SRS

Multi-Role Shadow Robotic System for Independent Living

Small or medium scale focused research project (STREP)

DELIVERABLE **D1.2**

Technology Assessment

Contract number :	247772
Project acronym :	SRS
Project title :	Multi-Role Shadow Robotic System for Independent Living

Deliverable number :	**D1.2**
Nature :	R – Report
Dissemination level :	PU – Public
Delivery date :	27 May 2010

Author(s) :	Georg Arbeiter
Partners contributed :	CU, CLMI-BAS, Fraunhofer, HdM, HPIS, INGEMA, PROFACTOR, ROBOTNIK and BED
Contact :	Georg Arbeiter, Fraunhofer IPA, Nobelstrasse 12, 70569 Stuttgart Phone: +49(0)711/970-1299, Fax +49(0)711/970-1008 Email: georg.arbeiter@ipa.fraunhofer.de

The SRS project was funded by the European Commission under the 7th Framework Programme (FP7) – Challenges 7: Independent living, inclusion and Governance

Coordinator: Cardiff University

Table of Contents

1. Introduction

We live in an environment of population ageing and many concerns have been discussed in most developed countries about the impacts of this increasing ageing process, including how to support the anticipated increases in demands for health and other support services. One side effect of this rapidly expanding older adult population is a significant increase in caregiving responsibilities being performed by family and friends. Smaller family sizes, along with geographically dispersed family members, make it difficult to provide this kind of unpaid care. Technology is seen as having significant potential in better equipping societies to address these issues and to cope with these increasing pressures.

Assistive technology (AT) are often considered as any service or tool that helps the elderly or disabled perform activities they have always performed but must now do differently. The scope of such a wide definition makes it possible for AT to cover any kind of equipment or service capable to fulfil the aforementioned definition: telecommunication equipment, computers, access systems, tools for independent living, mobility aids, video monitoring, remote health monitoring, electronic sensors and robotic-care systems are all considered AT [1]. Assistive technologies may not only support the aging adult but also their family and friends who serve as caregivers; devices that increase the independence of an older adult are intended to decrease the time required for caregiving. Access to these technologies influences how an elderly adult will be able to live independently, extending active and independent lives.

Until recently, attention in AT field has been mainly devoted to the assistive technology needs of elderly people with severe disabilities such as Alzheimer disease. Nevertheless, an increasing effort is being currently place in elderly people with mild disabilities, chronically ill or frailty.

AT can benefit the elderly and disabled people, improving their quality of life by enhancing their independence, increasing their participation in daily routines and activities, facilitating their mobility, communication, and other primary life functions. Technology has the potential to extend their physical independence, so they can stay for longer in their homes, living them a more dignified life. In this regard, a considerable interest has been placed in better enabling this population to receive care in their own homes, access to quality care services and extend their ability to remain in their own homes.

1.1. Assistive technologies and daily life difficulties.

Most classifications about disabilities and daily life difficulties rely on the *International Classification of Functioning, Disability and Health* (ICF) set by the World Health Organization [2]. The ICF is a checklist and a practical tool to elicit and record information on the functioning and disability of an individual. This information may be summarized for case records (for example, in clinical practice or social work). The ICF comprises are four sections:

- Part 1: Part 1a: Impairments of body functions
  Part 1b: Impairments of body structures
- Part 2: Activity limitations & participation restriction
- Part 3: Environmental factors
- Part 4: Other contextual information

Every part is segmented in smaller categories and each is given a value scale to determine the extent of impairments, activity and participation restrictions, and environmental barriers or facilitators. Part 2 and 3 are especially relevant for our research (see Table below).

Table

WHO International Classification of Functioning, Disability and Health (ICF) (extract)

Part 2: Activity limitations & participation restriction

Short List of A&P domains

d1. Learning and applying knowledge
- d110 Watching
- d115 Listening
- d140 Learning to read
- d145 Learning to write
- d150 Learning to calculate (arithmetic)
- d175 Solving problems

d2. General tasks and demands
- d210 Undertaking a single task
- d220 Undertaking multiple tasks

d3. Communication
- d310 Communicating with -receiving-spoken messages
- d315 Communicating with -receiving-non-verbal messages
- d330 Speaking
- d335 Producing non-verbal messages
- d350 Conversation

d4. Mobility
- d430 Lifting and carrying objects
- d440 Fine hand use (picking up, grasping)
- d450 Walking
- d465 Moving around using equipment (wheelchair, skates, etc.)
- d470 Using transportation (car, bus, train, plane, etc.)
- d475 Driving (riding bicycle and motorbike, driving car, etc.)

d5. Self care
- d510 Washing oneself (bathing, drying, washing hands, etc)
- d520 Caring for body parts (brushing teeth, shaving, grooming, etc.)
- d530 Toileting
- d540 Dressing
- d550 Eating
- d560 Drinking
- d570 Looking after one's health

d6. Domestic life
- d620 Acquisition of goods and services (shopping, etc.)
- d630 Preparation of meals (cooking etc.)
- d640 Doing housework (cleaning house, washing dishes laundry, ironing, etc.)
- d660 Assisting others

d7. Interpersonal interactions and relationships
- d710 Basic interpersonal interactions
- d720 Complex interpersonal interactions
- d730 Relating with strangers
- d740 Formal relationships
- d750 Informal social relationships
- d760 Family relationships
- d770 Intimate relationships

d8. Major life areas
- d810 Informal education
- d820 School education
- d830 Higher education
- d850 Remunerative employment
- d860 Basic economic transactions
- d870 Economic self-sufficiency

d9. Community, social and civic life
- d910 Community Life
- d920 Recreation and leisure
- d930 Religion and spirituality
- d940 Human rights
- d950 Political life and citizenship

Any other activity and participation

Part 3: Environmental factors

Short List of Environment

e1. Products and technology
- e110 For personal consumption (food, medicines)
- e115 For personal use in daily living
- e120 For personal indoor and outdoor mobility and transportation
- e125 Products for communication
- e150 Design, construction and building products and technology of buildings for public use
- e155 Design, construction and building products and technology of buildings for private use

e2. Natural environment and human made changes to environment
- e225 Climate
- e240 Light
- e250 Sound

e3. Support and relationships
- e310 Immediate family
- e320 Friends
- e325 Acquaintances, peers, colleagues, neighbours and community members
- e330 People in position of authority
- e340 Personal care providers and personal assistants
- e355 Health professionals
- e360 Health related professionals

e4. Attitudes
- e410 Individual attitudes of immediate family members
- e420 Individual attitudes of friends
- e440 Individual attitudes of personal care providers and personal assistants
- e450 Individual attitudes of health professionals
- e455 Individual attitudes of health related professionals
- e460 Societal attitudes
- e465 Social norms, practices and ideologies

e5. Services, systems and policies
- e525 Housing services, systems and policies
- e535 Communication services, systems and policies
- e540 Transportation services, systems and policies
- e550 Legal services, systems and policies
- e570 Social security, services, systems and policies
- e575 General social support services, systems n policies
- e580 Health services, systems and policies
- e585 Education and training services, systems n policies
- e590 Labour and employment services, systems and policies

Any other environmental factors

Source: ICF Checklist, World Health Organization.

1.1.1. International standards about assistive technologies.

The WHO-ICF classification rose key implication for other sorting made by international organizations, business associations, or academic work. The WHO-ICF acknowledged that human beings can experience a decrease in health throughout the lifespan and hence experience some degree o disability at any given point of their lives. Furthermore, the WHO-ICF classification shifted the focus from cause to impact and placed all health conditions on an equal footing allowing them to be compared using a common metric – the ruler of health and disability.
The International Organization for Standardization (ISO) embraced the WHO-ICF classification to build the commonly accepted international standard "*Assistive products for persons with disability - Classification and terminology*", also known as ISO 9999:2007,[3] which is the most widespread norm worldwide regarding AT and also the most used benchmark by governments, intergovernmental and business organizations to address the AT market.
The ISO 9999:2007 makes use of the term "Assistive Product" (AP) instead of "Assistive Technology" as a means to include in the definition more than just technological developments. Nevertheless, we will stick to using AT and not AP for practical reasons.
The ISO 9999:2008 classifies assistive products according to their function. This is a step ahead from previous ISO classifications, which mainly focused on the product and not on the user. This is a result of the integration of the WHO-ICF into the norm.
In sum, **Assistive Technology products consist on** any product (including devices, equipment, instruments, technology and software) specifically made or available in the market to prevent, compensate, control, alleviate, or counteract impairments, activity limitations and restrictions to participation (ISO, 2007). The ISO 9999:2007 classification groups products into classes and then into subclasses. For our work only the first class is relevant (See Table - the number indicates the class level and it is part or the standard).

Table
ISO 9999:2007. Assistive products for persons with disability - Classification and terminology.

