., Feil-Seifer D.: Socially assistive robotics for post-stroke rehabilitation. *Journal of Neuro Engineering and Rehabilitation* 4.5 (2007)
 8. Scassellati, B.: How social robots will help us to diagnose, treat, and understand autism. *Robotics research*. Springer Berlin Heidelberg, (2007), 552-563
 9. Cory, K., Breazeal, C.: Robots at home: Understanding long-term human-robot interaction. *IEEE/RSJ IROS International Conference on Intelligent Robots and Systems*, IEEE (2008)
 10. Pineau, J., Montemerlo, M., Pollack, M., Roy, N., Thrun, S.: Towards robotic assistants in nursing homes: Challenges and results. Special issue on Socially Interactive Robots, Robotics and Autonomous Systems 42.3 (2003), 271-281
 11. Rincon, F., Vibbert, M., Childs, V., Fry, R., Caliguri, D., Urtecho, J., Rosenwasser, R., Jallo, J.: Implementation of a model of robotic tele-presence (RTP) in the Neuro-ICU: effect on critical care nursing team satisfaction. *Neurocritical care* 17.1 (2012), 97-101
 12. Wada, K., Shibata, T.: Robot Therapy for Elders Affected by Dementia. *Engineering in medicine and biology magazine*, 27 (2008), 53–60
 13. Wong, B., Norman, D.: Evaluation of a Novel Medication Aid, the Calendar Blister-pak, and its effect on drug compliance in a geriatric outpatient clinic. *Journal of the American Geriatrics Society* (1987)
 14. Castellano, G., Paiva, A., Kappas, A., Aylett, R., Hastie, H., Barendregt, W., Nabais, F., Bull, S.: Towards Virtual and Robotic Empathic Tutors. 16th International Conference on Artificial Intelligence in Education (AIED'13), Memphis, USA. In H.C. Lane & K. Yacef (eds), *Artificial Intelligence in Education* Springer-Verlag, Berlin Heidelberg (July 2013)
 15. Gonçalves, D., Arsenio, A.: Human-driven multi-robot design process for social interactions with children on complex environments. 2015 6th International Conference on Automation, Robotics and Applications (Feb. 2015), 70–76
 16. Pinto, A., Almeida, J., Pinto, S., Pereira, J., Oliveira, A., de Carvalho, M.: Home telemonitoring of non-invasive ventilation decreases healthcare utilisation in a prospective controlled trial of patients with amyotrophic lateral sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry* (2010)
 17. Lim, J., Zitnick, C., Dollár, P.: Sketch Tokens: A Learned Mid-level Representation for Contour and Object Detection. *CVPR*, Portland, Oregon (2013)
 18. Girshick, R., Felzenszwalb, P., Mcallester, D.: Object detection with grammar models." *Advances in Neural Information Processing Systems* (2011)
 19. Arsenio, A.: Cognitive-developmental learning for a humanoid robot: a caregiver's gift. MIT. PhD thesis (2004)
 20. Yang, Y., Ge, S., Lee, T., Wang, C.: Facial expression recognition and tracking for intelligent human-robot interaction. *Intelligent Service Robotics* 1.2 (2008), 143-157
 21. Saponaro, G., Salvi, G., Bernardino, A.: Robot Anticipation of Human Intentions through Continuous Gesture Recognition, CR-HRI (2013)
 22. Goffredo, M., Seely, R., Carter, J., Nixon, M.: Markerless view independent gait analysis with self-camera calibration, In Proceedings of International Conference on Automatic Face and Gesture Recognition (2008)
 23. Chang, C., Lange, B., Zhang, M.: Towards Pervasive Physical Rehabilitation using Microsoft Kinect. 6th International Conference on Pervasive Computing Technologies for

- Healthcare (PervasiveHealth) (2012), 159–162
24. Scano, A., Caimmi, M., Malosio, M., Tosatti, L.: Using Kinect for upper-limb functional evaluation in home rehabilitation: A comparison with a 3D stereoscopic passive marker system. 5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics (Aug. 2014), 561–566
 25. Borghese, N., Mainetti, R., Pirovano, M., Lanzi, P.: An intelligent game engine for the at-home rehabilitation of stroke patients. IEEE 2nd International Conference on Serious Games and Applications for Health (SeGAH) (May 2013), 1–8
 26. Kitsunezaki, N., Adachi, E., Masuda, T., Mizusawa, J.: KINECT applications for the physical rehabilitation. IEEE International Symposium on Medical Measurements and Applications (MeMeA) (May 2013), 294–299
 27. Gama, A., Chaves, T., Figueiredo, L., Teichrieb, V.: Guidance and Movement Correction Based on Therapeutics Movements for Motor Rehabilitation Support Systems. 14th Symposium on Virtual and Augmented Reality (May 2012), 191–200
 28. Burke, J., McNeill, M., Charles, D., Morrow, P., Crosbie, J., McDonough, S.: Serious Games for Upper Limb Rehabilitation Following Stroke. Conference in Games and Virtual Worlds for Serious Applications (Mar. 2009), 103–110
 29. Sadikhov, D., Migge, B., Gassert, R.: Prototype of a VR upper-limb re-habilitation system enhanced with motion-based tactile feedback. World Haptics Conference (WHC) (Apr. 2013), 449–454
 30. Schonauer, C., Pintaric, T.: Full Body Interaction for Serious Games in Motor Rehabilitation. Proceedings of the 2nd Augmented Human International Conference, ACM Press (2011), 1–8
 31. Rego, P., Moreira, P., Reis, L.: Serious games for rehabilitation: A survey and a classification towards a taxonomy. 5th Iberian Conference on Information Systems and Technologies (CISTI 2010).
 32. Adherence to long-term therapies: evidence for action. World Health Organization (2003)
 33. Huang, H., Ingalls, T., Olson, L., Ganley, K., Rikakis, T., He, J.: Interactive multimodal biofeedback for task-oriented neural rehabilitation. 27th Annual International Conference of the Engineering in Medicine and Biology Society, IEEE-EMBS (2005), 2547–2550
 34. Garn, A., Baker, B., Beasley, E., Solmon, M.: What are the benefits of a commercial exergaming platform for college students? Examining physical activity, enjoyment, and future intentions. Journal of physical activity & health 9.2 (2012), 311–8
 35. Staiano, A., Calvert, S.: Exergames for physical education courses: Physical, social, and cognitive benefits. Child Development Perspectives 5.2 (2011), 93–98
 36. Singh, A., Klapper, A., Jia, J.: Motivating people with chronic pain to do physical activity: opportunities for technology design. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2014), 2803–2812
 37. Lohse, K., Shirzad, N., Verster, A.: Video Games and Rehabilitation: Using Design Principles to Enhance Engagement in Physical Therapy. Journal of Neurologic Physical Therapy 37, 4 (Dec. 2013), 166–75
 38. Laver, K., George, S., Thomas, S., Deutsch, J., Crotty, M.: Cochrane review: virtual reality for stroke rehabilitation,” Eur J Phys Rehabil Med, 48.3 (Sep. 2012), 523–530
 39. Cameirão, M., Bermúdez, S., Verschure, P.: Virtual reality based upper extremity rehabilitation following stroke: A review. Journal of CyberTherapy & Rehabilitation, 1.1 (2008)
 40. Lucca, L.: Virtual reality and motor rehabilitation of the upper limb after stroke: a generation of progress?. Journal Rehabilitation Medicine, 41.12 (Nov. 2009), 1003–1100
 41. Raskara, R., Ziegler, R., Willwacher, T.: Cartoon dioramas in motion. Proceedings of the 2nd international symposium on non-photorealistic animation and rendering (2002), 7-13
 42. Bandyopadhyay, D., Raskar, R., Fuchs, H.: Dynamic Shader Lamps: Painting on Movable Objects. IEEE/ACM international symposium on mixed and augmented reality (2001)
 43. Raskar, R., van Baar, J., Beardsley, P., Willwacher, T., Rao, S., Forlines, C.: iLamps: Geometrically Aware and Self-Configuring Projectors. ACM Transactions on Graphics (TOG), ISSN: 0730-0301, 22.3 (July 2003), 809-818

44. Wilson, A., Benko, H.: Combining Multiple Depth Cameras and Projectors for Interactions On, Above, and Between Surfaces. In Proceedings of ACM UIST '10 (2010), 273-282
45. Yang, R., Gotz, D., Hensley, J., Towles, H., Brown, M.: Pixelflex: A reconfigurable multi-projector display system. Proceedings of the conference on Visualization (2001), 167-174
46. Azuma, K., Miyashita, S.: Anamorphosis Projection by Ubiquitous Display in Intelligent Space. Proc. int. conf. universal access in human-computer interaction (2009), 209–217
47. Harris, C.: Tracking with rigid models. In Active vision, MIT Press Cambridge, MA, USA (1993), 59-73
48. Schkolne, S., Pruett, M., Schröder, P.: Surface drawing: creating organic 3D shapes with the hand and tangible tools. Proceedings of SIGCHI Conference on Human Factors in Computing Systems (2001), 261–268
49. Jung, H., Nam, T., Lee, H., Han, S.: Spray modeling: Augmented reality based 3D modeling interface for intuitive and evolutionary form development. Proceedings of the International Conference on Artificial Reality and Telexistence (2004)
50. Sodhi, R., Benko, H., Wilson, A.: LightGuide: projected visualizations for hand movement guidance. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2012)
51. Costa, N., Arsenio A.: Augmented Reality behind the wheel - Human Interactive Assistance by Mobile Robots. 6th International Conference on Automation, Robotics and Applications (Feb. 2015), 63 - 69

MINERVA Project, mid- To near Infrared Spectroscopy for Improved Medical Diagnostics

Valery Naranjo¹, Francisco Peñaranda¹, Mariano Alcañiz¹, Bruce Napier²,
Mark Farries³, Gary Stevens³, John Ward⁴, Cestmir Barta⁵, Radek Hasal⁵,
Angela Seddon⁶, Slawomir Sujecki⁶, Samir Lamrimi⁷, Uffe Møller⁸, Ole Bang⁸,
Peter M. Moselund⁹, Munir Abdalla¹⁰, Danny De Gaspari¹⁰, Rosa M. Vinella¹⁰,
Hedda Malm¹¹, Gavin R. Lloyd¹², Nick Stone¹³, Jayakrupakar Nallala¹³,
Juergen Schnekenburger¹⁴, Lena Kastl¹⁴ and Björn Kemper¹⁴

¹ Inst. Int. de Investigación en Bioingeniería y Tecnología Orientada al Ser Humano,
Universitat Politècnica de Valencia, Valencia, Spain

² Vivid Components, Dr.-Rörig-Damm 22, 33102, Paderborn, Germany

³ Gooch & Housego (Torquay) Ltd., Broomhill Way, Torquay, Devon, TQ2 7QL, U.K.

⁴ Gooch & Housego (UK) Ltd., Dowlish Ford, Ilminster, Somerset, TA19 0PF, U.K.

⁵ BBT-Materials Processing SRO, Doubicka 11, Praha 8, 184 00, Czech Republic

⁶ George Green Institute for Electromagnetics Research, Faculty of Engineering,
University Park, University of Nottingham, Nottingham, NG7 2RD, U.K.

⁷ LISA Laser Products OHG, Fuhrberg & Teichmann Max-Planck-Str. 1,
37191, Katlenburg-Lindau, Germany

⁸ DTU Fotonik, Department of Photonics Engineering, Technical University of Denmark,
2800 Kgs. Lyngby, Denmark

⁹ NKT Photonics A/S, Blokken 84, 3460, Birkerød, Denmark

¹⁰ Xenics NV, Ambachtelaan 44, BE-3001, Leuven, Belgium

¹¹ IR Nova AB, Electrum 236, 164 40, Kista, Sweden

¹² Biophotonics Research Unit, Gloucestershire Hospitals NHS Foundation Trust,
Gloucester, U.K.

¹³ Department of Physics, Exeter University, Exeter, U.K.

¹⁴ Biomedical Technology Center, University of Muenster, D-48149, Muenster, Germany
vnaranjo@labhuman.com

Abstract. The main idea behind the MINERVA project is the recognition that for the first time, through breakthroughs in photonic technology, it is possible to open the mid-IR electromagnetic spectrum (3-12 μm) for rapid medical imaging. In particular this could greatly improve the chances of early cancer diagnosis. MINERVA will exploit and develop the advances in soft glass optical fibres, acousto-optic (AO) modulator design, crystal growth, fibre lasers, supercontinuum sources and detectors in the mid-IR. Two specific high impact applications will be addressed: high volume pathology screening (i.e. automated microscope-based examination of routine patient samples) and human skin surface examination (i.e. non-invasive investigation of suspected skin cancer). In an Integrating Project of this scale it is possible to pursue several targets in parallel, each of which alone brings significant benefits. Together they could begin a new branch of the bio-medical imaging industry.

1 Introduction

MINERVA is a project funded by the European Commission through its Seventh Framework Programme (FP7) [1]. It brings together thirteen partners from across Europe with the common objective of developing mid-infrared (mid-IR) technology to improve the early diagnosis of cancer (Fig. 1).

Mid-IR radiation is an exciting new area for real-time molecular-sensing with applications in different areas: medicine and healthcare (e.g. early cancer detection: the MINERVA application space), environment and energy (e.g. monitoring exhaust gases), security (e.g. detection of narcotics or explosives, food security), chemical and industrial manufacturing (e.g. process control and quality assurance).

The MINERVA mid-IR range (1.5 to 12 μm) is rich in spectroscopic absorption peaks of biomolecules such as fats, proteins and carbohydrates. In particular it has been shown that, through the latest data analysis techniques, this spectral region can be used to identify the presence of early cancer. Currently there is a lack of practical sources and components for this spectral region, and so these mid-IR diagnostic techniques are restricted to laboratory demonstrations.

MINERVA aims to develop fibre, lasers and broadband sources, components, modulators and detectors to access this important part of the spectrum. In parallel it will identify analytical techniques using the new photonic hardware to improve early skin cancer diagnosis and the rapid and automatic assessment of biopsy samples using a microscope.



Fig. 1. Logos of the thirteen partners of MINERVA's consortium.

1.1 mid-IR Spectroscopy: A New Tool for Pathologists

The spectral region studied in MINERVA (1.5-12 μm) includes the so-called “fingerprint region” in which many biomolecules have tell-tale absorption peaks. By studying the pattern of absorbed radiation it is possible to deduce details of the type and distribution of these molecules, which in turn provides important information for disease diagnosis.

It is emphasised that this process is not as straightforward as simply spotting certain chemicals, or “cancer markers”. The information is buried in the inter-related distribution of species and subtle biochemical changes. It requires a powerful mathematical

technique known as multi-variate analysis to extract useful information from the reams of spectral data in order to spot the warning signs of cancer.

One form of multi-variate analysis is correlation mapping, which enables the visualisation of diseased cells or regions from spectral data (Fig. 2). MINERVA combines novel mid-IR spectroscopy with correlation mapping and hopes to lead to a breakthrough diagnostic technology.

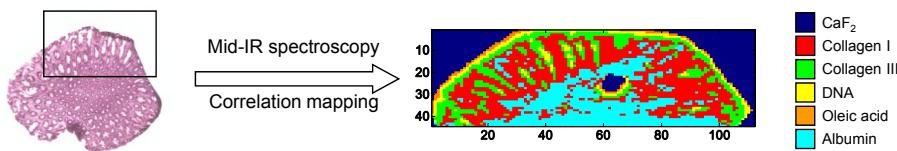


Fig. 2. Correlation mapping enables the visualisation of diseased cells or regions from spectra.

MINERVA will develop a suite of mid-IR photonic hardware to improve access to this information. Working in the mid-IR is extremely challenging, and MINERVA will need to break new ground in several technical areas:

- *Gooch & Housego* (G&H), the project coordinator, will develop mid-IR components such as fused combiners (glass devices used to combine or separate signals into different optical fibres), and acousto-optic (AO) modulators (to switch the signals and separate wavelengths at high speed).
- These AO devices will need new types and sizes of calomel crystals from *BBT Materials Processing SRO* (BBT).
- Mid-IR glass fibre to carry the radiation efficiently and conveniently is being produced at *University of Nottingham* (NOTT).
- Novel pump lasers at $2.9\text{ }\mu\text{m}$ and $4.5\text{ }\mu\text{m}$ from *LISA Laser Products OHG* (LISA) will be used by the *Technical University of Denmark* (DTU) and *NKT Photonics A/S* (NKT) to generate a range of supercontinuum sources in ZBLAN, indium fluoride and chalcogenide glasses, spanning the MINERVA range from $1.5\text{ }\mu\text{m}$ to $12\text{ }\mu\text{m}$.
- *Xenics* and *IRnova* are advancing the state-of-the-art in Type II superlattice detectors, which offer a cost effective route to highly efficient detection in the mid-IR.
- *University of Exeter* and *Gloucestershire Hospitals NHS Trust* (GHNT) will develop the multivariate algorithms and techniques for high volume screening of human samples.
- *Westfaelische Wilhelms-Universitaet Muenster* (WWU) will develop a skin cancer diagnostic process.
- *Universitat Politècnica de València* (UPV) is working on novel algorithms for the analysis of histopathological images and the recognition and classification of hyperspectral data of cancer samples.
- The project is managed and administrated by *Vivid Components*.

In the next sections it will be presented an overview of the expectancies of the project and the main preliminary advances reached by the different groups.

2 Mercurous Halides: Unique Acousto-Optic Materials for IR from BBT

BBT is a world leader in the growth and processing of Mercurous Chloride (Hg_2Cl_2 , Calomel) single crystals with excellent AO properties and is thus in favourable position to address this question. Calomel single crystals (Fig. 3.a) exhibit a wide range of optical transmission, high indices of refraction, extremely high value of acousto-optic figure of merit M2, very low velocity of shear acoustic wave, high value of birefringence (four times higher than Calcite), etc. The Calomel crystals are well adapted to fabricate acousto-optic devices operating in the mid and far IR (3 to 20 μm).

At the beginning of the MINERVA project the production technology enabled the growth of calomel crystal boules with a diameter of 26 to 29 mm (Fig. 3.b) and length of 45 to 60 mm (typically 55 mm). Within the MINERVA project the new technology is being developed enabling the growth of cylindrical crystal boules with a diameter up to 35 mm, which is necessary for the further manufacturing of acousto-optical tuneable filters of new design proposed by G&H.

The Calomel crystal growth process is highly demanding, difficult and complex, especially in case of bigger 35mm boules. The growing process is powered by a dynamic temperature field and corresponding axial and radial temperature gradients. The whole process has to be carefully maintained within narrow physical condition limits. All the equipment and accessories have to be newly developed by BBT and adjusted to the specific conditions for growing of the 36mm diameter crystals including the temperature controllers. These controllers are equipped with brand new cultivation programs with respect to the bigger material mass. A total of six crystallizers will be built within the project. Currently, four units are operational and tested (Fig. 3.c).

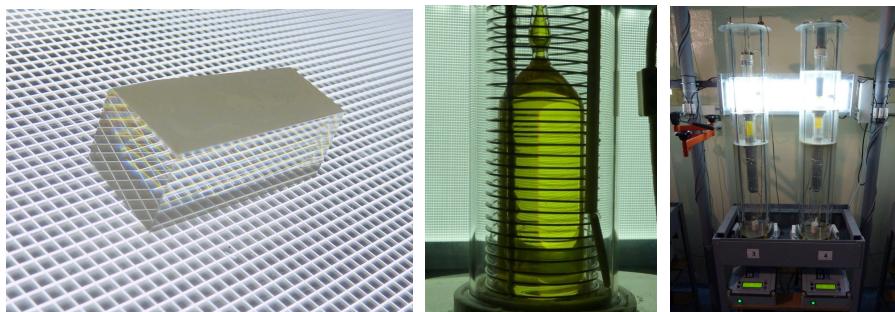


Fig. 3. (a) Polished Calomel AOTF substrate. (b) Growing Calomel crystal, diameter 28mm. (c) Two cultivation crystallizer units with Calomel crystals.

3 Er:ZBLAN Fibre Laser at 2.9 μm from University of Nottingham and LISA

The partners NOTT and LISA will develop a 2.9 μm laser based on Er-doped ZBLAN fibres diode-pumped at 976 nm. This fibre laser will be used as pump source for ultra-long wavelength supercontinuum generation (3-9 μm). The first step is the development of a fibre laser in an external cavity configuration. For that purpose a simulation model based on the rate equation and signal propagation equations is implemented by the NOTT group (Fig. 4). Different parameters will be studied, e.g. absorption cross-sections, emission cross-sections, and gain cross-sections, to predict the optimum laser performance.

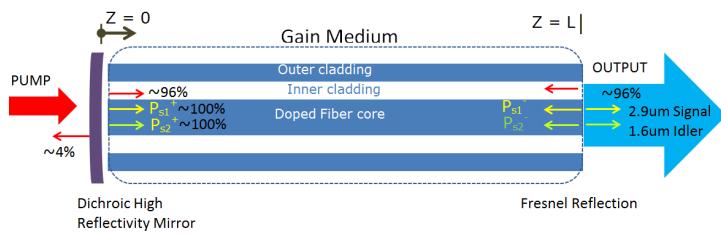


Fig. 4. Modelling scheme of the Er:ZBLAN fibre laser with an external cavity configuration. For the exact prediction both the 2.9 μm and the 1.6 μm laser signal were analysed in forward and backward propagation.

In parallel, LISA will carry out experiments for the handling (stripping, cleaving, splicing) of the soft glass fibre and target both high-power and high-energy laser operation with different resonator configurations (Fig. 5).

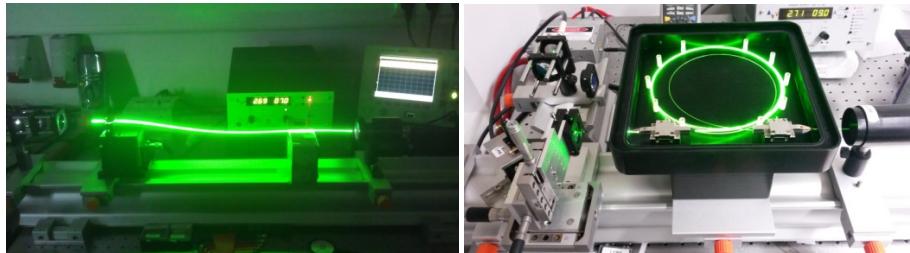


Fig. 5. (a) Set-up of the Er:ZBLAN fibre for absorption studies. The green fluorescence results from up-conversion processes. (b) Set-up of the Er:ZBLAN fibre in an external cavity configuration pumped with high-power fibre-coupled diodes. First experiments showed a good agreement with the simulations carried out at Nottingham.

Coated focussing and collimating optics have to be designed and manufactured for the laser studies. After the evaluation of the first tests in CW operation LISA's scientists and engineers will design a compact and robust cooled housing for the 2.9 μm laser.

Regarding high-energy operation special acousto-optic modulators (AOM) based on TeO₂ will be designed and built by G&H and delivered to LISA.

Further information about MINERVA's fibre laser can be found in [2–10].

4 Extreme IR Supercontinuum Modelling at DTU

DTUs team has the task of fibre modelling in MINERVA in close collaboration with the fibre manufacturing group at NOTT. The DTU group also models dynamic supercontinuum generation along the fibres using both measured material data and calculated fibre properties. This advanced modelling requires extensive computational resources in order to accurately follow the rapid spectral broadening, which covers over four octaves (from 1 μm to 16 μm); made possible by the strong non-linearity of chalcogenide glasses and the extremely high numerical aperture (NA) of the Nottingham fibres. Figure 6 shows two graphics with some results of the numerical modeling of mid-IR supercontinuum generation. Thorough analysis of the modelling has been presented in [11–13].

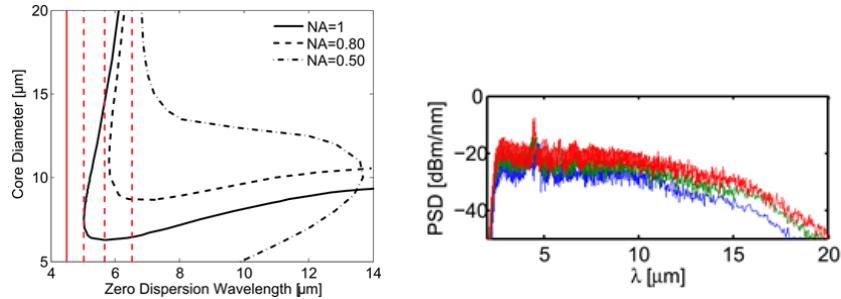


Fig. 6. (a) Zero dispersion wavelengths versus core diameter for step-index fibres (based on fibres fabricated at the University of Nottingham) with NA as given in the legend. (b) Modelling shows that a fibre with core diameter 10 μm and NA = 1.0 exhibiting no second zero dispersion is optimum. Super-continuum generation beyond 12 μm is observed numerically.

5 MINERVA Supercontinuum Sources from NKT

It has been mentioned that the mid-IR region contains a wealth of spectral data which can yield important information on the chemical composition of samples from gases and liquids to living cells. However, the investigation of this topic has been limited by the available photonic sources. Researchers had to choose between a very low intensity broadband source such as a “globar” thermal source, or a high intensity but narrowband source, such as a laser diode.

NKT Photonics is dedicated to providing flexible sources of high intensity light in an easy to use format (Fig. 7.a). It has already established commercial supercontinuum systems which can deliver any wavelength from 400 to 2000 nm on demand. It has recently launched the EXTEND-UV accessory which can extend the wavelength

coverage to cover the 270-400 nm UV region. The company would now like to push the limits of supercontinuum sources at longer wavelengths, reaching into the mid-IR region.

In MINERVA NKT is developing zirconium fluoride (ZrF_4) glass fibre supercontinuum sources to cover the 1.5-4.5 μm spectrum. Subsequently it will investigate even longer wavelengths by utilising newly developed indium fluoride (InF_3) fibres to extend the spectrum beyond the transmission band of ZrF_4 glasses.

These sources could detect changes in cells by monitoring absorption in the 2.6-3.8 μm region which relates to the balance between lipids and proteins (Fig. 8.a). The increase in wavelength from 4.5 μm up to $>5 \mu\text{m}$ would make it possible to interrogate additional important gas absorption lines such as carbon monoxide.

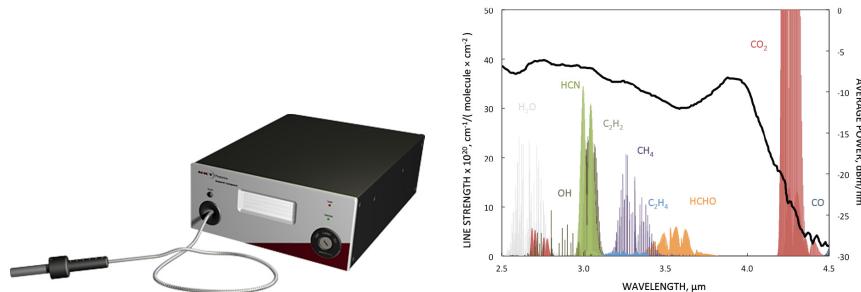


Fig. 7. (a) Schematic of a typical NKT supercontinuum source. (b) Graph showing absorption spectra of some key bio-molecules in the lower end of the “fingerprint region”.

In the first 2 years of the MINERVA project NKT has already developed several supercontinuum sources with output power of up to 2.5 W. These sources are more than a million times brighter than most thermal light sources and even brighter than a Synchrotron. We have shown the limits of zirconium-fluoride based systems by setting a new record for the longest wavelength supercontinuum generated at 4.75 μm . However, the chemometric specialists in MINERVA found that the main region of interest was the 2.5-3.8 μm region so we have also shown how the main power in the output spectrum can be shifted down to the main region of interest by altering the design of the nonlinear fiber. These first Supercontinuum sources are already at work in the development of the MINERVA-lite microscopy setup which will soon be applied to bio sample imaging.

Meanwhile NKT is pushing onward in the development of the mid-IR supercontinuum sources. The initial sources were based on rather long pulse nanosecond pump lasers with relatively low pulse repetition frequencies. This made the sources incompatible with most of the Fourier transform spectrometers (FTIRs) that many researchers use in the mid-IR region. In addition, the low repetition rate made it time consuming to counter any noise in the source by averaging over many pulses. NKT is therefore now developing sources based on much shorter pump pulses and with higher repetition rate in order to reduce noise and make the sources compatible with standard FTIRs.