First Level Classification

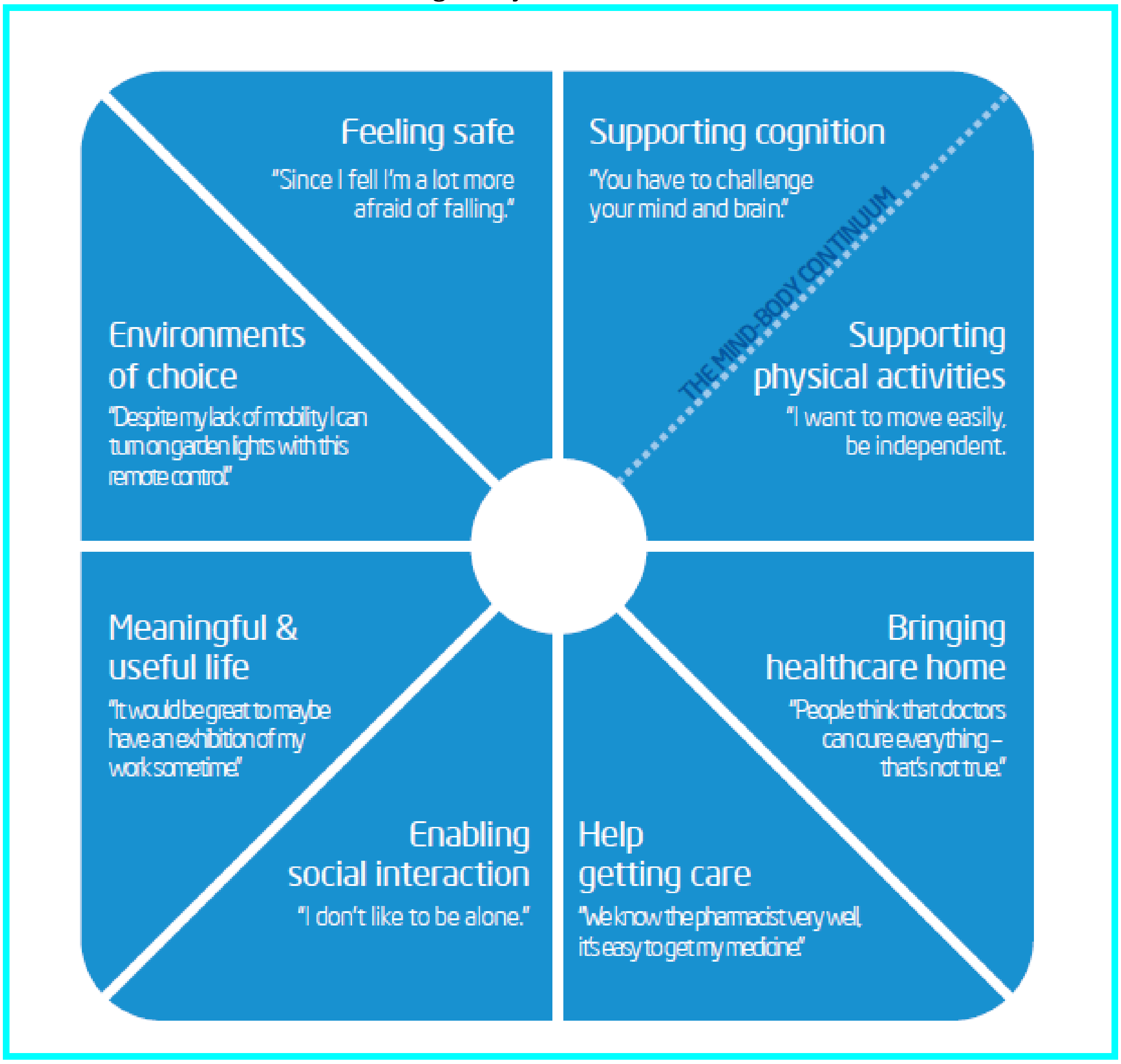
- 04 assistive products for personal medical
- 05 assistive products for training in skills
- 06 orthoses and prostheses
- 09 assistive products for personal care and protection
- 12 assistive products for personal mobility
- 15 assistive products for housekeeping
- 18 furnishings and adaptations to homes and other premises
- 22 assistive products for communication and information
- 24 assistive products for handling objects and devices
- 27 assistive products for environmental improvement, tools and machines
- 30 assistive products for recreation

Source: AENOR (2007). **Productos de apoyo para personas con discapacidad. Clasificación y terminología**. UNE-EN ISO 9999:2007. Madrid: AENOR Publications, page 12.

There are other important AT classifications. Most of them are based on the WHO-ICF. Others are founded on their own research results. For instance, The Global Aging Experience Project, carried out by Intel [4], proposes eight key areas of need for an ageing population when dealing with technology development (see Figure). These eight areas are the result of an ethnographic research undertaken in 2006 in seven European countries, along with an examination of world and European trends in ageing and technology. The 8 Key Areas intend to take into account diversity and the mind-body continuum. Regarding SRS project, our research is mainly, but not exclusively related to "Feeling safe" and "Environments of choice" areas.

This research points out key quality of life-related issues for technology development (Intel,

Figure
The Global Aging Experience Project. INTEL.
Ethnographic research. 2006
Eight Key Areas of Need



Source: INTEL Corporation. (2007).

2007: 27-31):

- People want to focus on what they CAN do, not what they <u>cannot</u> do.
- Healthy aging and independent living mean far more than "health".
- "Health" is not an objective quality; it's defined collaboratively and culturally.
- People mark the progression of aging by watershed events.
- Healthy aging is inextricably linked to social participation.
- The lived-in space is crucial to the experience of aging.
- Networks are large and increasingly complex just about everywhere.

1.1.2. Assistive technologies and personal rights.

The Universal Declaration of Human Rights (UDHR) soil for the extraordinary anti-discrimination legislation enacted during the second half of the 20th century and during the first decade of the current century [5]. In fact some experts say UDHR Article 7 (the right to equal protection against any discrimination) is the foundation to many other rights. Not only is UDHR Article 7 pertinent to our work, but so are Article 22 (right to social security), Article 25 (right to health and social care, especially for the elderly) and Article 27 (right to share in scientific advancement). The right to share in scientific advancement is considered an underdeveloped right as no comprehensive study and analysis have been conducted regarding the matter, while it entails serious implications for information and communication technologies, as will be seen later.

Although non-binding, the UDHR ploughed the ground for the adoption twenty years later of two major binding documents: the 1966 International Covenant on Economic, Social and Cultural Rights (ICESCR) [6] and the 1966 International Covenant on Civil and Political Rights (ICCPR) [7]. ICCPR Article 26 and ICESCR Article 2(§2) establish the prohibition of discrimination and the right to equal treatment, which is one of the European Commission's concerns when analysing access to AT in the Union by people with disabilities, elderly adults, and in the workplace (EC, 2003). ICESCR Article 9 establishes the right to social security and Article 15 (§1a) sets the right to enjoy the benefits of scientific progress and its applications. This is particularly relevant to any analysis of the potential role of the social security system in providing AT products to the citizens or to any private business aiming at developing a market niche on AT.

1.2. Summary of the technology assessment

For semi-autonomous tele-operation of service robots a variety of different technologies are evaluated in this report.

In section 2, telerobotics in general is investigated. Since telerobotics is one of the earliest fields of robotics, a lot of research has taken place. The research focuses on recent activities and research projects.

Information provided by intelligent homes can be used to enhance the robot's capabilities. An environment that is self-aware about its state is very useful for the operation of a robot. However, not every home can be converted into an intelligent one and the effort to build an intelligent home is extensive. A research about recent work in that field is provided by section 3.

In section 4, middlewares and communication concepts for robotics are introduced. Especially in tele-robotics, operation of a robot without an appropriate communication concept is to fail. Different software modules have to run on different machines and nevertheless communicate over large distances. An overview of existing software frameworks for robotics is also given in section 4. Some noticeable frameworks have evolved recently. They provide a base for the SRS framework development and are worth therefore an investigation.

Another important topic is robot safety. A robot that is not able to interact with the environment in a safe way cannot fulfil its tasks in a satisfactory manner. Because of that, safety concepts and research activities for robots are research in section 5.

Sections 6 and 7 deal with interaction technology. Whereas section 6 gives an overview of user interfaces and related input devices, section 7 shows haptics as an important technology for intuitive user interaction.

The results of the research in the field of cognitive capabilities are shown in sections 8-12. The first four sections focus on human motion analysis, environment perception and mobile manipulation. Those technologies enable the robot to obtain knowledge about itself and the environment it is operating in. Section 12 finally gives an introduction to state of the art in machine learning.

2. Telerobotics Control Model

2.1. Introduction

Telerobotics is one of the earliest fields of robotics. Since the first teleoperator was designed in 1940 the focus had been primarily on the nuclear, space, and underwater applications. The recent advances in technical areas such as speech and gestures recognition (AI, computer vision) and human machine interfaces (Haptics and Augmented reality) led to the emergence of new applications such as telesurgery, semiautonomous telerobotics. In this report existing control models for telerobotics will be explored. The result will be applied in the SRS as guidance for its high level control strategy.

2.2. Supervisory Control

A common model for telerobotics are developed by Ferell and Sheridan as supervisory control [8], in which the operator provides system commands to a human interactive computer that controls a task interactive computer, which translates higher level goals into a set of commands. The supervisor functions are:

(1) **Plan**, which includes the sub-activities of

- (a) Modelling the physical system,
- (b) Trading off objectives to decide what is satisfactory ("satisfying"), and
- (c) Formulating a strategy.

(2) **Teach**, including the distinctly different activities of

- (a) Deciding what to have the telerobot do, and
- (b) Deciding how to tell the telerobot to do it.

(3) **Monitor**, which includes

- (a) Deciding how to allocate attention among all the various signals that can be observed,
- (b) Estimating current system state or "situation", and
- (c) Detecting /diagnosing any abnormality in what is currently happening.

(4) **Intervene**, which in the case of abnormality means

- (a) Deciding on and effecting minor adjustments if they will suffice, or
- (b) Complete manual takeover, or
- (c) System shutdown; or
- (d) If the programmed action has come to a normal conclusion, it means reverting back to step (2).

(5) **Learn** from experience to improve future planning.

According to Sheridan a teleoperator is a machine enabling a human operator to move about, sense and mechanically manipulate objects at a distance. It usually has artificial sensors and effectors for manipulation and/or mobility, plus a means for the human to communicate with both. Most generally, any tool which extends a person's mechanical action beyond his reach is a teleoperator. A telerobot is a subclass of teleoperator in which the machine acts as a robot for short periods, but is monitored by a human supervisor and reprogrammed from time to time.

The telesensor programming (TSP) approach is a semi-autonomous concept under supervisory control that distributes intelligence between man and machine [9]. It was implemented in space robotic system.

Main Advantage:

- Can deal with large delay and unstable communication network between the robot and the remote human operator;
- Allow a shift of more and more autonomy and intelligence to the robot system during the operation.

2.3. Shared Control

Shared control is mainly developed and implemented in telesurgery [10]. It is based on local sensory feedback loops at the teleoperator site by which combined commands were refined autonomously to simplify the operation of teleoperator [11].

Main advantage

- Improve the efficiency of the task execution for the teleoperator.

The control model was further examined in 2004 by Defense Advanced Research Projects Agency (DARPA)/National Science Foundation (NSF) [12] as part of study on HRI. Teleoperation for remote presence applications is of great value to HRI research for several reasons, including how to represent the situational context, expectations, and social models for an application. Such a representation should model the relationships, frequency and content of interactions, richness of relationships, and amount of communications in such a way as to predict the impact of interactions on performance. The study also cited social informatics issues that could be captured by a model, such as who is accountable for what actions and what responsibilities and skills can be expected from the human and the robot.

2.4. FP6/FP7 research

Name	Status	Research Focuses	Main Outcomes	Web address
Improvement of the emergency risk management through secure mobile Mechatronic support to bomb disposal (**Rescuer**)	IST-511492 Completed 2008-April-30	The RESCUER project will endow an outdoor mechatronic system with set of two cooperative manipulators, so as to demonstrate the ability of the system to act as a human rescue professional in search and rescue-like operation scenarios. Search and Rescue (SAR) operations are a challenging application, due to their nature, they naturally foster advances in Artificial Vision, Navigation in outdoors unstructured environments.	Control Strategy of Two Robot Arms for Rescue and Bomb Disposal Missions A Hardware Architecture for Real-Time Extraction of Disparity Maps from Large Images	http://www.rescuer-ist.net
Vehicle Tele-operation System For Application		This project intends the development of a highly efficient vehicle teleoperation system for	Prototype of a teleoperated vehicle. (adaptation of commercial one) suitable for the specific project	http://laimuz.unizar.es/teledrive/

To Entertainment, Research And Remote Interventions (**TELEDRIVE**)		remote interventions, entertainment and research applications.	application	

3. Intelligent Home

3.1. Introduction

Several terms are used nowadays as almost synonyms of Intelligent Home like, Smart Home, Intelligent Space and Domotics[1].
Related fields are: Ambient Intelligence[2], Pervasive computing[3], Ubiquitous computing and Context awareness.
In the beginning the focus in the Smart Home field was mainly on single-system home automation, i.e. on automatic control of single home systems, such as lighting or heating. Security was also important with many houses equipped with alarm systems.
For the integration of multiple components a communication bus and protocol were needed. The development of the X10 protocol in the seventies introduced a new wave of commercial domotics components mostly using power line wiring for communication. Other technologies followed, such as CEBus and EIB, but X10 is still in use today due to its low cost and ease of installation. See [13] for a comparison of these home communication technologies.
Remote monitoring, especially for health-related variables, came along in the nineties, with the advent of the Internet; see for example [14].
Since the nineties monitoring of a person's daily activities has been considered as a simple way to monitor her health status [15]. The idea of monitoring people's activities of daily life (ADL) at home has since gained importance as one of the main branches of Smart Home / Ambient Intelligence research. The aim is to build and dynamically refine a model of human activities which enables the smart home to infer what inhabitants are doing so that the system can better support them.
The monitoring branch of this field, as can be seen from the mind map in Figure 3.1, is the richest in both researches and applications. In the last ten years lots of smart home systems have been presented which monitor several aspects of people home life, from presence to position to activities, from ambient parameters to personal health parameters, from intrusion detection to fall detection; some authors like [16] even propose to monitor emotions[4].

[1] We prefer to use the term Smart Home rather than Intelligent Home because a computer that could succeed in Turing's test [Turing] is still to be announced, let alone a home!
[2] See [Ramos] for a state of the art on Ambient Intelligence (dated 2007). The paper devotes some space also to possible applications in Elderly and Health Care, domotics and robotics integration.
[3] According to [Mihailidis] *A central theme in pervasive computing is building rich predictive models of human behavior from sensor data, enabling the environment to be aware of the activities performed within it. Using an accurate model, the environment can first determine the actions and intent of the user and then make considered decisions about the type of assistance to provide.* The same description can be applied to all the researches reported under the following section on Activity Monitoring.
[4] The research branch in this case is called Affective Computing, see http://affect.media.mit.edu/

In the following we reported separately references to health monitoring from activity monitoring.
In the following sections references are classified according to their category. Security and privacy issues are a common concern across all branches, to prevent both physical as well as digital intrusion in the smart home, thus we decided not to create a specific category.
Some recent references are reported on Activity and Health monitoring, but also on Home Automation systems, Infrastructure topics, and integration of robots in a smart home. Our literature research is concluded with some references on design issues and a glimpse on possible future research directions.

3.2. Literature research

Figure 3.1 it is shown a mind map of the main concepts found during this literature research.

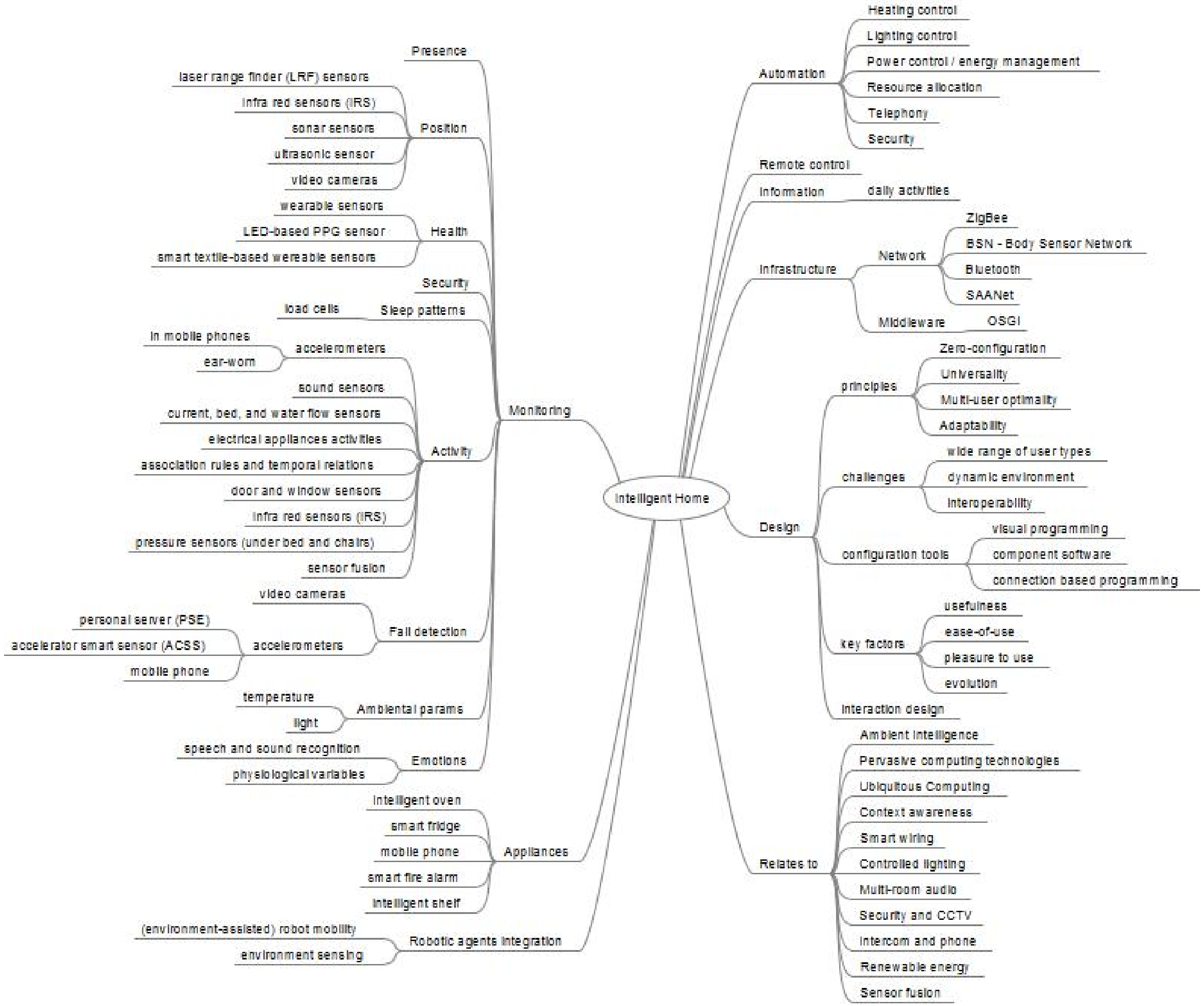


Figure 3.1 - Mind map of Intelligent Home concepts

3.2.1. Activity Monitoring

Activity monitoring is an expanding research field which aims at building predictive models of human behaviour starting from data gathered from several types of different sensors installed in a home environment. When the smart home environment is aware of the activities performed by the user, it can determine which kind of assistance the user may need and activate the proper controls or appliances to provide it.