As these mid-IR supercontinuum sources become available and known in the field, the MINERVA consortium expects the emergence of new markets. For example, an

important spectroscopic application in the petrochemical industry is to monitor single wavelengths in the 3-3.5 μm band in order to optimise the refining processes. Monitoring the whole spectrum simultaneously would allow a full real-time chemical analysis of the output chemicals.

Some relevant references concerning the supercontinuum sources within the MINERVA framework have been already published [14–24]

6 MINERVA type-II Superlattice IR Detectors from IRnova

Type-II superlattice (T2SL) is a material/technology that can be used for high quality cooled photon detectors, with tailorabile bandgap from 2 μm and upwards. The name comes from the fact that the conduction and valence bands display a so-called “broken type-II” (sometimes also called “type-IIb” or even “type-III”) alignment between the constituent materials, which can be InAs/GaSb/AlSb, or alloys thereof (Fig. 8.a). In contrast to typical quantum well devices, e.g. the active regions of semiconductor lasers, the superlattice layers in the T2SL material are so thin (typically 3 nm) that mini-bands are formed in the material. These mini-bands resemble the conduction and valence bands of a bulk semiconductor material. By carefully selecting the superlattice layer thicknesses and compositions, novel materials can be defined to meet widely different needs.

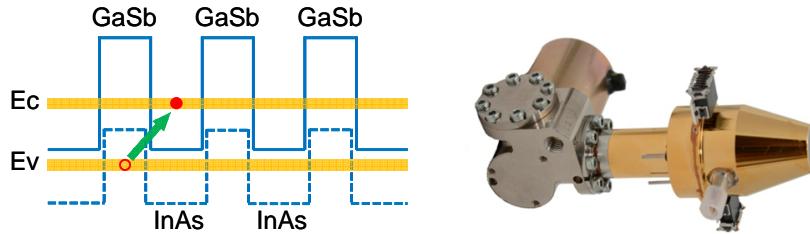


Fig. 8. (a) Schematic of T2SL band-gap structure. (b) Detector/Dewar/cooler assembly for T2SL from IRnova.

Compared with a traditional bulk material for the 3-5 μm range, such as InSb, T2SL requires less cooling and thus draws less power, which allows for longer cooler lifetime and consequently lower life-cycle cost. For the 8-12 μm range, the traditional alloy bulk material HgCdTe (or “MCT”) is difficult to fabricate with high yield, partly due to the extreme sensitivity of the bandgap to composition (particularly the HgTe:CdTe alloy ratio). Here T2SL materials have distinct advantages in fabrication.

Focal plane arrays comprising hundreds of thousands of T2SL detector pixels are flip-chip bonded to a CMOS read-out-circuit and then mounted on a ceramic carrier, which in turn is glued to a cold finger in a vacuum Dewar housing, complete with an IR window. The cold finger is cooled to detector operating temperature by a Stirling rotary cooler. IRnovas detector-Dewar-cooler assembly for T2SL can be seen in Fig. 8.b.

IRnova has recently worked on improving the quantum efficiency (QE) of the detection by applying anti-reflective coatings to the detector surface. By this method, the

QE was increased from approximately 50% to 80% in the wavelength region of interest. This improves the signal-to-noise ratio and allows a reduced integration time for each image frame.

Apart from MINERVA applications, IRnova plans to use T2SL technology for gas detection of key greenhouse gases with absorption lines in the atmospheric transmission bands, such as methane and perhaps also sulphur hexafluoride (SF₆).

More information about the superlattice IR detectors can be found in [25]

7 Infrared Megapixel Camera Development at Xenics and IRnova

The sensing unit for MINERVA is being developed in a joint effort between Xenics and IRnova. From the start of the project Xenics has been working on the design of a Read-Out IC (ROIC), to be integrated through flip-chip technology with the T2SL (Type-2 Super Lattice) photodiode material, which is being developed by IRnova.

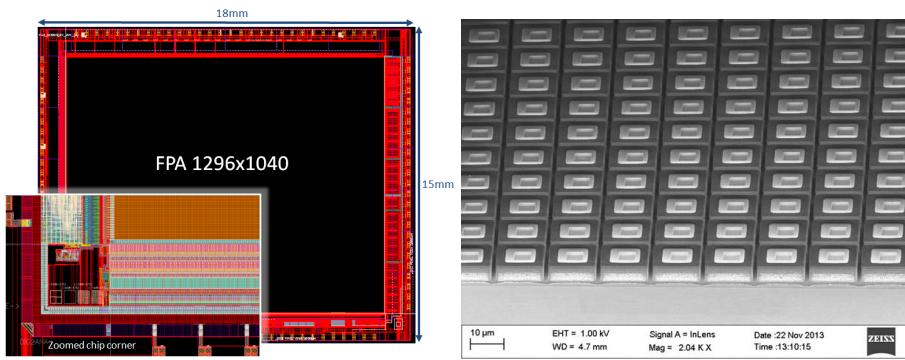


Fig. 9. (a) Global and zoomed view of the designed ROIC, currently in manufacturing. (b) Microscopic view of manufactured array of T2SL photo diodes.

To provide sufficient resolution for reliable data analysis, a 1280×1024 pixel array was chosen, on an aggressive pitch of 12 μm. Project requirements including frame rate, sensitivity and noise level were taken into account in the design process. After extensive simulations and test sample manufacturing, the final design was taped out to a manufacturing foundry (Fig. 9.a). The use of advanced 0.18 μm CMOS technology is required to allow all necessary functionality within the available space of 144 μm² per pixel. The first wafers are currently available for post-processing to be followed by wafer-level verification of the electrical functionality.

In parallel, IRnova has been working on the optimisation of the design and processing of the T2SL material, towards cut-off wavelength matching and dark current minimisation (Fig. 9.b). Once the ROIC chips become available later this year, they will be hybridized to the sensor chip, and IRnova will integrate the resulting hybrid in a Sterling-cooled Dewar. The so-called engine will in its turn be integrated into a full camera by Xenics, to be supplied to the other MINERVA project partners to be used for capturing spectroscopic images of prepared tissue samples and live cell phantoms.

8 Pattern Recognition and Data Analysis at GHNT

The first task of GHNT was to provide supporting evidence for the MINERVA instrument specifications. This was achieved by analysing an existing dataset and applying pattern recognition techniques to discriminate between benign and cancerous samples from human colon tissue biopsies (Fig. 10). Sensitivity and specificity of up 86-99% can be achieved with the existing dataset. Using this study as a baseline GHNT was able to assess the impact of various factors that will affect the quality and speed of the MINERVA instrument.

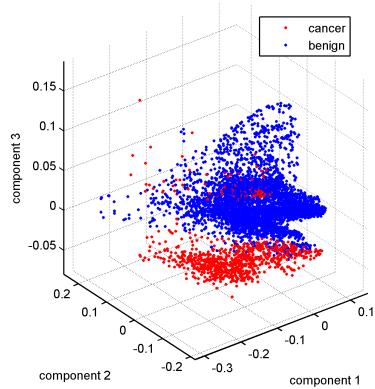


Fig. 10. Partial Least Squares (PLS) scores plot showing the separation between benign and cancer samples in baseline study.

Reducing the number of data points per spectrum is one way to potentially speed up the system; measuring fewer wavenumbers means a faster total acquisition time. Multivariate pattern recognition algorithms were used to identify potential wavenumber targets for the MINERVA instrument. Figure 11 shows the wavenumber regions identified as ‘important’ for the baseline study.

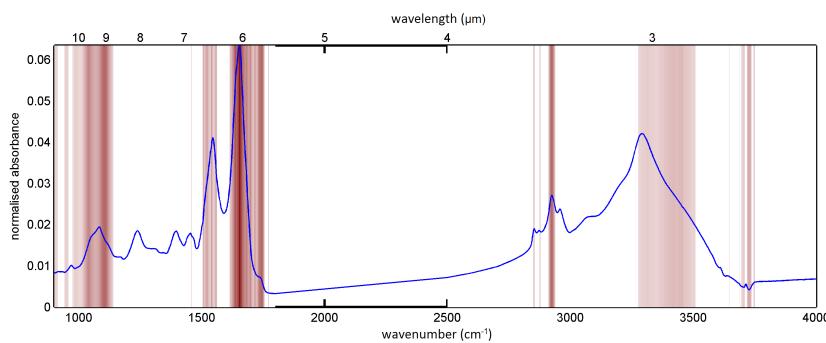


Fig. 11. VIP identified wavenumber targets for the MINERVA system (red). Reference spectrum (blue).

A minimum acquisition time per spectrum means that the MINERVA instrument will be able to rapidly assess samples in a clinical timeframe. However, reducing acquisition time also increases the amount of noise. To determine what level of noise can be tolerated by the pattern recognition algorithms GHNT simulated the addition of noise to the baseline study until it was no longer able to discriminate between pathology groups. This allowed a minimum target signal-to-noise ratio (SNR) to be determined for the MINERVA instrument whilst maintaining an acceptable ability to discriminate between pathology types.

9 High Resolution mid-IR Imaging at University of Exeter

One of the main objectives of Exeters group within the MINERVA project is large scale pathology screening using mid-infrared (mid-IR) spectroscopy. Currently FTIR spectral histopathology, which has the potential to develop as a cancer diagnostic tool, is carried out using a heated silicon carbide rod (“Globar”) as the mid-IR light source and focal plane array (FPA) based detectors. This technology is limited by the low flux of the light source and the limited tissue area that can be measured in a given amount of time.

The novel technologies being developed in the MINERVA project; consisting of mid-IR super-continuum light source (instead of a “Globar”) and new generation mega-pixel (FPA) detectors (instead of a 128×128 pixel FPA), will be tested on pathological samples at the University of Exeter.

Currently the base instrument, a commercially available Agilent FTIR imaging system, in addition to the conventional Globar source coupled to FPA based imaging, has been retro-fitted with a new high-resolution imaging capability. The FTIR images acquired using this set-up provided a five-fold improvement in image resolution from $5.5 \mu\text{m}^2$ of the current technology to $1.1 \mu\text{m}^2$ using the high magnification optics (Fig. 12).

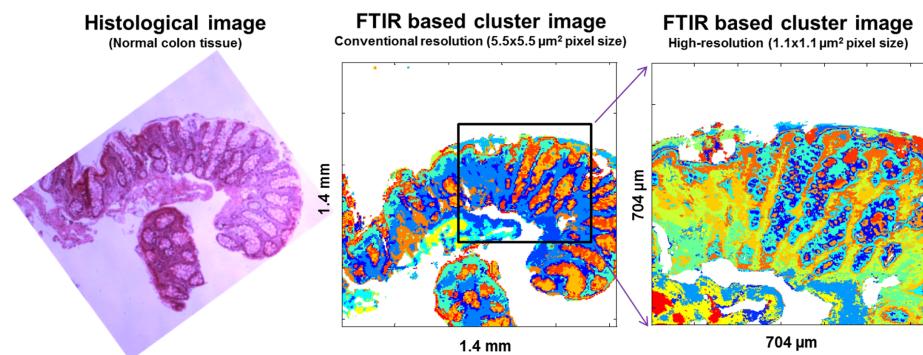


Fig. 12. FTIR based K-means cluster images obtained using conventional and high-resolution imaging compared with the histological image. Histological features based on the bio-molecular composition are partitioned in the cluster images. In the high-resolution imaging, tissue and cellular features are more apparent.

Future work in MINERVA will combine these novel technologies for large scale pathology screening, and also high-resolution imaging in tissue regions of interest, with

the aim to develop faster and accurate cancer diagnostic tools. Initially this will integrate with a 4.5 μm NKT MINERVA source, and later in the project it will be extended to very long mid-IR wavelengths: possibly out beyond 10 μm .

10 Development of Standardised Samples for mid-IR Spectrometer Instruments Testing at WWU

A key task of WWU is to transfer the MINERVA technologies to skin diagnostics and to use mid-IR spectroscopy for the fast screening of human body surfaces and identification of patho-physiologically altered cells and tissue lesions. This requires standardised cell and tissue sample standards with marker spectra for technology performance testing of the novel optical components and systems and for training of novel approaches for advanced data analysis.

The work of WWU in the first MINERVA project period thus focused on the establishment of standard samples with representative spectral information of human skin and skin cancer cells. WWU has established cell culture models which represent major cellular skin constituents and skin cancer cell types. Furthermore, sample preparation procedures on mid-IR compatible substrates have been developed that allow long-term storage of cell lines without significant losses in the quality of the spectral properties.

In order to identify suitable marker spectra of human skin, sample sets with different preparations and cell types were analysed with mid-IR spectroscopy in collaborative work with GHNT to retrieve reference data for technology performance testing and for the evaluation novel algorithms for sample analysis and classification developed by UPV. The principle component analysis (PCA) of mid-IR spectra from different cell types shows an excellent distinct grouping of skin components as fibroblasts and keratinocytes and cancer cells.

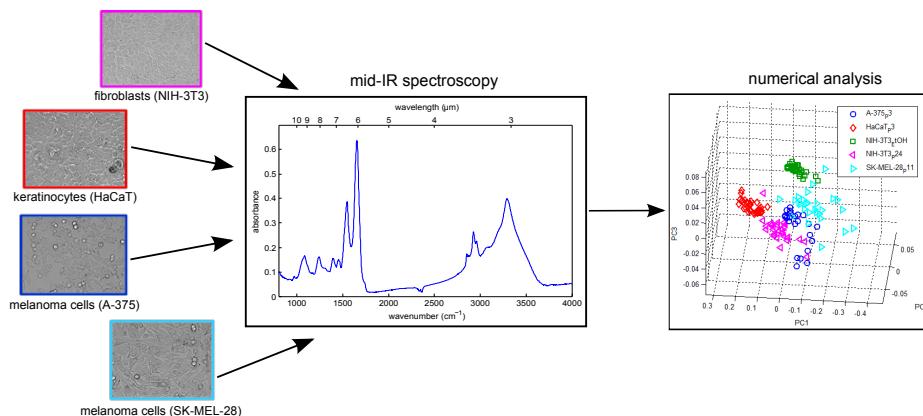


Fig. 13. Principle component analysis (PCA) of mid-IR spectral data from fibroblasts (NIH-3T3), keratinocytes (HaCaT) and skin cancer cells (A-375, SK-MEL-28) illustrates the differentiation between different cell types.