The reviewed literature about activity monitoring could be further classified according to the type of sensors and the type of model used. The most used sensors are accelerometers, like in [17] [18] and [19], but also cameras ([20] and [21]), infrared ([17] and [22]), electrical water flow and floor pressure ([23] and [21]) sensors are used. To identify emotions a combination of physiological and environment sensors are used ([16] and [24]).
[17] reports the results of a research project on ”Ambient Intelligence”, “Sensor Fusion” and “Wireless Sensor Networks”. The proposed demonstrator provides health and activity monitoring, indoor tracking and identification of dangerous events. [18] proposes the use of accelerometers embedded in off-the-shelf cellular handsets to monitor elderly activity levels. [19] proposes the use of an ear-worn accelerometer for patients' activity monitoring. [25] proposes the use of Personal Server (PSE) accelerometers as fall detection sensors and proposes an algorithm to analyze related data.
[21] proposes an approach to home activity monitoring based on data fusion from multiple sensors (video, contact, pressure, flow, electrical, presence).
[22] analyzes a location-recognition algorithm using a Bayesian classifier which, starting from infrared sensor data, is able to detect people's position and trajectory in the home. The PILAS smart home management provides location-based services to inhabitants. [23] reports on the use of wireless sensors such as current, bed, and water flow sensors, to monitor elderly people activity and detect abnormality patterns. [26] presents an approach to elderly activity monitoring in their home based on information from electrical appliances.
[24] reports on using a global speech and sound recognition system set-up in a flat to detect distress situations. [16] proposes the architecture of a middleware for smart homes and reports initial experimental results. The system measures physiological and environment (context) variables to identify human emotions and automates appropriate home services to adapt the residents’ environment.

As for the type of model the most used is Bayesian networks ([19, 27] and [22]), but also Fuzzy logic ([20], [28] and [29]) and association rules [30].
[30] proposes and analyzes a method for elderly activity monitoring based on "the data mining technique of association rules and Allen’s temporal relations". The work starts from the observation that "similar activities lead to the appearance of a common set of temporal relations between specific sensors". [20] proposes a configurable system based on video cameras to monitor elderly activities in a home environment. A 3-D representation of the human is built in real time and fuzzy logic is used to model human activity. The system has been configured and evaluated for fall detection. [28] proposes a combined Fuzzy State / Q-learning algorithm to predict future user activities in an assistive environment. [27] proposes a method to build a smart home hybrid model based both on activity monitoring and preference learning. The model is dynamically adjusted by considering services acceptance / usage as user feedback. The focus is on activity recognition in a dynamic multi-user environment. Parameters of the preference model are trained with an Expectation Maximization (EM) algorithm. To model time information and to predict probability of an activity a Dynamic Bayesian Network (DBN) is used.

3.2.2. Health monitoring

Health monitoring is very close to activity monitoring. As noted above monitoring of a person's daily activities has been considered as a simple way to monitor her health status [15]. Many references could have been classified either under activity or health monitoring. References provided in this section are more specific to health monitoring and well-being solutions for elderly people, such as an exercise-management system [31] or a wearable airbag [32]. The most used sensors in this branch are wearable and equipped with a wireless connection ([33] and [34]), but also load cells ([35] and [36]) and microphones ([37] and [38]) are used. The monitored parameters range from the simplicity of cough sounds [37] to the complexity of sleep patterns [36].

[31] proposes an exercise-management system which provides dynamic exercise prescription based on awareness of the user's status monitored through sensors of a U-health environment. In [32] accelerometers are used for fall-detection and as a trigger for the inflation of a wearable airbag to protect elderly falling backwards.
[39] analyzes the state of the art in smart textile-based wearable biomedical systems and discusses related technical, medical and business issues. [33] presents a framework performing real-time analysis of physiological data to monitor people's health conditions.
[35] reports about a ZigBee-based wireless sensor network built to provide health monitoring into ordinary homes; several different sensors used (bed, heat, ECG, pedometer, microwave radar) are also discussed. [36] proposes the use of load cells placed under the bed for unobtrusive continuous monitoring of sleep patterns.
[37] proposes a health monitor system able to detect cough sounds. [38] presents SHA, a smart home for quadriplegia patients with voice-activated functionalities and particular attention to security issues.

3.2.3. Automation

Automation is the oldest application in smart homes, thus few related researches can be found in recent literature. Wide adoption of home automation technologies is hindered by two main issues. One is the high cost, both in absolute terms and compared to benefits, of home automation solutions and the other can be seen as lack of standards which results in poor integration and interoperability between different devices and systems. Recent researches try to cope with the two issues above by proposing novel approaches to interoperation [40] and by exploring which savings could be realized to improve the cost/benefit ratio ([41] and [42]). Another research topic is about handling the needs of multiple inhabitants in the same smart home [43].
As for the methods used we observe a prevalence of ontologies and rule-based systems ([40] , [42] and [43]), but also operational research is used [44].

[40] proposes a rule-based approach to domotic devices interoperation. [41] presents the DOMOSEC home automation solution based on standard domotic technologies and providing remote control, security system, eHealth, elderly adaptation, greenhouse automation and energy efficiency; all with a 3D interface.
In [42] an intelligent home system is proposed to handle security / intruder detection, local and remote control of appliances, daily activity reminder, solar panel orientation, etc. [43] proposes a home resource management system based on agents using ontologies and case-based reasoning. Resource allocation is done for power, water, gas, network bandwidth and other resources common to inhabitants of the smart home.

According to [44] smart homes will be equipped with smart shelves / cabinets using automatic identification (such as RFID) and alert systems. These appliances will be able to notify you when you have to re-order food items and may even automatically place orders to your preferred online grocery chains. The paper proposes an automated grocery ordering system (Grocery Shopping Agent) that minimizes the total shopping cost.
An alternative approach to home automation is proposed by [45] who proposes the smart home as a teaching environment for the occupants to acquire healthier habits. The approach to home automation is one which uses subtle and unobtrusive reminders rather than active enactment, to let the user more in control. The view is one of technology which supports human activities rather than automating them. The article also reports about physical-digital interaction studies in home environments.

3.2.4. Infrastructure

The researches reviewed under this category report about infrastructure technologies applied to smart home environments. Most of the works in this branch are about networking technologies (ZigBee and Bluetooth, [46], [47]), communication technologies (videoconferencing [48], telephony [49]) and user interface technologies (RFID [50], flex-sensor gloves [51]). Other works are about the management of home infrastructures (energy [52], oven [53]) and related security issues ([54]).
As for the methods used we find, as in other branches, mostly fuzzy logic [29] and Bayesian classifiers [55].

[46] praises ZigBee vs Bluetooth as the wireless network technology of choice for intelligent home systems. [47] reports about the development of a smart home network based on ZigBee technology in combination with smart home appliance communication protocol (SAANet).
[48] introduces a videoconferencing service integrated in an ambient intelligence (AmI) environment. [49] presents the results of a work aiming at improving the social presence experience in home telephony. Two different telephony systems are compared; one with loudspeakers and microphones embedded in the ceiling and one with a distributed set of clearly visible and tangible speakerphone units.
[50] The paper proposes a way to control home music playback using RFIDs embedded in a CD. The scenario is: a passive RFID reader is embedded into a digital table; once a resident put the Virtual-CD on it, the system identifies the object through its RFID and can for example directly deliver songs in streaming fashion to the networked media player embedded and hidden in the wall. [51] presents a system that interprets simple gestures from a couple of flex-sensor gloves, connected through a Body Sensor Network (BSN), as commands for intelligent home devices. Gestures are not intuitive at all. Much more could have been accomplished by just applying some imagination to a good technology starting point such as BSN-based flex-sensor gloves.
[52] presents a home energy management system based on power line communication and offering both auto-configuration and remote monitoring. [53] proposes an intelligent oven for manipulating recommended healthy recipe choices to suit each family member health concerns.
As noted by [54] *the idea of using existing electronics in smart home appliances and connecting them to the Internet is a new dimension along which technologies continues to grow*, but *there are serious security challenges that have to be addressed*. The author analyzes security incidents related to networked smart home appliances occurred in Japan and identifies existing challenges. Possible countermeasures are also discussed.

[29] proposes an adaptive control algorithm based on fuzzy logic for intelligent routing in a smart home environment. [55] proposes a context-aware middleware for ubiquitous computing based on multi-agent technology and naive Bayesian classifiers.

3.2.5. Robot integration

The integration of robots in a smart home environment is a recent research topic. Some researches focus on merging data provided by smart home sensors with data collected by the robot ([56] and [57]). Other researches rely more on robot's sensors to obtain monitoring and fall detection ([58] and [59]).

[56] reports building a low cost robotic assistant by using a distributed system of embedded microcontrollers, some on the mobile robot and some in the intelligent environment, for guiding tasks such as path following (using fuzzy logic), obstacle avoidance and data acquisition. [57] proposes a robotic service framework, called ubiquitous robotic space (URS), comprising three conceptual spaces: physical, semantic and virtual. The framework is experimented by developing a robotic security application for an office environment. - The physical space, with its Localization Sensor Network and Wireless Sensor Network, enhances both mobility and environment sensing capabilities of the robot. The semantic space hosts the rule engine-based Context Aware Service Execution (CASE) system and the dynamically reconfigurable service logic. The virtual space provides a 3-D virtual model of the physical space, enabling intuitive interaction about the physical space. The 3-D model is built starting from fusion of range data with images and by 3-D reconstruction from a sequence of 2-D images.

[58] proposes an approach for the service robot to detect abnormal state of the elderly in a home environment. The system performs location monitoring through laser range finder (LRF) sensors. The model of home-wide movements of the person is built by analyzing data and extracting location patterns. At run-time location information is used as input to continuously train the model and detect whether person is in abnormal state. [59] presents the architecture of a mobile robot assistant equipped with a facial expression system to increase user's enjoyment. The robot uses ultrasonic rangers and a laser scanner to understand its environment. It also provides event reminder functionalities and nursing care functions such as emergency calls, fall detection and physiological signals monitoring. The user-worn RFID provides temperature measuring and emergency call functionalities. [60] proposes a software-framework which simplifies programming of system skills for a rehabilitation robot interacting with a smart home environment.

3.2.6. Architecture and Design

Some of the smart home researches published in the last years focus on Architecture and Design and are reported below. OSGi seems to be one the most promising frameworks on which to base a smart home architecture. Its use is explicitly reported by [61] and [62], but also by [55], [54] and [63]. The reported researches also identify useful design principles ([11], [64] and [65]) learned from the presented experiences. The best strategy for creating a successful system is to collect requirements directly from final users, as SRS is doing. [66] reports a research done on a group of 430 elderly people in Taipei. The research uses focus groups and questionnaires to define the level of demand for intelligent environment technologies. The results show that a safe environment is the most demanding issue for the elderly to participate in recreational activities.

[61] presents a Connected Home Platform (CHP) and Development Framework for easy design, development and deployment of smart home services. For building applications the

ROCob API Specification is proposed. The networking technology is based on the electrical power line. [62] proposes a framework based on OSGi to facilitate application deployment and management in smart home environments. [63] reports about the architecture and experimentation of an ultrasonic-based indoor location tracking service integrated in the OSGi framework.

[67] presents the user-centered design of an activity management system for elderly people. Ten key design factors are identified and used to design the presented TIVIPOL system. [64] proposes a smart-home environment design based on ZUMA principles: Zero-configuration, Universality, Multi-user optimality, and Adaptability.

[65] presents the design of the user interactions for the multimodal user interface of a smart home system (the Millennium Home). The paper details several lessons learned in interaction design in the form of key factors to be considered in the design of similar systems; some high-level contextual issues are also discussed. The most important one is about *Multiuser interaction*, which is still a topical research issue, even if already in 2000 [68] proposed a system for multi-person tracking using multiple stereo/color cameras. The main point is that *it is not always obvious for a computer system to detect which user has issued a command or requires a system output*. This same issue has security aspects too, because *we need also to be aware that not all passers-by of a ubiquitous system are permitted users*: think about guests for example or even a thief!

3.2.7. Future directions

The only recent criticism we found published about smart homes, or better in this case Ambient Intelligence, is Marc Böhlen's *Second order ambient intelligence* [69]. The author starts reporting a visit to a smart house presentation to demonstrate his statement that "many AmI systems you can actually experience today leave you wondering what all the excitement is about". Then the author tries to imagine the future of Ambient Intelligence assuming that the current research challenges are met and home controls and systems are automatically operated when needed, reliably and invisibly, reacting both to predicted inhabitants' needs and weather conditions.

Böhlen then forecasts that not only technical but also social aspects will be taken into account (social networking features such as those of Facebook for example, but integrated in the home communication and multimedia system). He imagines that gesture based interfaces may evolve into behaviour based UIs and that subtle emotion-recognition systems will be the norm. Also argues that smart activity monitoring systems "that deserve the term" may get to the point that they "should attempt to listen to those who seem to be saying nothing and find some meaning in it". The author finally widens the scope of AmI to the external environment and even dares to push as far as imagining "a distributed sensor network in the Atlantic Ocean that recorded weather data and kept track of maritime traffic, but also listened to whale song and diverted ocean traffic when whales are in the vicinity". Although this may seem a vision for the far future, HP Labs are working on something similar named CeNSE or Central Nervous System for the Earth [70], envisioning "trillions of interconnected inexpensive sensors measuring and processing all sorts of data in real time, to improve safety, sustainability, and security of people and businesses".

3.3. Open Source

In the following sections some open source frameworks for home automation are listed. Further searches can be done in the future to find other open source software for specific functionalities such as activity monitoring.

3.3.1. WOSH Framework

URL: http://wosh.sourceforge.net/
What is it?: WOSH is a scalable, multi-platform, message oriented, home automation software
License: CC By-Nc, Creative Commons Attribution-Noncommercial 3.0 Unported (http://creativecommons.org/licenses/by-nc/3.0/)
Latest version: v 0.8.161 [icarus]
State: This project is actually released for developers, the current stage is not yet ready for end-users.
Features: communication, remote control, entertainment, graphical console, multi-user, role based access, support for X10
Description: WOSH (Wide Open Smart Home) is an open source, multi-platform framework (message oriented middleware) written in ANSI C++, designed to enable (smart) home automation.
WOSH is a Service Oriented Framework (SOA) providing a (Network) OS independent infrastructure for developing component based software (services, aka bundles), but it ships with many implemented services and some end-user applications (as woshsrv, WorkShop).
WOSH and some components are (partially) based on QT4.

3.3.2. MisterHouse

URL: http://misterhouse.sourceforge.net/
What is it?: MisterHouse is an open source home automation program. Written in Perl, it fires events based on time, web, socket, voice, and serial data. It currently runs on Windows 95/98/NT/2k/XP and on most Unix based platforms, including Linux and Mac OSX.
License: Creative Commons Attribution Share-Alike Non-Commercial 2.5 License.
Latest version: v2.105
State: Stable
Features: action execution based on voice input, time of day, file data, serial port data, and socket data; support both wired and wireless (?) X10; supports serial, CPU-XA, IR and other interfaces; supports RSS to provide various logs;can control RoboSapien, ESRA, and ER1 robots; supports xAP and xPL home automation protocols.
Description: MisterHouse is an open source home automation program written in Perl, it fires events based on time, web input, email messages, instant messages, socket messages, voice commands, serial data, bluetooth proximity, infrared signals, X10 and Insteon powerline signals, and many more. It currently runs on Windows 95/98/NT/2k/XP/Vista/7 and on most Unix based platforms, including Linux and Mac OSX.

3.3.3. LinuxMCE

URL: http://www.linuxmce.com/
What is it?: LinuxMCE is an add-on to Ubuntu Linux, presented as a "complete whole-house media solution with pvr + distributed media, and the most advanced smarthome solution available."
License: GPL
Latest version: 8.10

State: Released, Actively developing new versions
Features: (from http://ostoolbox.blogspot.com/2007/05/linuxmce.html)
Media & Entertainment
• 3D alpha-blended GUI optimized for displaying on a TV and using a remote control
• Media browser presenting all content on all devices in the home on a 3D rotating cube
• Plug-and-play detection and aggregation of network storage and DMA's
• Built-in NAS providing centralized backup and whole-house media server
• "Follow Me" Media, each family member's media follows him/her through the house
• Automatically controls all existing av devices, like TV's, Stereo's, etc.)
• Many control options: mobile phone, webpad, pda, phone

Smart Home
• Home Automation: Control lighting, climate, security, camera surveillance, and more
• Communication: Phone system with auto-attendant, voice mail, call forwarding/routing for VOIP and POTS lines
• Security: Uses your existing home alarm, surveillance cameras, lights, phones and tv's to notify you on your mobile phone of any security alerts with the option of reseting the alarm or broadcasting your voice in the house over the tv's

3.4. FP6/FP7 research

Relevant technologies from recent projects
High availability of current technologies spread upon sensor devices, has opened the door to many intelligent home project aiming to secure, control, and support daily life of elders affected by health mainly mental desease.

In the range of devices mainly used, microphones, cameras, health condition like internal pressure, temperature, sweet, etc, localization, electric motors control, white good control, etc.

Many projects offer middleware to seal the plenty of signals collected, and some reasoning systems to infer current status, like neuronal networks and semantic and ontologies.

The project currently discovered have been classified depending on the kind of usage they are addressing toward the elder:

- monitoring and Prevention, to provide alert or corrective actions based on collected information
- Assistive Technology to provide a direct help to Elders
- Independent Life addressing easy and automated control of home devices
- Infrastructure Design addressing the design of complex home device monitoring
- Architecture building middleware that manages and controls parameters flow

Based on the SRS features and on the availability of project results the following projects have been selected, as candidates to SRS integration:

- Isisemd – providing a remote tool to interact with the assisted elder
- Acube – a monitoring environment to manage complex building hosting many patients
- Companionable – a project trying to exploit integration of static domotic environment and mobile robots.
- Oasis – a project ontology.