Figure 13 illustrates the analysis and differentiation of different cell types (cancer/non-cancer) that have been prepared at WWU for the example of results PCA of mid-IR spectral data from fibroblasts (NIH-3T3), keratinocytes (HaCaT), and skin cancer cells (A-375, SK-MEL-28). Based on these results, current and future activities at WWU in MINERVA focus on the development of novel mid-IR standards models for skin cancer detection that are based on 3D human skin equivalents in vitro. Further information about the standardization of the cell samples was presented in [26].

11 First Steps with MINERVA Image Processing at UPV

The first objectives of the image and signal processing group at UPV are focused on segmentation and registration of different kinds of images (Fig. 14): infra-red spectral images (IR), white light (WL), and those most used by clinicians at present, the haematoxylin and eosin (H&E) stained images. The latter is the current “gold standard” used to distinguish between a healthy or pathological patient sample.



Fig. 14. (a) Infra-red (IR) image. (b) White light (WL) image. (c) Haematoxylin and eosin (H&E) image.

The objective within MINERVA is to automatically segment regions of interest (healthy and pathological) in the H&E images and look for their features in the infrared spectrum. To achieve this goal the H&E image must be registered with the WL image (which is already registered with the infrared volume). So, the work is focused on two interactive steps: registration and segmentation.

Registration allows the matching of elements that clinicians considered important in the H&E images with the spectral images. A successful registration task would allow users to learn, and later identify, the areas from which diseased and healthy cells and patients can be distinguished (Fig. 15.a).

Segmentation concerns the accurate extraction of the cell contours (Fig. 15.b). This would reduce the huge amount of data to be analysed looking for subtle biochemical changes (“cancer markers”). Once the contours have been identified, the regions must be classified as healthy or cancerous depending on subtle features including shape, texture and clustering. This is an extremely difficult task, but the use of the spectral information in the mid-IR should eventually aid clinicians to improve on the current gold standard.

More details about the work on image processing in MINERVA project was presented in [27, 28].

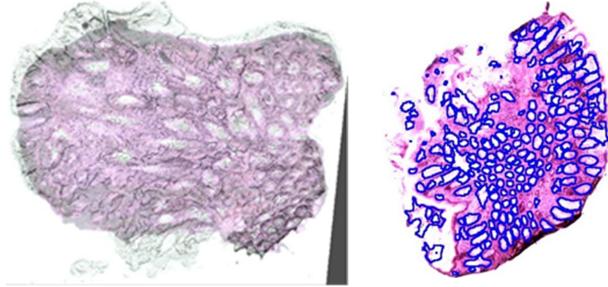


Fig. 15. (a) Projective registration test. (b) Segmentation sample test.

12 MINERVA Lite

A prototype MINERVA system that operates in the 2-4.5 μm wavelength band has been assembled so that the individual parts being developed by partners in the project can be evaluated together. The final MINERVA system will operate at even longer wavelengths. The key components in the integrated system are: NKT supercontinuum source, G&H acousto-optic tunable filter, Xenics IR camera, commercial microscope and IR optics and control electronics.

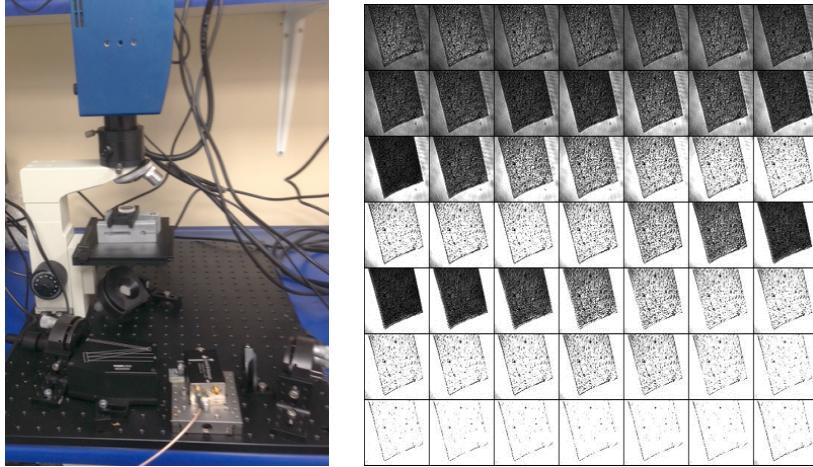


Fig. 16. (a) Photo of part of the “MINERVA Lite” laboratory set-up. (b) 0.3 Mpixel images at 49 wavelengths which can be used to form an (x, y, λ) image cube in 0.6 s.

The breadboard system is shown on Fig. 16.a. This system can take 0.3 megapixel images with 20 μm spatial resolution at a rate of 85 frames per second. Each image is taken at a different wavelength so that a set of 49 spectral images can be built up in 0.6 s (Fig. 16.b). These images can form an (x, y, λ) “image cube”. Each pixel records a spectrum and this has enabled MINERVA researchers to identify a polymer film in

the sample image. This important preliminary result will be extended in MINERVA to identify spectra from cancerous cells in tissue samples and in real time on live patients.

More related work in MINERVA project has been published in [29–31]. In addition, the research done in MINERVA has been mentioned and reviewed in [32, 33].

References

1. Minerva | Mid- to NEar infrared spectroscopy for imProVed medical diAgnostics. (<http://minerva-project.eu/>) Accessed: 2015-05-22.
2. Seddon, A.B.: Mid-infrared (IR)—A hot topic: The potential for using mid-IR light for non-invasive early detection of skin cancer *in vivo*. *Physica Status Solidi (B)* 250 (2013) 1020–1027
3. Oladeji, A., Sojka, L., Tang, Z., Furniss, D., Phillips, A., Seddon, A., Benson, T., Sujecski, S.: Numerical investigation of mid-infrared emission from Pr³⁺ doped GeAsGaSe fibre. *Optical and Quantum Electronics* 46 (2014) 593–602
4. Sakr, H., Tang, Z., Furniss, D., Sojka, L., Moneim, N., Barney, E., Sujecski, S., Benson, T., Seddon, A.: Towards mid-infrared fiber-lasers: rare earth ion doped, indium-containing, selenide bulk glasses and fiber. In: SPIE BiOS, International Society for Optics and Photonics (2014) 89380V
5. Sojka, L., Tang, Z., Furniss, D., Sakr, H., Oladeji, A., Bereš-Pawlak, E., Dantanarayana, H., Faber, E., Seddon, A., Benson, T., et al.: Broadband, mid-infrared emission from Pr³⁺ doped GeAsGaSe chalcogenide fiber, optically clad. *Optical Materials* 36 (2014) 1076–1082
6. Sujecski, S.: An efficient algorithm for steady state analysis of fibre lasers operating under cascade pumping scheme. *International Journal of Electronics and Telecommunications* 60 (2014) 143–149
7. Dantanarayana, H.G., Abdel-Moneim, N., Tang, Z., Sojka, L., Sujecski, S., Furniss, D., Seddon, A.B., Kubat, I., Bang, O., Benson, T.M.: Refractive index dispersion of chalcogenide glasses for ultra-high numerical-aperture fiber for mid-infrared supercontinuum generation. *Optical Materials Express* 4 (2014) 1444–1455
8. Seddon, A.: Mid-infrared photonics for early cancer diagnosis. In: Transparent Optical Networks (ICTON), 2014 16th International Conference on, IEEE (2014) 1–4
9. Sakr, H., Furniss, D., Tang, Z., Sojka, L., Moneim, N., Barney, E., Sujecski, S., Benson, T., Seddon, A.: Superior photoluminescence (PL) of Pr³⁺-In, compared to Pr³⁺-Ga, selenide-chalcogenide bulk glasses and PL of optically-clad fiber. *Optics express* 22 (2014) 21236–21252
10. Sujecski, S., Oladeji, A., Sojka, L., Phillips, A., Seddon, A., Benson, T., Sakr, H., Tang, Z., Furniss, D., Scholle, K., et al.: Modelling and design of mir chalcogenide glass fibre lasers. In: Numerical Simulation of Optoelectronic Devices (NUSOD), 2014 14th International Conference on, IEEE (2014) 111–112
11. Agger, C., Kubat, I., Møller, U., Moselund, P.M., Petersen, C., Napier, B., Seddon, A., Sujecski, S., Benson, T., Farries, M., et al.: Numerical demonstration of 3–12 μm supercontinuum generation in large-core step-index chalcogenide fibers pumped at 4.5 μm. In: Nonlinear Optics, Optical Society of America (2013) NW4A–09
12. Møller, U., Yu, Y., Petersen, C.R., Kubat, I., Mechlin, D., Brilland, L., Troles, J., Luther-Davies, B., Bang, O.: High Average Power Mid-infrared Supercontinuum Generation in a Suspended Core Chalcogenide Fiber. In: Nonlinear Photonics, Optical Society of America (2014) JM5A–54
13. Møller, U., Yu, Y., Kubat, I., Petersen, C.R., Gai, X., Brilland, L., Méchin, D., Caillaud, C., Troles, J., Luther-Davies, B., et al.: Multi-milliwatt mid-infrared supercontinuum generation in a suspended core chalcogenide fiber. *Optics express* 23 (2015) 3282–3291

14. Thomsen, C.L., Nielsen, F.D., Johansen, J., Pedersen, C., Moselund, P.M., Møller, U., Sørensen, S.T., Larsen, C., Bang, O.: New horizons for supercontinuum light sources: from UV to mid-IR. In: SPIE OPTO, International Society for Optics and Photonics (2013) 86370T
15. Møller, U., Bang, O.: Intensity noise of normal-pumped picosecond supercontinuum generation. In: Lasers and Electro-Optics Europe (CLEO EUROPE/IQEC), 2013 Conference on and International Quantum Electronics Conference, IEEE (2013) 1
16. Kubat, I., Agger, C., Moselund, P., Bang, O.: Mid-infrared supercontinuum generation in tapered ZBLAN fiber with a standard Erbium mode-locked fiber laser. In: Lasers and Electro-Optics Europe (CLEO EUROPE/IQEC), 2013 Conference on and International Quantum Electronics Conference, IEEE (2013) 1
17. Kubat, I., Agger, C., Moselund, P.M., Bang, O.: Optimized ZBLAN fiber for efficient and broadband mid-infrared supercontinuum generation through direct pumping at 1550nm. In: 1st International Workshop on Spatio-Temporal Complexity in Optical Fibers. (2013)
18. Kubat, I., Agger, C.S., Moselund, P.M., Bang, O.: Mid-infrared supercontinuum generation to $4.5 \mu\text{m}$ in uniform and tapered ZBLAN step-index fibers by direct pumping at 1064 or 1550 nm. *JOSA B* 30 (2013) 2743–2757
19. Moselund, P., Petersen, C., Leick, L., Seidelin Dam, J., Tidemand-Lichtenberg, P., Pedersen, C.: Highly Stable, All-fiber, High Power ZBLAN Supercontinuum Source Reaching $4.75 \mu\text{m}$ used for Nanosecond mid-IR Spectroscopy. In: Advanced Solid State Lasers, Optical Society of America (2013) JTh5A–9
20. Møller, U.V., Sørensen, S.T., Petersen, C.R., Kubat, I., Moselund, P.M., Bang, O.: Supercontinuum generation from ultraviolet to mid-infrared. (In: 15th Conference on Optical Fibers and Their Applications (OFTA 2014))
21. Kubat, I., Rosenberg Petersen, C., Møller, U.V., Seddon, A., Benson, T., Brilland, L., Méchin, D., Moselund, P.M., Bang, O.: Thulium pumped mid-infrared $0.9\text{--}9\mu\text{m}$ supercontinuum generation in concatenated fluoride and chalcogenide glass fibers. *Optics express* 22 (2014) 3959–3967
22. Kubat, I., Petersen, C.R., Møller, U., Seddon, A., Benson, T., Brilland, L., Méchin, D., Moselund, P., Bang, O.: Mid-infrared supercontinuum generation in concatenated fluoride and chalcogenide glass fibers covering more than three octaves. In: CLEO: Science and Innovations, Optical Society of America (2014) STh3N–1
23. Kubat, I., Agger, C.S., Møller, U., Seddon, A.B., Tang, Z., Sujecki, S., Benson, T.M., Furniss, D., Lamrini, S., Scholle, K., et al.: Mid-infrared supercontinuum generation to $12.5 \mu\text{m}$ in large na chalcogenide step-index fibres pumped at $4.5 \mu\text{m}$. *Optics express* 22 (2014) 19169–19182
24. Petersen, C.R., Møller, U., Kubat, I., Zhou, B., Dupont, S., Ramsay, J., Benson, T., Sujecki, S., Abdel-Moneim, N., Tang, Z., et al.: Mid-infrared supercontinuum covering the $1.4\text{--}13.3 \mu\text{m}$ molecular fingerprint region using ultra-high NA chalcogenide step-index fibre. *Nature Photonics* 8 (2014) 830–834
25. Martijn, H., Asplund, C., von Würtemberg, R.M., Malm, H.: High performance MWIR type-II superlattice detectors. In: Proc. of SPIE Vol. Volume 8704. (2013) 87040Z–1
26. Kastl, L., Rommel, C.E., Kemper, B., Schnekenburger, J.: Standardized cell samples for midir technology development. In: SPIE BiOS, International Society for Optics and Photonics (2015) 931507
27. Naranjo, V., Villanueva, E., Lloyd, G.R., Stone, N., Lopez-Mir, F., Alcaniz, M.: Stained and infrared image registration as first step for cancer detection. In: Biomedical and Health Informatics (BHI), 2014 IEEE-EMBS International Conference on, IEEE (2014) 420–423
28. López-Mir, F., Naranjo, V., Morales, S., Angulo, J.: Probability density function of object contours using regional regularized stochastic watershed. In: Image Processing (ICIP), 2014 IEEE International Conference on, IEEE (2014) 4762–4766

29. Stevens, G., Woodbridge, T.: Development of low loss robust soft-glass fiber splices. In: Workshop on Specialty Optical Fibers and their Applications, Optical Society of America (2013) W3–21
30. Markos, C., Kubat, I., Bang, O.: Hybrid polymer photonic crystal fiber with integrated chalcogenide glass nanofilms. *Scientific reports* 4 (2014)
31. Valle, S., Ward, J., Pannell, C., Johnson, N.: Acousto-optic tunable filter for imaging application with high performance in the ir region. In: SPIE OPTO, International Society for Optics and Photonics (2015) 93590E–93590E
32. Maragkou, M.: Supercontinuum: Reaching the mid-infrared. *Nature Photonics* 8 (2014) 746–746
33. Steinmeyer, G., Skibina, J.S.: Supercontinuum: Entering the mid-infrared. *Nature Photonics* 8 (2014) 814–815

El MUNDO: Embedding Measurement Uncertainty in Decision Making and Optimization

Carmen Gervet and Sylvie Galichet

LISTIC, Laboratoire d'Informatique, Systems, Traitement de l'Information et de la Connaissance, Université de Savoie, BP 8043974944, Annecy-Le-Vieux Cedex, France
`{gervetec, sylvie.galichet}@univ-savoie.fr`

Abstract. In this project we address the problem of modelling and solving constraint based problems permeated with data uncertainty, due to imprecise measurements or incomplete knowledge. It is commonly specified as bounded interval parameters in a constraint problem. For tractability reasons, existing approaches assume independence of the data, also called parameters. This assumption is safe as no solutions are lost, but can lead to large solution spaces, and a loss of the problem structure. In this paper we present two approaches we have investigated in the El MUNDO project, to handle data parameter dependencies effectively. The first one is generic whereas the second one focuses on a specific problem structure. The first approach combines two complementary paradigms, namely constraint programming and regression analysis, and identifies the relationships between potential solutions and parameter variations. The second approach identifies the context of matrix models and shows how dependency constraints over the data columns of such matrices can be modeled and handled very efficiently. Illustrations of both approaches and their benefits are shown.