Project Map

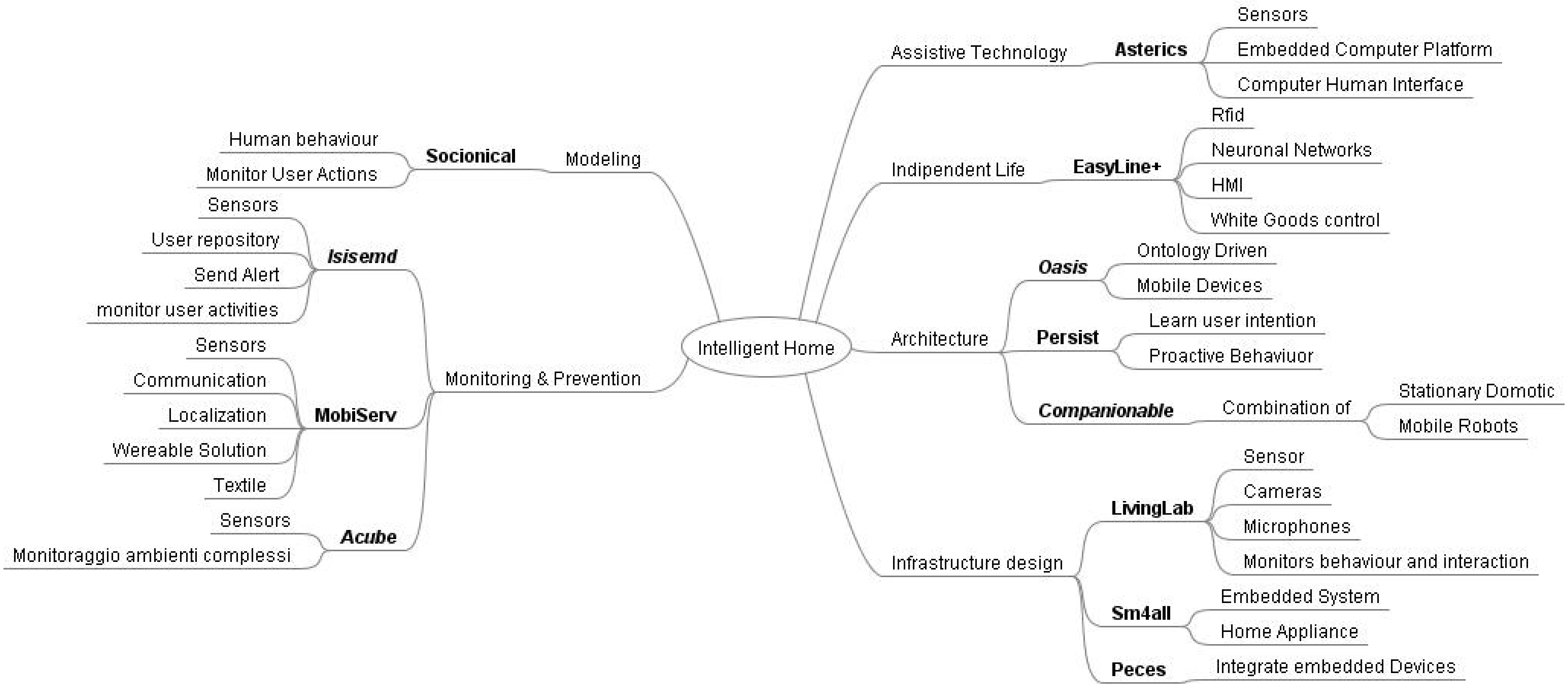


Figure 2 - Mind map of FP6/FP7 projects related to Intelligent Home

3.4.1. Architecture

Project 1 - Companionable

Title: Integrated cognitive assistive and domotic companion robotic systems for ability and security
Start date: 2008-01-01
End date 2011-12-31
Project Reference: 216487
Programme type: Seventh Framework Programme
Subprogramme Area: ICT and ageing
Objective: CompanionAble addresses the issues of social inclusion and homecare of persons suffering from chronic cognitive disabilities prevalent among the elderly, a rapidly increasing population group. Those people need support of carers and are at risk of social exclusion, yet this problem not well addressed by ICT technology, but would lead to a social and economical pressure for staying at home as long as possible.

The main unique selling point of the Companionable project lies in the synergetic combination of the strengths of a mobile robotic companion with the advantages of a stationary smart home, since neither of those approaches alone can accomplish the demanding tasks to be solved. Positive effects of both individual solutions shall be combined to demonstrate how the synergies between a stationary smart home solution and an embodied mobile robot companion can make the care and the care person's interaction with her assistive system significantly better.

Starting with a profound requirements engineering for ICT-supported care and therapy management for the care persons, basic technologies for multimodal user observation and human-machine interaction will provide the fundamentals for the development of a stationary smart home assistive system and a mobile robot assistant, building the cornerstones of the overall system integrating the promising solutions of both parts. Substantial support comes from the research activities focusing on an architectural framework, allowing such a complex care scenario solution be achievable. After the realization of the respective scenarios, long lasting field experiments will be carried out to evaluate and test the system, and both scenarios can be evaluated to show their strength and weaknesses. This will initiate the development of an overall, integrated care scenario (smart home with embedded robot companion).

The realization of this integrated care concept is to be seen as the in-principal vision of Companionable.

Project 2 – Persist

Title: Personal self-improving smart spaces
Start Date: 2008-04-01
End Date: 2010-11-31
Duration: 30 months
Project Reference: 215098
Programme type: Seventh Framework Programme
Subprogramme Area: Service and Software Architectures, Infrastructures and Engineering
Objective: Current trends in the design of pervasive systems have concentrated on the provision of isolated smart spaces via a fixed infrastructure. This is likely to lead to the evolution of islands of pervasiveness separated by voids in which there is no support for pervasiveness. The user experience will be all or nothing, with no graceful degradation from the former to the latter.

The vision of PERSIST is of a Personal Smart Space which is able to provide pervasiveness and context awareness to a user everywhere and all the time. Personal Smart Spaces will be adaptable and capable of self-improvement.

The objective of PERSIST is to develop Personal Smart Spaces that provide a minimum set of functionalities which can be extended and enhanced as users encounter other smart spaces during their everyday activities. They will be capable of learning and reasoning about users, their intentions, preferences and context. They will be endowed with pro-active behaviours, which enable them to share context information with neighbouring Personal Smart Spaces, resolve conflicts between the preferences of multiple users, make recommendations and act upon them, prioritise, share and balance limited resources between users, services and devices, reason about trustworthiness to protect privacy and be sufficiently fault-tolerant to guarantee their own robustness and dependability.

Project 3 – Oasis

Title: Open architecture for accessible services integration and standardisation
Start date: 2008-01-01
End date: 2011-12-31
Duration: 48 months
Project Reference: 215754
Programme type: Seventh Framework Programme
Subprogramme Area: ICT and ageing
Objective: OASIS introduces an innovative, Ontology-driven, Open Reference Architecture and Platform, which will enable and facilitate interoperability, seamless connectivity and sharing of content between different services and ontologies in all application domains relevant to applications for the elderly and beyond.

The OASIS platform is open, modular, holistic, easy to use and standards abiding. It includes a set of novel tools for content/services connection and management, for user interfaces creation and adaptation and for service personalization and integration. Through this new Architecture, over 12 different types of services are connected with the OASIS Platform for the benefit of the elderly, covering user needs and wants in terms of Independent Living Applications (nutritional advisor, activity coach, brain and skills trainers, social communities platform, health monitoring and environmental control), Autonomous Mobility and Smart Workplaces Applications (elderly-friendly transport information services, elderly-friendly route guidance, personal mobility services, mobile devices, biometric authentication interface and multimodal dialogue mitigation and other smart workplace applications).

Applications are all integrated as a unified, dynamic service batch, managed by the OASIS Service Centre and supporting all types of mobile devices (tablet PC, PDA, smartphone, automotive device, ITV, infokiosk,) and all types of environments (living labs, sheltered homes, private homes, two car demonstrators, public transport, DSRT, etc.) in 4 Pilot sites Europewide. As user friendliness and acceptability is a top priority for the project, a user-centred-design approach is followed along the service and application development.

Tested iteratively and thoroughly by hundreds of end users, their caregivers and other stakeholders, the OASIS platform and applications will be optimized and submitted for standardization by the purpose-established OASIS world-wide Industrial Forum.

3.4.2. Infrastructure Design

Project 1 – Living Lab

Title: Design study for the living lab research Infrastructure, to research human interaction with, and stimulate the adoption of, sustainable, smart and healthy innovations around the home
Start Date: 2008-01-01
End Date: 2010-03-31
Duration: 27 months
Project Reference: 212498
Programme type: Seventh Framework Programme
Subprogramme Area: Design studies for research infrastructures in all S&T fields
Objective: The objective of this design study is to address all key issues related to the feasibility of a new research infrastructure with a clear European dimension, named LIVING LAB, that will: Advance the field of user centred research (i.e. by studying the interaction of people with innovations for the ***home***), Test, evaluate and improve innovations for the ***home***, Foster societal needs such as sustainability and quality of life, Stimulate competitiveness of European industry (that brings these innovations to the market).

LIVING LAB will bring together Europe's top research institutes and companies to study the interaction of people with technology in the ***home*** environment, and to stimulate cooperative projects in the fields of user centred research and product development. A LIVING LAB-core infrastructure will look like an ordinary house, but (invisible to its inhabitants, who are all volunteers) it will have sensors, cameras and microphones that record every aspect of ***home*** life. The behaviour and interactions of the volunteers can be monitored at any point in the day throughout the duration of their stay. One key advantage of the LIVING LAB over other simulation setups is that products can be evaluated in a real-life environment, over a prolonged period of time.

Project 2 – SM4All

Title: *Smart* homes for all; an embedded middleware platform for pervasive and immersive environments for-all
Start Date: 2008-09-01
End Date: 2011-08-31
Duration: 36 months
Project Reference: 224332
Programme type: Seventh Framework Programme
Subprogramme Area: Network embedded and control systems
Objective: Embedded systems are specialised computers used in larger systems or machines to control equipments such as automobiles, ***home*** appliances, communication, control and office machines. Such pervasivity is particularly evident in immersive realities, i.e., scenarios in which invisible embedded systems need to continuously interact with human users, in order to provide continuous sensed information and to react to service requests from the users themselves.

The SM4ALL project will investigate an innovative middleware platform for inter-working of ***smart*** embedded services in immersive and person-centric environments, through the use of composability and semantic techniques for dynamic service reconfiguration. By leveraging on P2P technologies, the platform is inherently scalable and able to resist to devices' churn and failures, while preserving the privacy of its human users as well as the security of the whole environment. This is applied to the challenging scenario of private houses and ***home***-care assistance in presence of users with different abilities and needs (e.g., young able bodied, aged and disabled).

The specific composition of the Consortium, consisting of top-class universities and research centres (UOR, TUW, RUG, KTH and FOI), of user partners specialized in domotics and ***home***-care assistance (FSL and THFL) and a SME specialized in specific brain-computer interfaces (GTEC), and of leader companies in the embedded sector (TID and ED) guarantees a widespread dissemination and exploitation of the project results, coupled with a privileged position inside ARTEMIS and ARTEMISIA (due to the presence of UOR, TUW and ED in such bodies).

Project 3 – Peces

Title: *Smart* homes for all; an embedded middleware platform for pervasive and immersive environments for-all
Start Date: 2008-12-31
End Date: 2011-08-31
Duration: 39 months
Project Reference: 224342
Programme type: Seventh Framework Programme
Subprogramme Area: Network embedded and control systems
Objective: The overarching goal of the PECES project is the creation of a comprehensive software layer to enable the seamless cooperation of embedded devices across various smart spaces on a global scale in a context-dependent, secure and trustworthy manner.
The increasing number of devices that is invisibly embedded into our surrounding environment as well as the proliferation of wireless communication and sensing technologies are the basis for visions like ambient intelligence, ubiquitous and pervasive computing.

The benefits of these visions and their undeniable impact on the economy and society have led to a number of research and development efforts. These include various European projects such as EMMA or AMIGO that develop specialized middleware abstractions for different application areas such as automotive and traffic control systems or home automation. These efforts have enabled smart spaces that integrate embedded devices in such a way that they interact with a user as a coherent system. However, they fall short of addressing the cooperation of devices across different environments. This results in isolated 'islands of integration' with clearly defined boundaries such as the smart home or office.

For many future applications, the integration of embedded systems from multiple smart spaces is a primary key to providing a truly seamless user experience. Nomadic users that move through different environments will need to access information provided by systems embedded in their surroundings as well as systems embedded in other smart spaces. Depending on their context and on the targeted application, this can be smart spaces in their vicinity such as 'smart stores' or distant places with a specific meaning such as their home or their office or dynamically changing places.

PECES is committed to developing the technological basis to enable the global cooperation of embedded devices residing in different smart spaces in a context-dependent, secure, and trustworthy manner.

3.4.3. Indipendent Life

Project 1 – Easy Line+

Title: Low cost advanced white goods for a longer independent life of elderly people

Start date:2007-01-01
End date:2009-12-31
Duration:36 months
Project Reference:045515
Programme type: Sixth Framework Programme
Subprogramme Area: Ambient Assisted Living (AAL) in the Ageing Society

Objective: The elderly suffer some disabilities that get worst with the passing years. These disabilities will make carrying out the tasks of an independent life more difficult. It is a reality that the main disabilities (42%) prevent individuals from carrying out home tasks and that, about a fourth part of the household accident are produced in the kitchen, where the "white goods" are key elements. Facing this situation, the project consortium, has decided to carry out the EASY LINE+ project in order to develop prototypes near to market of advanced white goods in order to support elderly persons with or without disabilities to have a longer independent life and will compensate their loss of physical and/or cognitive abilities.

The project foresees using the integrated RFID, Neuronal Networks and HMI technologies to build a system that can capture data of the home environment, and can control via wireless communication (Zigbee) or the mains electricity (EMS PLC), any white good in the home. The users, elderly persons, may actuate by himself any white good in the home, or may leave the "e-servant" to do the actuation. The e-servant will be a white good control system, based on the sensor information and the habits of the user that can program any application without/or with user cooperation. The e-servant, also will be a learning system that detects the loss of abilities of the user and tries to compensate them.

The consortium of this project will be led BSH-E as European level and third at world level in White Goods manufacturing, jointly with R&D experts in new technologies suitable for increasing the functionalities of white goods like I3A (Neuronal Networks, Zigbee, RFID sensors), NEWI (Human Machine Interfaces). Accessibility expert as SBS C-LAB, important industries in RFID applications (IDENT), software/HM (ADSS) and domotic implementation (G2V)

3.4.4. Modeling

Project 1 – Socionical
Title: Complex socio-technical system in ambient intelligence
Start date:2009-02-01
End date:2013-01-31
Duration:48 months
Project Reference: 231288
Programme type: Seventh Framework Programme
Subprogramme Area: Science of complex systems for socially intelligent ICT
Objective: We will develop Complexity Science based modelling, prediction and simulation methods for large scale socio-technical systems. We focus on the specific example of Ambient Intelligence (AmI) based smart environments. A key component of such environments is the ability to monitor user actions and to adjust its configuration and functionality accordingly. Thus, the system reacts to human behaviour while at the same influencing it. This creates a feedback loop and leads to a tight entanglement between the human and the technical system. At the same time there is dynamic, heterogeneous human-human, human-technology, and technology-technology communication leading to ad-hoc coupling between components and different feedback loops. The

project will study global properties and emergent phenomena that arise in AmI based socio-technical systems from such local feedback loops and their coupling on two concrete scenarios: transportation and emergency/disaster.

SOCIONICAL takes a parallel, multi facetted research approach. Thus, we will investigate analytical methods, complex networks based representations, and agent based models. The advances in modelling and prediction will be verified by large scale, distributed simulation driven by real life data. We will develop a methodology by which a small number of instrumented users can be realistically integrated in a large scale simulation as additional 'agents', experiencing the system and driving it. A separate WP is devoted to the integration of different approaches into a coherent framework. Another ensures generalization.

To take into account all technological, psychological and social dimensions and realistic diversity of behaviours we have assembled a multi disciplinary consortium with separate WPs for technology analysis and the modelling of human technology interactions.

SOCIONICAL has a WP devoted to the development and dissemination of guidelines and recommendation for businesses and policy makers.

3.4.5. Monitoring & Prevention

Project 1 – Acube
Title: Ambient Aware Assistant
Start Date: 2008-10-01
End Date: 2011-09-30
Duration: 36 months
Project Reference: 248434
Programme type: Bando Grandi Progetti 2006 - Italy
Objective: Improving the quality of life for the elderly and disabled through technological progress. That is the goal of project Acube. The project's goal is to study technologies for monitoring complex environments that can be applied in areas such as assisted living homes to help personnel, as well as to support the independence and safety of users. Using distributed sensor networks it will be possible to monitor the areas of a rehabilitation center (e.g., an institution for Alzheimer sufferers) or similar facilities in order to detect events, situations, and activities even in complex scenarios with many people. The base technology is a distributed sensor network (DSN) for collecting environmental data, which is connected to a computing system able to comprehend perceived changes and to develop and appropriate response.
The major technical outcome of the project is the development of a monitoring system that requires the integration of a wide variety of heterogeneous technologies (video, audio, rfid, wsn, biomedical). The development of advanced algorithms to recognize events, situations, activities, behaviors in complex multi-person scenarios will enable the smart environment to understand who is doing what, where, when and how. This knowledge allows the system intelligence to take decisions (e.g. rising alarms). Processing includes adaptation capabilities, to fit different environments and users. ACube will need to configure itself automatically and operate intelligently according to the data it senses. In the future, the system could be exported to different application domains, such as the intelligent monitoring and surveillance of public spaces (museums, schools, stations).

Project 2 – MobiServ
Title: An integrated intelligent home environment for the provision of health, nutrition and mobility services to the elderly

Start Date: 2009-12-01
End Date: 2012-11-30
Duration: 36 months
Project Reference: 248434
Programme type: Seventh Framework Programme
Subprogramme Area: ICT & Ageing
Objective: Life expectancy increases, and the wish to prolong independent living remains strong. The objective of the MOBISERV project is to develop a proactive personal service robotics for supporting independent living. The project will develop a personalized system, orchestrating vital signs recording and analysis, warnings, and alerts to health and emergency assistance networks. Existing solutions are closed to external developers and address only few problems pertinent to the elderly. MOBISERV will deliver a robotic prototype of an open standard-based personal platform capable of sensing the user s personal environment and adapting to the user s patterns of behaviour.

By early detection of threatening environmental and/or emerging medical conditions, harmful consequences will be mitigated by issuing warnings and providing guidance; in case adverse events cannot be evaded, alarms will be issued. The platform will be an integration of innovative components delivered by the project and of existing standards-compliant technologies. Innovative wireless (bio-) sensor-actuators, localisation and communication technologies, smart textiles and clothing and a wearable solution hosting monitoring equipment will be integrated into an existing robotic platform capable of self-learning and able to support elderly in indoor contexts. Tele-alarm applications will be developed to enhance health and integrated care services. A user-centred participatory design process will be adopted, with iterative design and evaluation. The system will be evaluated under real life conditions.