1 Introduction

Data uncertainty due to imprecise measurements or incomplete knowledge is ubiquitous in many real world applications, such as network design, renewable energy economics, and production planning (e.g. [16, 22]). Formalisms such as linear programming, constraint programming or regression analysis have been extended and successfully used to tackle certain forms of data uncertainty. Constraint Programming (CP), is a powerful paradigm used to solve decision and optimization problems in areas as diverse as planning, scheduling, routing. The CP paradigm models a decision problem using constraints to express the relations between variables, and propagates any information gained from a constraint onto other constraints. When data imprecision is present, forms of uncertainty modeling have been embedded into constraint models using bounded intervals to represent such imprecise parameters, which take the form of coefficients in a given constraint relation. The solution sought can be the most robust one, that holds under the largest set of possible data realizations, or a solution set containing all solutions under any possible realization of the data. In such problems, uncertain data dependencies can exist, such as an upper bound on the sum of uncertain production rates per machine, or the sum of traffic distribution ratios from a router over several links. To our knowledge, existing approaches assume independence of the data when tackling real

world problems essentially to maintain computational tractability. This assumption is safe in the sense that no potential solution to the uncertain problem is removed. However, the set of solutions produced can be very large even if no solution actually holds once the data dependencies are checked. The actual structure of the problem is lost. Thus accounting for possible data dependencies cannot be overlooked.

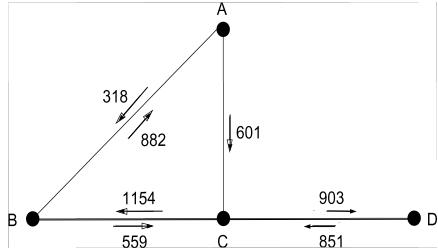


Fig. 1. Sigcomm4 network topology and mean values for link traffic.

Traditional models either omit any routing uncertainty for tractability reasons, and consider solely the shortest path routing or embed the uncertain parameters but with no dependency relationships. Values for the flow variables are derived by computing bounded intervals, which are safe enclosing of all possible solutions. Such intervals enclose the solution set without relating to the various instances of the parameters. For instance, the traffic between A and C can also pass through the link $A \rightarrow B$. Thus the flow constraint on this link also contains $0.3..0.7 * F_{AC}$. However, the parameter constraint stating that the sum of the coefficients of the traffic F_{AC} in both constraints should be equal to 1 should also be present. Assuming independence of the parameters for tractability reasons, leads to safe computations, but at the potential cost of a very large solution set, even if no solution actually holds. Not only is the problem structure lost, but there is not insight as to how the potential solutions evolve given instances of the data.

The question remains as to how to embed this information in a constraint model that would remain tractable. To our knowledge this issue has not been addressed. This is the purpose of our work.

In this paper we present two approaches to account for data dependency constraints in decision problems. We aim to more closely model the actual problem structure, refine the solutions produced, and add accuracy to the decision making process. The first one uses regression analysis to identify the relationship among various instances of the uncertain data parameters and potential solutions. Regression analysis is one of the most widely used statistical techniques to model and represent the relationship among variables. Recently, models derived from fuzzy regression have been defined to represent incomplete and imprecise measurements in a contextual manner, using intervals [6]. Such models apply to problems in finance or complex systems analysis in engineering whereby a relationship between crisp or fuzzy measurements is sought.

The basic idea behind our approach is a methodological process. First we extract the parameter constraints from the model, solve them to obtain tuple solutions over the parameters, and run simulations on the constraint models using the parameter tuples

as data instances that embed the dependencies. In the example above this would imply for the two constraints given, that if one parameter takes the value 0.3, the other one would take the value 0.7. A set of constraint models can thus be generated and solved efficiently, matching a tuple of consistent parameters to a potential solution. Finally, we run a regression analysis between the parameter tuples and the solutions produced to determine the regression function, i.e. see how potential solutions relate to parameter variations. This multidisciplinary approach is generic and provides valuable insights to the decision maker. However, while applying it to different problems we identified a certain problem structure that could be tackled without the need for the use of constraint problem set.

The second approach was then designed. We identified the context of matrix models, and showed how constraints over uncertain data can be handled efficiently in this context making powerful use of mathematical programming modeling techniques. For instance in a production planning problem, the rows denote the products to be manufactured and the columns the machines available. A data constraint such as an upper bound on the sum of uncertain production rates per machine, applies to each column of the matrix. Matrix models are of high practical relevance in many combinatorial optimization problems where the uncertain data corresponds to coefficients of the decision variables. Clearly, the overall problem does not need to be itself a matrix model. With the imprecise data specifying cells of an input matrix, the data constraints correspond to restrictions over the data in each column of the matrix. In this context, we observe that there is a dynamic relationship between the constraints over uncertain data and the decisions variables that quantify the usage of such data. Uncertain data are not meant to be pruned and instantiated by the decision maker. However, decision variables are meant to be instantiated, and the solver controls their possible values. This leads us to define a notion of relative consistency of uncertain data constraints, in relationship with the decision variables involved, in order to reason with such constraints. For instance, if an uncertain input does not satisfy a dependency constraint, this does not imply that the problem has no solution! It tells us that the associated decision variable should be 0, to reflect the fact that the given machine cannot produce this input.

The main contributions of our work lies in identifying and developing two multidisciplinary means to study the efficient handling of uncertain data constraints. Both approaches are novel towards the efficient handling of uncertain data constraints in combinatorial problems. We illustrate the benefits and impacts of both approaches respectively on a network flow and a production planning problem with data constraints.

The paper is structured as follows. Section 2 summarizes the related work. Section 3 describes the combination of constraint reasoning and regression analysis. Section 4 describes the concept of matrix models to handle data dependency constraints. A conclusion is finally given in Section 5.

2 Background and Related Work

The fields of regression analysis and constraint programming are both well established in computer science. While we identified both fields as complementary, there has been little attempt to integrate them together to the best of our knowledge. The reason is, we

believe, that technology experts tackle the challenges in each research area separately. However, each field has today reached a level of maturity shown by the dissemination in academic and industrial works, and their integration would bring new research insights and a novel angle in tackling real-world optimization problems with measurement uncertainty. There has been some research in Constraint Programming (CP) to account for data uncertainty, and similarly there has been some research in regression modeling to use optimization techniques.

CP is a paradigm within Artificial Intelligence that proved effective and successful to model and solve difficult combinatorial search and optimization problems from planning and resource management domains [19]. Basically it models a given problem as a Constraint Satisfaction Problem (CSP), which means: a set of variables, the unknowns for which we seek a value (e.g. how much to order of a given product), the range of values allowed for each variable (e.g. the product-order variable ranges between 0 and 200), and a set of constraints which define restrictions over the variables (e.g. the product order must be greater than 50 units). Constraint solving techniques have been primarily drawn from Artificial Intelligence (constraint propagation and search), and more recently Operations Research (graph algorithms, Linear Programming). A solution to a constraint model is a complete consistent assignment of a value to each decision variable.

In the past 15 years, the growing success of constraint programming technology to tackle real-world combinatorial search problems, has also raised the question of its limitations to reason with and about uncertain data, due to incomplete or imprecise measurements, (e.g. energy trading, oil platform supply, scheduling). Since then the generic CSP formalism has been extended to account for forms of uncertainty: e.g. numerical, mixed, quantified, fuzzy, uncertain CSP and CDF-interval CSPs [7]. The fuzzy and mixed CSP [11] coined the concept of parameters, as uncontrollable variables, meaning they can take a set of values, but their domain is not meant to be reduced to one value during problem solving (unlike decision variables). Constraints over parameters, or uncontrollable variables, can be expressed and thus some form of data dependency modeled. However, there is a strong focus on discrete data, and the consistency techniques used are not always effective to tackle large scale or optimization problems. The general QCSP formalism introduces universal quantifiers where the domain of a universally quantified variable (UQV) is not meant to be pruned, and its actual value is unknown *a priori*. There has been work on QCSP with continuous domains, using one or more UQV and dedicated algorithms [2, 5, 18]. Discrete QCSP algorithms cannot be used to reason about uncertain data since they apply a preprocessing step enforced by the solver QCSPsolve [12], which essentially determines whether constraints of the form $\forall X, \forall Y, C(X, Y)$, and $\exists Z, \forall Y, C(Z, Y)$, are either always true or false for all values of a UQV. This is a too strong statement, that does not reflect the fact that the data will be refined later on and might satisfy the constraint.

Example 1. Consider the following constraint over UQV:

$$\forall X \in \{1, 2, 3\}, \forall Y \in \{0, 1, 2\}, X \geq Y$$

Using QCSPsolve and its peers, this constraint would always be false since the possible parameter instance ($X = 1, Y = 2$) does not hold. However all the other

tuples do. This represents one scenario among 9 for the data realization and thus is very unlikely to occur if they are of equal opportunity. A bounds consistency approach is preferable as the values that can never hold would determine infeasibility, and if not the constraints will be delayed till more information on the data is known. In this particular example the constraint is bounds consistent. \square

Frameworks such as *numerical, uncertain, or CDF-interval* CSPs, extend the classical CSP to approximate and reason with continuous uncertain data represented by intervals; see the real constant type in Numerica [21] or the bounded real type in ECLiPSe [8]. Our previous work introduced the *uncertain and CDF-interval* CSP [23, 20]. The goal was then to derive efficient techniques to compute reliable solution sets that ensure that each possible solution corresponds to at least one realization of the data. In this sense they compute an enclosure of the set of solutions. Even though we identified the issue of having a large solution set, the means to relate different solutions to instances of the uncertain data parameters and their dependencies were not thought of.

On the other hand, in the field of regression analysis, the main challenges have been in the definition of optimization functions to build a relevant regression model, and the techniques to do so efficiently. Regression analysis evaluates the functional relationship, often of a linear form, between input and output parameters in a given environment. Here we are interested in using regression to seek a possible relation between uncertain constrained parameters in a constraint problem, e.g. distribution of traffic among two routers on several routes and the solutions computed according to the parameter instances.

We note also that methods such as sensitivity analysis in Operations Research allow to analyze how solutions evolve relative parameter changes. However, such models assume independence of the parameters. Related to our second approach, are the fields of Interval Linear Programming [17, 9] and Robust Optimization [3, 4]. In the former technique we seek the solution set that encloses all possible solutions whatever the data might be, and in the latter the solution that holds in the larger set of possible data realization. They do offer a sensitivity analysis to study the solution variations as the data changes. However, uncertain data constraints have been ignored for computational tractability reasons.

In the following section, we present the first approach, showing how we can seek possible relationships between the solutions, and the uncertain data variations while accounting for dependencies.

3 On Combining Constraint Reasoning and Regression Analysis

Let us first give the intuition behind this methodology through a small example.

3.1 Intuition

The core element is to go around the solving of a constraint optimization problem with uncertain parameter constraints by first identifying which data instances satisfy the parameter constraints alone. This way we seek tuples of data that do satisfy the uncertain

data constraints. We then substitute these tuples in the original uncertain constraint model to solve a set of constraint optimization problems (now without parameter constraints). Finally to provide further insight, we run a regression between the solutions produced and the corresponding tuples.

Consider the following fictitious constraint between two unknown positive variables, X and Y ranging in $0.0..1000.0$, with uncertain data parameters A, B taking their values in the real interval $[0.1..0.7]$:

$$A * X + B * Y = 150$$

The objective is to compute values for X and Y in the presence of two uncertain parameters (A, B). Without any parameter dependency a constraint solver based on interval propagation techniques with bounded coefficients, derives the ranges $[0.0..1000.0]$ for both variables X and Y [8]. Let us add to the model a parameter constraint over the uncertain parameters A and B : $A = 2 * B$. Without adding this parameter constraint to the model, since it is not handled by the solver, we can manually refine the bounds of the uncertain parameters in the constraint model such that the parameter constraint holds over the bounds, thus accounting partially for the dependency. We obtain the constraint system:

$$[0.2..0.7] * X + [0.1..0.35] * Y = 150$$

The solution returned to the user is a solution space:

$$X \in [0.0..750.0], Y \in [0.0..1000.0]$$

The actual polyhedron describing the solution space is depicted in Fig. 2.

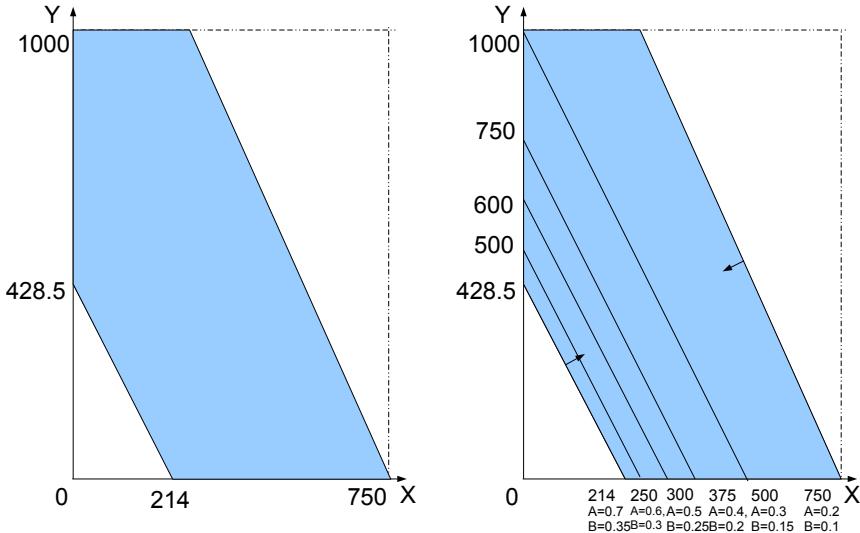


Fig. 2. Left: Solution space. Tight and certain bounds for the decision variables: $[0, 750]$ $[0, 1000]$. Right: Solution vectors of problem instances with consistent parameter solutions.

We now give the intuition of our approach. The idea is to first solve the parameter dependency constraints alone to obtain solution tuples, not intervals. To do so we use a traditional branch and bound algorithm. We obtain a set of tuples for A and B such that for each tuple the constraint $A = 2 * B$ holds. The idea is to have a well distributed sample of solutions for this parameter constraint.

We obtain a set of tuples that satisfy the parameter constraint, in this case for instance $(0.2, 0.1), (0.3, 0.15), (0.4, 0.2), (0.5, 0.25), (0.6, 0.3), (0.7, 0.35)$. We then substitute each tuple in the uncertain constraint model rendering it a standard constraint problem, and solve each instance. We record the solution matching each tuple instance. The issue now is that even though we have a set of solutions for each tuple of parameters, there is no indication how the solutions evolve with the data. The tuples might only represent a small set within the uncertainty range. The idea is to apply a regression analysis between both. The regression function obtained shows the potential relationship between the data parameters, that do satisfy the parameter constraints, and the solutions. In this small example we can visualize how the solution evolves with the data, see Fig. 2 on the right. In the case of much larger data sets, a tool like Matlab can be used to compute the regression function and display the outcome. The algorithm and complexity analysis are given in the following section.

3.2 Methodology and Algorithm

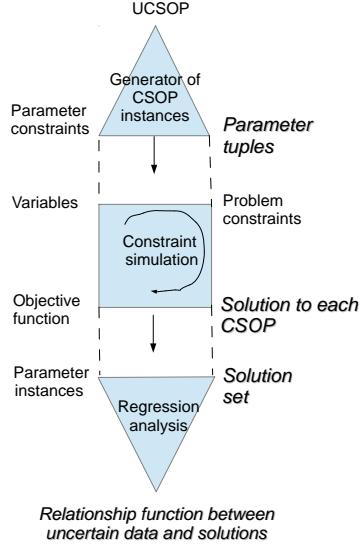
Our methodology is a three-steps iterative process: 1) Extract the uncertain parameter constraints from the uncertain optimization problem and run branch and bound to produce a set of tuple solutions, 2) solve a sequence of standard constraint optimization problems where the tuples are being substituted to the uncertain parameters. This is a simulation process that produces, if it exists, one solution per tuple instance. And finally, 3) run a regression analysis on the parameter instances and their respective solution, to identify the relationship function showing how the solutions evolve relative to the with consistent parameters. The overall algorithmic process is given in Fig. 3. The outcome of each step is highlighted in italic bold.

A constraint satisfaction and optimization problem, or CSOP, is a constraint satisfaction problem (CSP) that seeks complete and consistent instantiations optimizing a cost function. We use the notion of uncertain CSOP, or UCSOP first introduced in [23]. It extends a classical CSOP with uncertain parameters.

Uncertain CSOP and Uncertain Parameter Constraints. We first recall a CSOP. It is commonly specified as a tuple $(\mathcal{X}, \mathcal{D}, \mathcal{C}, f)$, where

- \mathcal{X} is a finite set of variables,
- \mathcal{D} is the set of corresponding domains,
- $\mathcal{C} = \{c_1, \dots, c_m\}$ is a finite set of constraints,
- f is the objective function over a subset of the variables.

Definition 1 (UCSOP). An uncertain constraint satisfaction and optimization problem is a classical CSOP in which some of the constraints may be uncertain, and is specified by the tuple $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}}, \Lambda, \mathcal{U}, f)$. The finite set of parameters is denoted by Λ , and the

**Fig. 3.** Process.

set of ranges for the parameters by \mathcal{U} . A solution to a UCSOP is a solution space enclosing safely the set of possible solutions.

Example 2. Let $X_1 \in D_1$ and $X_2 \in D_2$ both have domains $D_1 = D_2 = [1.0..7.0]$. Let λ_1 and λ_2 be parameters with uncertainty sets $U_1 = [2.0..4.0]$ and $U_2 = [1.0..6.0]$ respectively. Consider three constraints:

$$C_1 : X_1 > \lambda_1, C_2 : X_1 = X_2 + \lambda_2, C_3 : X_2 > 2$$

and the objective function to maximize $f(X_1, X_2) = X_1 + X_2$. This problem denotes the UCSOP $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}}, \Lambda, \mathcal{U}, f)$ where $\mathcal{X} = \{X_1, X_2\}$, $\mathcal{D} = \{D_1, D_2\}$, $\Lambda = \{\lambda_1, \lambda_2\}$, $\mathcal{U} = \{U_1, U_2\}$, and $\mathcal{C}_{\mathcal{X}} = \{C_1, C_2, C_3\}$.

Note that C_3 is a classical certain constraint; C_1 and C_2 are both uncertain constraints because they contain uncertain parameters. If now we add a constraint over the parameters such as $C_4 : \lambda_2 = \lambda_1 + 3$, the set of parameter constraints is $\mathcal{C}_{\Lambda} = \{C_4\}$.

Constraint Simulation. We now present our approach to solve a UCSOPs with parameter constraints, by transforming it into a set of tractable CSOPs instances where the parameter constraints hold. More formally, we consider a UCSOP $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}} \cup \mathcal{C}_{\Lambda}, \Lambda, \mathcal{U}, f)$.

Definition 2 (Instance of UCSOP). Let us denote by n the number of variables, m the number of uncertain parameters, p the number of parameter constraints, and $inst(\mathcal{U}_i)$ a value within the range of an uncertainty set. An instance of a UCSOP is a certain CSOP $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}})$ such that for each uncertain constraint $C_i(X_1..X_m, \lambda_1, ..\lambda_m)$, we have $\lambda_j = inst(\mathcal{U}_j)$, such that $\forall k \in \{1, .., p\}$, the parameter constraint $C_k(\lambda_1, ..\lambda_m)$ is satisfied.

Example 3. Continuing example 2, the UCSOP has two possible instances such that the parameter constraint $\lambda_2 = \lambda_1 + 3$ holds, given that $\lambda_1 \in \mathcal{U}_1, \lambda_2 \in \mathcal{U}_2$. The valid tuples (λ_1, λ_2) are $(2, 5)$, and $(3, 6)$. The CSOP instances we generate are:

$$C_1 : X_1 > 2, C_2 : X_1 = X_2 + 5, C_3 : X_2 > 2$$

and

$$C_1 : X_1 > 3, C_2 : X_1 = X_2 + 6, C_3 : X_2 > 2$$

with the same objective function to maximize $f = X_1 + X_2$.

The generator of CSOP instances extracts the parameter constraints, polynomial in the number of constraints in the worst case, then generates a set of parameter tuples that satisfy these constraints. We can use a branch and bound search on the parameter constraints of the UCSOP. The constraint simulation then substitutes the tuple solutions onto the original UCSOP to search for a solution to each generated CSOP. This is polynomial in the complexity of the UCSOP. The process is depicted in Algorithm 1.

Algorithm 1. Generate and solve CSOPs from one UCSOP.

Input: A UCSOP $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}} \cup \mathcal{C}_A, \Lambda, \mathcal{U}, f)$
Output: Solutions to the CSOPs

```

1 SolsTuples  $\leftarrow \emptyset$ 
2 extract( $\mathcal{C}_A$ )
3 Tuples  $\leftarrow \text{solveBB}(\Lambda, \mathcal{U}, \mathcal{C}_A)$ 
4 for  $T_i \in \text{Tuples}$  do
5   substitute  $\Lambda$  with  $T_i$  in  $(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}}, \Lambda, f)$ 
6    $S_i \leftarrow \text{solveOpt}(\mathcal{X}, \mathcal{D}, \mathcal{C}_{\mathcal{X}}, T_i, f)$ 
7   SolsTuples  $\leftarrow \text{SolsTuples} \cup \{(S_i, T_i)\}$ 
8 return SampleSols
```

Regression Analysis. The final stage of our process is to run a regression analysis between the parameter solution tuples $T \in \text{Tuples}$ and the corresponding $sol \in \text{SolsTuples}$ to estimate the relationship between the variations in the uncertain parameters called independent variables in regression analysis, and the solutions we computed, called dependent variables. Using the common approach we can model a linear regression analysis or one that minimizes the least-squares of errors.

Let us consider a linear regression, and the notation we used for the constraint model, where T_i is one parameter tuple, S_i the associated solution produced. We assume that the parameter instances were selected such that they are normally distributed in the first place. There are d of them. The regression model takes the following form. β is the regression coefficient to be found, and ϵ the noise. The regression model is then solved using MATLAB as a blackbox.

$$S = \beta T + \epsilon$$

where

$$S = \begin{pmatrix} S_1 \\ S_2 \\ \dots \\ S_d \end{pmatrix}, T = \begin{pmatrix} T_1 \\ T_2 \\ \dots \\ T_d \end{pmatrix}, \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_d \end{pmatrix}, \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \dots \\ \epsilon_d \end{pmatrix}$$

3.3 Illustration of the Methodology

We illustrate the benefits of our approach by solving an uncertain constraint optimization problem, the traffic matrix estimation for the sigcomm4 problem, given in Fig. 1. The topology and data values can be found in [16, 23]. Given traffic measurements over each network link, and the traffic entering and leaving the network at the routers, we search the actual flow routed between every pair of routers. To find out how much traffic is exchanged between every pair of routers, we model the problem as an uncertain optimization problem that seeks the min and max flow between routers such that the traffic link and traffic conservation constraints hold. The traffic link constraints state that the sum of traffic using the link is equal to the measured flow. The traffic conservation constraints, two per router, state that the traffic entering the network must equal the traffic originating at the router, and the traffic leaving the router must equal the traffic whose destination is the router.

We compare three models. The first one does not consider any uncertain parameters and simplifies the model to only the variables in bold with coefficient 1. The traffic between routers takes a single fixed path, as implemented in [16]. The second model extends the first one with uncertain parameters but without the parameter dependency constraints. The third one is our approach with the parameter dependency constraints added. A parameter constraint, over the flow F_{AB} , for instance, states that the coefficients representing one given route of traffic from A to B take the same value; and the sum of coefficients corresponding to different routes equals to 1. Note that the uncertain parameter equality constraints are already taken into account in the link traffic constraints. The uncertain parameters relative to flow distributions are commonly assumed between 30 and 70 % [23]. The distribution of split traffic depends mainly on the duration of traffic sampling, the configuration of the routers, and the routing protocol itself.

Decision variables:

$$[F_{AB}, F_{AC}, F_{AD}, F_{BA}, F_{BC}, F_{BD}, F_{CA}, f_{CB}, F_{CD}, F_{DA}, F_{DB}, F_{DC}] \in 0.0..100000$$

Parameters:

$$[\lambda_{1_{AB}}, \lambda_{1_{AC}}, \lambda_{1_{AD}}, \lambda_{1_{BC}}, \lambda_{1_{BD}}, \lambda_{2_{AB}}, \lambda_{2_{AC}}, \lambda_{2_{AD}}, \lambda_{2_{BC}}, \lambda_{2_{BD}}] \in 0.3..0.7$$

Link traffic constraints:

$A \rightarrow B$	$\lambda_{1_{AB}} * \mathbf{F}_{AB} + \lambda_{1_{AC}} * F_{AC} + \lambda_{1_{AD}} * F_{AD}$	$= 309.0..327.82$
$B \rightarrow A$	$\mathbf{F}_{BA} + \mathbf{F}_{CA} + \mathbf{F}_{DA} + \lambda_{1_{BC}} * F_{BC} + \lambda_{1_{BD}} * F_{BD}$	$= 876.39..894.35$
$A \rightarrow C$	$\lambda_{2_{AC}} * \mathbf{F}_{AC} + \lambda_{2_{AD}} * \mathbf{F}_{AD} + \lambda_{2_{AB}} * F_{AB} + \lambda_{1_{BC}} * F_{BC} + \lambda_{1_{BD}} * F_{BD}$	$= 591.93..612.34$
$B \rightarrow C$	$\lambda_{2_{BC}} * \mathbf{F}_{BC} + \lambda_{2_{BD}} * \mathbf{F}_{BD} + \lambda_{1_{AC}} * F_{AC} + \lambda_{1_{AD}} * F_{AD}$	$= 543.30..562.61$
$C \rightarrow B$	$\lambda_{2_{AB}} * F_{AB} + \mathbf{F}_{CB} + \mathbf{F}_{CA} + \mathbf{F}_{DA} + \mathbf{F}_{DB}$	$= 1143.27..1161$
$C \rightarrow D$	$\mathbf{F}_{CD} + \mathbf{F}_{BD} + \mathbf{F}_{AD}$	$= 896.11..913.98$
$D \rightarrow C$	$\mathbf{F}_{DC} + \mathbf{F}_{DB} + \mathbf{F}_{DA}$	$= 842.09..861.35$

Parameter constraints

$$\lambda_{1_{AB}} + \lambda_{2_{AB}} = 1, \lambda_{1_{AC}} + \lambda_{2_{AC}} = 1, \lambda_{1_{AD}} + \lambda_{2_{AD}} = 1, \lambda_{1_{BC}} + \lambda_{2_{BC}} = 1$$

Traffic conservation constraints

<i>A origin</i>	$F_{AD} + F_{AC} + F_{AB}$	= 912.72..929.02
<i>A destination</i>	$F_{DA} + F_{CA} + F_{BA}$	= 874.70..891.00
<i>B origin</i>	$F_{BD} + F_{BC} + F_{BA}$	= 845.56..861.86
<i>B destination</i>	$F_{DB} + F_{CB} + F_{AB}$	= 884.49..900.79
<i>C origin</i>	$F_{CD} + F_{CB} + F_{CA}$	= 908.28..924.58
<i>C destination</i>	$F_{DC} + F_{BC} + F_{AC}$	= 862.53..878.83
<i>D origin</i>	$F_{DC} + F_{DB} + F_{DA}$	= 842.0..859.0
<i>D destination</i>	$F_{CD} + F_{BD} + F_{AD}$	= 891.0..908.0

Results. We first ran the initial model (constraints with variables in bold for the link traffic together with the traffic conservation constraints) and reproduced the model and results of [23]. We used the linear EPLEX solver. By adding the uncertain parameters there was no solution at all. This indicates that not all traffic could be rerouted given the traffic volume data given.

We then disabled the uncertainty over the traffic BD, maintaining its route through the router C only. A solution set was found, with solution bounds much larger than the initial bounds (without uncertain distribution). Indeed, the space of potential solutions expanded. However, when we run simulations using our approach on the model with dependency constraints, there was no solution to the model. This shows the importance of taking into account such dependencies, and also indicates in this case the data provided are very likely matching a single path routing algorithm for the sigcomm4 topology.

After enlarging the interval bounds of the input data we were able to find a solution with a 50 % split of traffic, but none with 40 – 60 or other combinations. This experimental study showed the strong impact of taking into account dependency constraints with simulations [13].

Exploiting the Problem Structure. After completing this study, and running the simulations, we identified that when the uncertain parameters are coefficients to the decision variables and follow a certain problem structure, we can improve the efficiency of the approach. Basically it became clear that the fact that the constraints on the uncertain parameters tell us about the potential values of the decision variables and not whether the problem is satisfiable or not. For instance, when we allow the traffic F_{BD} to be split, there was no solution. Possible interpretations are: 1) this distribution (30 – 70%) is not viable and the problem is not solvable, 2) that there is no traffic between B and D. In the latter case, the handling of the dependency constraints should be done hand in hand with the labeling of the decision variables. This has been the subject of our second approach [14].

4 Matrix Models

The main novel idea behind this approach is based on the study of the problem structure. We identify the context of matrix models where uncertain data correspond to coefficients of the decision variables, and the constraints over these apply to the columns of the input matrix. Such data constraints state restrictions *on the possible usage of the*

data, and we show how their satisfaction can be handled efficiently in relationship with the corresponding decision variables.

In this context, the role and handling of uncertain data constraints is to determine "which data can be used, to build a solution to the problem". This is in contrast with standard constraints over decision variables, which role is to determine "what value can a variable take to derive a solution that holds". We illustrate the context and our new notion of uncertain data constraint satisfaction on a production planning problem inspired from [15].

Example 4. Three types of products are manufactured, P_1, P_2, P_3 on two different machines M_1, M_2 . The production rate of each product per machine is imprecise and specified by intervals. Each machine is available 9 hrs per day, and an expected demand per day is specified by experts as intervals. Furthermore we know that the total production rate of each machine cannot exceed 7 pieces per hour. We are looking for the number of hours per machine for each product, to satisfy the expected demand. An instance data model is given below.

Product	Machine M1	Machine M2	Expected demand
P_1	[2, 3]	[5, 7]	[28, 32]
P_2	[2, 3]	[1, 3]	[25, 30]
P_3	[4, 6]	[2, 3]	[31, 37]

The uncertain CSP model is specified as follows:

$$[2, 3] * X_{11} + [5, 7] * X_{12} = [28, 32] \quad (1)$$

$$[2, 3] * X_{21} + [1, 3] * X_{22} = [25, 30] \quad (2)$$

$$[4, 6] * X_{31} + [2, 3] * X_{32} = [31, 37] \quad (3)$$

$$\forall j \in \{1, 2\} : X_{1j} + X_{2j} + X_{3j} \leq 9 \quad (4)$$

$$\forall i \in \{1, 2, 3\}, \forall j \in \{1, 2\} : X_{ij} \geq 0 \quad (5)$$

Uncertain data constraints:

$$a_{11} \in [2, 3], a_{21} \in [2, 3], a_{31} \in [4, 6], a_{11} + a_{21} + a_{31} \leq 7 \quad (6)$$

$$a_{12} \in [5, 7], a_{22} \in [1, 3], a_{32} \in [2, 3], a_{12} + a_{22} + a_{32} \leq 7 \quad (7)$$

Consider a state of the uncertain CSP such that $X_{11} = 0$. The production rate of machine M_1 for product P_1 becomes irrelevant since $X_{11} = 0$ means that machine M_1 does not produce P_1 at all in this solution. The maximum production rate of M_1 does not change but now applies to P_2 and P_3 . Thus $X_{11} = 0$ infers $a_{11} = 0$. Constraint (6) becomes:

$$a_{21} \in [2, 3], a_{31} \in [4, 6], a_{21} + a_{31} \leq 7 \quad (8)$$

Assume now that we have a different production rate for P_3 on M_1 :

$$a_{11} \in [2, 3], a_{21} \in [2, 3], a_{31} \in [8, 10], a_{11} + a_{21} + a_{31} \leq 7 \quad (9)$$

P_3 cannot be produced by M_1 since $a_{31} \in [8, 10] \not\leq 7$, the total production rate of M_1 is too little. This does not imply that the problem is unsatisfiable, but that P_3 cannot be produced by M_1 . Thus $a_{31} \not\leq 7$ yields $X_{31} = 0$ and $a_{31} = 0$. \square

In this example we illustrated our interpretation of uncertain data dependency constraints, and how the meaning, role and handling of such constraints differs from standard constraints over decision variables. It must be different, and strongly tied with the decision variables the data relates to, as these are the ones being pruned and instantiated. In this following we formalize our approach by introducing a new notion of relative consistency for dependency constraints, together with a model that offers an efficient means to check and infer such consistency.

4.1 Formalization

We now formalize the context of matrix models we identified and the handling of uncertain data constraints within it.

Problem Definition

Definition 3 (Interval Data). An interval data, is an uncertain data, specified by an interval $[\underline{a}, \bar{a}]$, where \underline{a} (lower bound) and \bar{a} (upper bound) are positive real numbers, such that $\underline{a} \leq \bar{a}$.

Definition 4 (Matrix Model with Column Constraints). A matrix model with uncertain data constraints is a constraint problem or a component of a larger constraint problem that consists of:

1. A matrix (A_{ij}) of input data, such that each row i denotes a given product P_i , each column j denotes the source of production and each cell a_{ij} the quantity of product i manufactured by the source j . If the input is bounded, we have an interval input matrix, where each cell is specified by $[\underline{a}_{ij}, \bar{a}_{ij}]$.
2. A set of decision variables $X_{ij} \in \mathbb{R}^+$ denoting how many instances of the corresponding input shall be manufactured
3. A set of column constraints, such that for each column j : $\Sigma_i [\underline{a}_{ij}, \bar{a}_{ij}] @ c_j$, where $@ \in \{=, \leq\}$, and c_j can be a crisp value or a bounded interval.

The notion of (interval) input matrix is not to be confused with the Interval Linear Programming matrix (ILP) model in the sense that an ILP matrix is driven by the decision vector and the whole set of constraints as illustrated in the example below.

Example 5. The interval input matrix for the problem in example 2 is:

$$\begin{pmatrix} [\underline{a}_{11}, \bar{a}_{11}] & [\underline{a}_{12}, \bar{a}_{12}] \\ [\underline{a}_{21}, \bar{a}_{21}] & [\underline{a}_{22}, \bar{a}_{22}] \\ [\underline{a}_{31}, \bar{a}_{31}] & [\underline{a}_{32}, \bar{a}_{32}] \end{pmatrix} = \begin{pmatrix} [2, 3] & [5, 7] \\ [2, 3] & [1, 3] \\ [4, 6] & [2, 3] \end{pmatrix}$$

The ILP matrix for this problem is:

$$\begin{bmatrix} [2,3] [5,7] & 0 & 0 & 0 & 0 \\ 0 & 0 & [2,3] [1,3] & 0 & 0 \\ 0 & 0 & 0 & 0 & [4,6] [2,3] \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

The decision vector for the ILP matrix is $[X_{11}, X_{12}, X_{21}, X_{22}, X_{31}, X_{32}]$. \square

To reason about uncertain matrix models we make use of the robust counterpart transformation of interval linear models into linear ones. We recall it, and define the notion of relative consistency of column constraints.

Linear Transformation. An Interval Linear Program is a Linear constraint model where the coefficients are bounded real intervals [9, 3]. The handling of such models transforms each interval linear constraint into an equivalent set of atmost two standard linear constraints. Equivalence means that both models denote the same solution space. We recall the transformations of an ILP into its equivalent LP counterpart.

Property 1 (Interval linear constraint and equivalence). Let all decision variables $X_{il} \in \mathbb{R}^+$, and all interval coefficients be positive as well. The interval linear constraint $C = \sum_i [\underline{a}_{il}, \bar{a}_{il}] * X_{il} @ [c_l, \bar{c}_l]$ with $@ \in \{\leq, =\}$, is equivalent to the following set of constraints depending on the nature of $@$. We have:

1. $C = \sum_i [\underline{a}_{il}, \bar{a}_{il}] * X_{il} \leq [c_l, \bar{c}_l]$ is transformed into: $C = \sum_i \underline{a}_{il} * X_{il} \leq \bar{c}_l$
2. $C = \sum_i [\underline{a}_{il}, \bar{a}_{il}] * X_{il} = [c_l, \bar{c}_l]$ is transformed into:

$$C = \{\sum_i \underline{a}_{il} * X_{il} \leq \bar{c}_l \wedge \sum_i \bar{a}_{il} * X_{il} \geq c_l\}$$

Note that case 1 can take a different form depending on the decision maker risk aversity. If he assumes the highest production rate for the smallest demand (pessimistic case), the transformation would be: $C = \sum_i \bar{a}_{il} * X_{il} \leq c_l$. The solution set of the robust counterpart contains that of the pessimistic model.

Example 6. Consider the following constraint $a_1 * X + a_2 * Y = 150$ (case 2), with $a_1 \in [0.2, 0.7]$, $a_2 \in [0.1, 0.35]$, $X, Y \in [0, 1000]$. It is rewritten into the system of constraints: $l_1 : 0.7 * X + 0.35 * Y \geq 150 \wedge l_2 : 0.2 * X + 0.1 * Y \leq 150$.

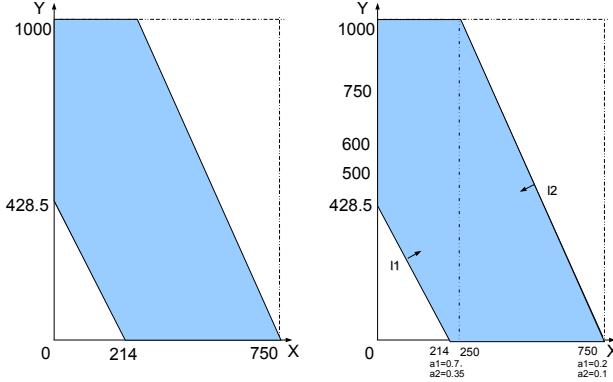


Fig. 4. Left: Solution space. Reliable bounds for the decision variables: $[0, 750]$ $[0, 1000]$. Right: Solution space bounded by the constraints l_1 and l_2 .

The polyhedron describing the solution space (feasible region) for X and Y is depicted in Fig. 4 (left), together with the boundary lines l_1 and l_2 , representing the two constraints above. \square

The transformation procedure also applies to the column constraints, and is denoted transf . It evaluates to true or false since there is no variable involved.

Relative Consistency. We now define in our context, the relative consistency of column constraints with respect to the decision variables. At the unary level this means that if $(X_{ij} = 0)$ then $(a_{ij} = 0)$, if $\neg \text{transf}(a_{ij} @ c_j)$ then $(X_{ij} = 0)$ and if $X_{ij} > 0$ then $\text{transf}(a_{ij} @ c_j)$ is true.

Definition 5 (Relative Consistency). A column constraint $\sum_i a_{il} @ c_l$ over the column l of a matrix $I * J$, is relative consistent w.r.t. the decision variables X_{il} if and only if the following conditions hold (C4. and C5. being recursive):

- C1. $\forall i \in I$ such that $X_{il} > 0$, we have $\text{transf}(\sum_i a_{il} @ c_l)$ is true
- C2. $\forall k \in I$ such that $\{\neg \text{transf}(\sum_{i \neq k} a_{il} @ c_l) \text{ and } \text{transf}(\sum_i a_{il} @ c_l)\}$ is true, we have $X_{kl} > 0$
- C3. $\forall i \in I$ such that X_{il} is free, $\text{transf}(\sum_i a_{il} @ c_l)$ is true
- C4. $\forall k \in I$, such that $\neg \text{transf}(a_{kl} @ c_l)$, we have $X_{kl} = 0$ and $\sum_{i \neq k} a_{il} @ c_l$ is relative consistent
- C5. $\forall k \in I$, such that $X_{kl} = 0$, we have $\sum_{i \neq k} a_{il} @ c_l$ is relative consistent

Example 7. Consider the Example 1. It illustrates cases C4. and C5, leading to a recursive call to C3. Let us assume now that the X_{i1} are free, and that we have the column constraint $[2, 3] + [2, 3] + [4, 6] = [7, 9]$. Rewritten into $2 + 2 + 4 \leq 9, 3 + 3 + 6 \geq 7$, we have $X_{31} > 0$, since $3 + 3 \not\geq 7$ and $3 + 3 + 6 \geq 7$. It is relative consistent with $X_{31} > 0$ (C2.).

Similarly if we had an uncertain data constraint limiting the total production rate of M_1 to 5 and $a_{31} \in [6, 7]$, yielding the column constraint:

$$a_{11} \in [2, 3], a_{21} \in [2, 3], a_{31} \in [6, 7], \quad a_{11} + a_{21} + a_{31} \leq 5$$

With all X_{i1} free variables this column constraint is not relative consistent, since the transformed relation is $2 + 2 + 6 \leq 5$. P_3 cannot be produced by M_1 because the maximum production rate of M_1 is too little ($6 \not\leq 5$, and condition 3 fails). This does not mean that the problem is unsatisfiable! Instead, the constraint can become relative consistent by inferring $X_{31} = 0$ and $a_{31} = 0$, since the remainder $a_{11} + a_{21} \leq 5$ is relative consistent (condition 3 holds). M_1 can indeed produce P_1 and P_2 . \square

4.2 Column Constraint Model

Our intent is to model column constraints and infer relative consistency while preserving the computational tractability of the model. We do so by proposing a Mixed Integer Interval model of a column constraint. We show how it allows us to check and infer relative consistency efficiently. This model can be embedded in a larger constraint model. The consistency of the whole constraint system is inferred from the local and relative consistency of each constraint.

Modeling Column Constraints. Consider the column constraint over column l :

$$\Sigma_i[\underline{a}_{il}, \overline{a}_{il}] @ c_l.$$

It needs to be linked with the decision variables X_{il} . Logical implications could be used, but they would not make an active use of consistency and propagation techniques. We propose an alternative MIP model.

First we recall the notion of bounds consistency we exploit here.

Definition 6 (Bound Consistency). [1] An n -ary constraint is Bound Consistent (BC), iff for each bound of each variable there exists a value in each other variable's domain, such that the constraint holds.

A constraint system with column constraints is BC if each constraint is BC.

Example 8. The constraint $X \in [1, 2], Y \in [2, 4], Z \in [1, 2], X + Y + Z = 5$ is not bounds consistent because the value $Y = 4$ cannot participate in a solution. Once the domain of Y is pruned to $[2, 3]$, the constraint is BC. \square

New Model. To each data we associate a Boolean variable. Each indicates whether: 1) the data must be accounted for to render the column constraint consistent, 2) the data violates the column constraint and needs to be discarded, 3) the decision variable imposes a selection or removal of the data. Thus the column constraint in transformed state is specified as a scalar product of the data and Boolean variables. The link between the decision variables and their corresponding Booleans is specified using a standard mathematical programming technique that introduces a big enough positive constant K , and a small enough constant λ .

Theorem 1 (Column Constraint Model). Let $X_{il} \in \mathbb{R}^+$ be decision variables of the matrix model for column l . Let B_{il} be Boolean variables. Let K be a large positive number, and λ a small enough positive number. A column constraint

$$\Sigma_i[\underline{a}_{il}, \overline{a}_{il}] @ c_l$$

is relative consistent if the following system of constraints is bounds consistent

$$\text{transf}(\Sigma_i[\underline{a}_{il}, \overline{a}_{il}] \times B_{il} @ c_l) \quad (10)$$

$$\forall i, 0 \leq X_{il} \leq K \times B_{il} \quad (11)$$

$$\forall i, \lambda \times B_{il} \leq X_{il} \quad (12)$$

Proof. The proof assumes that the system of constraints (10-12) is BC and proves that this entails that the column constraint is relative consistent. Note that constraint (10) is transformed into an equivalent linear model using the transformation procedure given in Section 4.1 relative to the instance of $@$ used.

If the system of constraints is BC, by definition each constraint is BC. If all X_{il} are free, so are the B_{il} (they do not appear elsewhere), and since (10) is BC then condition 3 holds and the column constraint is relative consistent. If some of the X_{il} are strictly positive and ground, the corresponding Booleans are set to 1 (since 11 is BC), and given that (10) is BC by supposition, condition 3 holds and the column constraint is relative consistent. If some of the X_{il} are ground to 0, so are the corresponding B_{il} since (12) is BC, and since (10) is BC, condition 3 holds (the remaining decision variables are either strictly positive or free and constraint (10) is BC). \square

Complexity. For a given column constraint, if we have n uncertain data (thus n related decision variables), our model generates n Boolean variables and $O(2n + 2) = O(n)$ constraints. This number is only relative to the size of a column and does not depend on the size or bounds of the uncertain data domain.

4.3 Column Constraints in Optimization Models

The notion of Bounds Consistency for a constraint system ensures that all possible solutions are kept within the boundaries that hold. As we saw, for the column constraints this means that resources that can't be used (do not satisfy the dependency) cannot be selected to contribute to the solutions, and those that must be used are included in the potential solution.

Such an approach is not limited to decision problems, it can be easily embedded in optimization models. An optimization problem with bounded data can be solved using several objective functions such as minmax regret or different notions of robustness. Our column constraint model can naturally be embedded in any uncertain optimization problems provided an input matrix model is associated with the specification of the uncertain data dependency constraints. Side constraints can be of any form.

The only element to be careful about is the transformation model chosen for the column constraints. The one given in Section 4.1. is the robust one that encloses all possible solutions. However, depending on the risk adversity of the decision maker some more restrictive transformations can be used as we discussed. Clearly if the whole problem can be modelled as a Mixed Integer Problem (MIP), MIP solvers can be used.

4.4 Illustration of the Matrix Model Approach

We illustrate the approach on the production planning problem. The robust model is specified below. Each interval linear core constraint is transformed into a system of two linear constraints, and each column constraint into its robust counterpart.

For the core constraints we have:

$$\begin{aligned} 2 * X_{11} + 5 * X_{12} &\leq 32, \quad 3 * X_{11} + 7 * X_{12} \geq 28, \\ 2 * X_{21} + X_{22} &\leq 30, \quad 3 * X_{21} + 3 * X_{22} \geq 25, \\ 4 * X_{31} + 2 * X_{32} &\leq 37, \quad 6 * X_{31} + 3 * X_{32} \geq 31, \\ \forall j \in \{1, 2\}, X_{1j} + X_{2j} + X_{3j} &\leq 9, \\ \forall i \in \{1, 2, 3\}, \forall j \in \{1, 2\} : X_{ij} &\geq 0, \\ \forall i, j, X_{ij} &\geq 0, B_{ij} \in \{0, 1\}, \end{aligned}$$

And for the column constraints:

$a_{11} \in [2, 3], a_{21} \in [2, 3], a_{31} \in [4, 6]$, $a_{11} + a_{21} + a_{31} \leq 7$ and
 $a_{12} \in [5, 7], a_{22} \in [1, 3], a_{32} \in [2, 3]$, $a_{12} + a_{22} + a_{32} \leq 7$ transformed into:

$$\begin{aligned} 2 * B_{11} + 2 * B_{21} + 4 * B_{31} &\leq 7, \\ 5 * B_{12} + B_{22} + 2 * B_{32} &\leq 7, \\ \forall i \in \{1, 2, 3\}, j \in \{1, 2\} \quad 0 \leq X_{ij} &\leq K * B_{ij}, \\ \forall i \in \{1, 2, 3\}, j \in \{1, 2\} \quad \lambda * B_{ij} &\leq X_{ij} \end{aligned}$$

We consider three different models: 1) the robust approach that seeks the largest solution set, 2) the pessimistic approach, and 3) the model without column data constraints. They were implemented using the ECLiPSe `ic` interval solver [8]. We used the constants $K=100$ and $\lambda = 1$. The column constraints in the tightest model take the form: $3 * B_{11} + 3 * B_{21} + 6 * B_{31} \leq 7$ and $7 * B_{12} + 3 * B_{22} + 3 * B_{32} \leq 7$.

The solution set results are summarized in the following table with real values rounded up to hundredth for clarity. The tightest model, where the decision maker assumes the highest production rates has no solution.

Variables	With column constraints		Without column constraints
	Robust model Booleans	Solution bounds	Solution bounds
X_{11}	0	0.0..0.0	—
X_{12}	1	4.0..4.5	—
X_{21}	1	3.33..3.84	—
X_{22}	1	4.49..5.0	—
X_{31}	1	5.16..5.67	—
X_{32}	0	0.0..0.0	—

Results. From the table of results we can clearly see that:

1. Enforcing Bounds Consistency (BC) on the constraint system without the column constraints, is safe since the bounds obtained enclose the ones of the robust model with column constraints. However, they are large, and the impact of accounting for the column constraints, both in the much reduced bounds obtained, and to detect infeasibility is shown.
2. The difference between the column and non column constraint models is also interesting. The solutions show that only X_{11} and X_{32} can possibly take a zero value from enforcing BC on the model without column constraints. Thus all the other decision variables require the usage of the input data resources. Once the column constraints are enforced, the input data a_{11} and a_{32} must be discarded since otherwise the column constraints would fail. This illustrates the benefits of relative consistency over column constraints.
3. The tightest model fails, because we can see from the solution without column constraints that a_{21} and a_{31} must be used since their respective X_{ij} are strictly positive in the solution to the model without column constraints. However from the

tight column constraint they can not both be used at full production rate at the same time. The same holds for a_{12} and a_{22} .

All computations were performed in constant time given the size of the problem. This approach can easily scale up, since if we have n uncertain data (thus n related decision variables) in the matrix model, our model generates n Boolean variables and $O(2n + 2) = O(n)$ constraints. This number does not depend on the size or bounds of the uncertain data domain, and the whole problem models a standard CP or MIP problem, making powerful use of existing techniques.

5 Conclusion

In this paper we introduced two multi-disciplinary approaches to account for dependency constraints among data parameters in an uncertain constraint problem. The first approach follows an iterative process that first satisfies the dependency constraints using a branch and bound search. The solutions are then embedded to generate a set of CSPs to be solved. However this does not indicate the relationship between the dependent consistent parameters and possible solutions. We proposed to use regression analysis to do so. The current case study showed that by embedding constraint dependencies only one instance had a solution. This was valuable information on its own, but limited the use of regression analysis. Further experimental studies are underway with applications in inventory management, problems clearly permeated with data uncertainty. This directed us towards the second approach where we identified the structure of matrix models to account for uncertain data constraints in relationship with the decision variables. Such models are common in many applications ranging from production planning, economics, inventory management, or network design to name a few. We defined the notion of relative consistency, and a model of such dependency constraints that implements this notion effectively. The model can be tackled using constraint solvers or MIP techniques depending on the remaining core constraints of the problem. Further experimental studies are underway with applications to large inventory management problems clearly permeated with such forms of data uncertainty and dependency constraints. An interesting challenge to our eyes, would be to investigate how the notion of relative consistency can be generalized and applied to certain classes of global constraints in a CP environment, whereby the uncertain data appears as coefficients to the decision variables. Even though our approaches have been applied to traditional constraint problems in mind, their benefits could be stronger on data mining applications with constraints [10].

References

1. Benhamou, F. Interval constraint logic programming. In Constraint Programming: Basics and Trends. LNCS, Vol. 910, 1-21, Springer, 1995.
2. Benhamou F. and Goualard F. Universally quantified interval constraints. In Proc. of CP-2000, LNCS 1894, Singapore, 2000.
3. Ben-Tal, A; and Nemirovski, A. Robust solutions of uncertain liner programs. Operations Research Letters, 25, 1-13, 1999.

4. Bertsimas, D. and Brown, D. Constructing uncertainty sets for robust linear optimization. *Operations Research*, 2009.
5. Bordeaux L., and Monfroy, E. Beyond NP: Arc-consistency for quantified constraints. In Proc. of CP 2002.
6. Boukezzoula R., Galichet S. and Bissier A. A MidpointRadius approach to regression with interval data. *International Journal of Approximate Reasoning*, Volume 52, Issue 9, 2011.
7. Brown K. and Miguel I. Chapter 21: Uncertainty and Change. *Handbook of Constraint Programming* Elsevier, 2006.
8. Cheadle A.M., Harvey W., Sadler A.J., Schimpf J., Shen K. and Wallace M.G. ECLiPSe: An Introduction. Tech. Rep. IC-Parc-03-1, Imperial College London, London, UK.
9. Chinneck J.W. and Ramadan K. Linear programming with interval coefficients. *J. Operational Research Society*, 51(2):209–220, 2000.
10. De Raedt L., Mannila H., O’Sullivan and Van Hentenryck P. organizers. Constraint Programming meets Machine Learning and Data Mining Dagstuhl seminar 2011.
11. Fargier H., Lang J. and Schiex T. Mixed constraint satisfaction: A framework for decision problems under incomplete knowledge. In Proc. of AAAI-96, 1996.
12. Gent, I., and Nightingale, P., and Stergiou, K. QCSP-Solve: A Solver for Quantified Constraint Satisfaction Problems. In Proc. of IJCAI 2005.
13. Gervet C. and Galichet S. On combining regression analysis and constraint programming. *Proceedings of IPMU*, 2014.
14. Gervet C. and Galichet S. Uncertain Data Dependency Constraints in Matrix Models. *Proceedings of CPAIOR*, 2015.
15. Inuiguchi, M. and Kume, Y. Goal programming problems with interval coefficients and target intervals. *European Journl of Oper. Res.* 52, 1991.
16. Medina A., Taft N., Salamatian K., Bhattacharyya S. and Diot C. Traffic Matrix Estimation: Existing Techniques and New Directions. *Proceedings of ACM SIGCOMM02*, 2002.
17. Oettli W. On the solution set of a linear system with inaccurate coefficients. *J. SIAM: Series B, Numerical Analysis*, 2, 1, 115-118, 1965.
18. Ratschan, S. Efficient solving of quantified inequality constraints over the real numbers. *ACM Trans. Computat. Logic*, 7, 4, 723-748, 2006.
19. Rossi F., van Beek P., and Walsh T. *Handbook of Constraint Programming*. Elsevier, 2006.
20. Saad A., Gervet C. and Abdennadher S. Constraint Reasoning with Uncertain Data usingCDF-Intervals *Proceedings of CP’AI-OR*, Springer,2010.
21. Van Hentenryck P., Michel L. and Deville Y. *Numerica: a Modeling Language for Global Optimization* The MIT Press, Cambridge Mass, 1997.
22. Tarim, S. and Kingsman, B. The stochastic dynamic production/inventory lot-sizing problem with service-level constraints. *International Journal of Production Economics* 88, 105119,2004.
23. Yorke-Smith N. and Gervet C. Certainty Closure: Reliable Constraint Reasoning with Uncertain Data *ACM Transactions on Computational Logic* 10(1), 2009.

Author Index

Abdalla, M.	53	Mosbahi, O.	25
Alcañiz, M.	53	Moselund, P.	53
Arsenio, A.	33	Nallala, J.	53
Bang, O.	53	Napier, B.	53
Barta, C.	53	Naranjo, V.	53
Farries, M.	53	Peñaranda, F.	53
Galichet, S.	70	Salem, M.	25
Gaspari, D.	53	Schnekenburger, J.	53
Gervet, C.	70	Seddon, A.	53
Hasal, R.	53	Serafin, S.	3
Jlalia, Z.	25	Sikström, E.	3
Kastl, L.	53	Smida, M.	25
Kemper, B.	53	Stevens, G.	53
Khalgui, M.	25	Stone, N.	53
Lamrini, S.	53	Sujecki, S.	53
Lloyd, G.	53	Vinella, R.	53
Maculewicz, J.	3	Ward, J.	53
Malm, H.	53		
Møller, U.	53		