Project 3 – Isisemd
Title: Intelligent System for independent living and self-care of seniors with cognitive problems or mild dementia
Start Date: 2009-03-01
End Date: 2011-08-31
Duration: 30 months
Project Reference: 238914
Programme type: ICT Policy Support Programme
Subprogramme Area: ICT for user friendly administrations, public services and inclusion
Objective: The aim of the ISISEMD project is to provide a pilot of innovative intelligent set of scalable services that will support the independent living of elderly people in general and in particular the group of elderly with cognitive problems or mild dementia and at the same time to support the formal and informal caregivers in their daily interaction with the elderly. The services will improve the elderly ability for self-care by support for their basic daily activities in way that prevents health risks in their homes. The services will also strengthen the daily interaction with their social sphere - partners and relatives, friends and care-givers, giving them the feeling of safety and preventing their social isolation. Last but not least, their cognitive training and activation will be strengthened. To prove wide applicability in Europe, the pilot will be validated and tested in realistic conditions for 12-month period in four Member States regions which have extensive experience from existing telehomecare services for elderly. The pilot set of services will integrate: a) several partial services, b) already tested prototype, and c) completed R&D work. The operation will be evaluated with three target end-user groups - elderly, formal and informal caregivers, addressing thus the diverse requirements of these groups. The pilot service will contain 3 different service bundles (basic services, intermediate and high level) that allow for escalation of the service provided to the end-users based on their needs and providing different pricing schemes. The

ISISEMD pilot has strong user centric focus and offers a wide range of individual activities that will allow for the e-Inclusion of the elderly people in every day social life. In the final period of the project, usability, business and cost-benefit analysis of the pilot will be carried out. ISISEMD is an expert consortium of 12 partners, built on public-private partnership.

3.4.6. Assistive Technology

Project 1 – Asterics
Title: Assistive technology rapid integration and construction set
Start Date: 2010-01-01
End Date: 2012-12-31
Duration: 36 months
Project Reference: 247730
Programme type: Accessible and Assistive ICT
Subprogramme Area: ICT for user friendly administrations, public services and inclusion
Objective: More than 2,6 million people in Europe have problems with their upper limbs and therefore many of them depend on Assistive Technologies (AT). As the potential of the individual user is very specific, adaptive, ICT-based solutions are needed to let this population group participate in modern society. Such solutions are rarely available on today's market.

AsTeRICS will provide a flexible and affordable construction set for realising user driven AT by combining emerging sensor techniques like Brain-Computer Interfaces and computer vision with basic actuators. People with reduced motor capabilities will get a flexible and adaptable technology at hand which enables them to access the Human-Machine-Interfaces (HMI) at the standard desktop but in particular also of embedded systems like mobile phones or smart home devices.

AsTeRICS will implement a set of building blocks for the realisation of AT:
- Sensors which allow the individual to exploit any controllable body or mind activity for interacting with HMI
- Actuators for interfacing to standard IT, to embedded systems and to interact with the environment
- An Embedded Computing Platform that can be configured to combine sensors and actuators to tailored AT-solutions which support the full potential of an individual user

The core of the software suite will be provided as Open Source. The complete system will be affordable for many people who cannot benefit from leading edge supportive tools today.

AsTeRICS revolutionises the concept of AT: AT today mostly focuses on a certain task or situation. Due to the growing importance of the PC, AT has been oriented towards standard Human-Computer (HCI) or desktop interfaces. AsTeRICS respects the strong need for flexible, adaptable AT functionalities accompanying people with disabilities away from the desktop, enabling them to interact with a divers' and fast changing set of deeply embedded devices in our modern environment.

4. Middleware's, Communication and Open Software Robotic Frameworks/Software Architectures

4.1. Introduction

Modern robot systems often operate under very complicated conditions. In order to complete their tasks, robots need to communicate with operator(s) and/or each other. Robot communication takes place whenever data/command transfer is required between robots, robots and human operator(s) and even between several modules on the same robot system (internal communication).
Today's aspects of robot communication require transfer of data and commands between various kinds of hardware platforms, operating systems and applications. In order to achieve a versatile concept of robot communications it is very useful to build communications middleware, helping to interconnect and interoperate all components of the system.
Another important aspect of Robot Communications is Teleoperation. It allows the Human Operator (HO) to remotely control one or more robots using some sort of physical user interface device (such as a joystick) and set robots' position, speed or both.
Robot Communications is normally integrated into the Robotics Frameworks/Software Architectures. This report focuses on Open Source Robotic Framework. The framework can be defined as a group of open source software packages that simplifies programming of robotic devices, thus providing a set of reusable components, a unified programming environment or common drivers or facilities for existing robotic hardware.

A deep survey of development environments for autonomous mobile robots is done at [16].

4.2. Literature research

Next, a list of the most important open software robotic frameworks currently available is presented.

4.2.1. OpenJAUS (Open Joint Architecture for Unmanned Systems) [71]

Purpose: Support the acquisition of Unmanned Systems by providing a mechanism for reducing system life-cycle costs. This is accomplished by providing a framework for technology reuse/insertion.

Technical constraints:

- Platform Independence
- Mission Isolation
- Computer Hardware Independence
- Technology Independence

4.2.2. ORCA [72]

Purpose: Orca is an open-source suite of tools for developing component-based robotic systems. It provides the means for defining and developing components which can be pieced together to form arbitrarily complex robotic systems, from single vehicles to distributed sensor networks. In addition it provides a repository of pre-made components which can be used to quickly assemble a working robotic system

Technical constraints:

- Little flexibility with regard to the implementation platform

4.2.3. OROCOS (Open Robot Control Software) [73]

Purpose: The Open Robot Control Software project provides a Free Software toolkit for real-time robot arm and machine tool control. Consists of two decoupled but integrated sub-projects:

- Open Real-time Control Services.
- Open Robot Control Software.

Technical constraints:

- The Orocos project seems to contain fine C++ libraries useful for industrial robotic applications and is focused on control software

4.2.4. ROS (Robot Operating System – Robot Open Source) [74]

Purpose: ROS is an open-source, meta-operating system for your robot. It provides the services you would expect from an operating system, including hardware abstraction, low-level device control, implementation of commonly-used functionality, message-passing between processes, and package management. It also provides tools and libraries for obtaining, building, writing, and running code across multiple computers

Technical constraints:

- No real limitations can be found except that ROS comes from a US based company (California).

4.2.5. PLAYER [75]

Purpose: The Player Project creates Free Software that enables research in robot and sensor systems. According to the Player Project, the Player robot server is probably the most widely used robot control interface in the world. Its simulation back-ends, Stage and Gazebo, are also very widely used. Released under the GNU General Public License, all code from the Player/Stage project is free to use, distribute and modify. Player is developed by an international team of robotics researchers and used at labs around the world.

Technical constraints:

- It is mostly US funded by NSF, DARPA and JPL and supported by US research institutions

4.2.6. MICROSOFT ROBOTICS [76]

Purpose: According to Microsoft, Microsoft Robotics products and services enable academic, hobbyist and commercial developers to easily create robotics applications across a wide variety of hardware.

Technical constraints:

- Dependency on Microsoft development tools

4.2.7. CLARAty (Coupled-Layer Architecture for Robotic Autonomy) [77]

Purpose: CLARAty is an integrated framework for reusable robotic software. It defines interfaces for common robotic functionality and integrates multiple implementations of any given functionality. Examples of such capabilities include pose estimation, navigation, locomotion and planning. In addition to supporting multiple algorithms, CLARAty provides adaptations to multiple robotic platforms. CLARAty, which was primarily funded by the Mars Technology Program, serves as the integration environment for the program's rover technology developments.

Technical constraints:

- Public access seems to be limited.
- The license and download policy has critics.
- CLARAty is incompatible with the GPL and cannot be used for commercial activities.

4.2.8. YARP (Yet Another Robot Platform) [78]

Purpose: It is a set of libraries, protocols, and tools to keep modules and devices cleanly decoupled. It is reluctant middleware, with no desire or expectation to be in control of your system. YARP is definitely not an operating system.

Technical constraints:

- Yarp / RoboCub were supported by European Union grant RobotCub (IST- 2004-004370) and by euCognition (FP6 Project 26408). These excellent projects have ended.

4.2.9. CARMEN (Carnegie Mellon Robot Navigation Toolkit) [79]

Purpose: CARMEN is an open-source collection of software for mobile robot control. CARMEN is modular software designed to provide basic navigation primitives including: base and sensor control, logging, obstacle avoidance, localization, path planning, and mapping.

Technical constraints:

- C programming language
- No graphical tools
- Not vision/speech processing

4.2.10. MOOS (Mission Oriented Operating Suite) [80]

Purpose: MOOS is a C++ cross platform middle ware for robotics research. It is helpful to think about it as a set of layers.

- Core MOOS - The Communications Layer: The most fundamental layer CoreMOOS is a very robust network based communications architecture (two libraries and a lightweight communications hub called MOOSDB) which for very little effort lets you build applications which communicate with each other.
- Essential MOOS - Commonly Used Applications: Essential MOOS is a layer of applications which use CoreMOOS. They offer a range of functionality covering common tasks for example process control, logging

Technical constraints:

- Oriented to autonomous marine vehicles

4.2.11. RoboComp [81]

Purpose: RoboComp is an open-source robotic software framework. It uses software component technology to achieve its goals: efficiency, simplicity and reusability. Its components can be distributed over several cores and CPU's. Existing software components, can be easily integrated with new components made by RoboComp users.

Technical constraints:
- Rough list of common software dependences
- Communication depends on the ICE framework
- Still under development

4.2.12. MARIE [82]

Purpose: MARIE is a free software tool using a component based approach to build robotics software systems by integrating previously existing and new software components.
MARIE's initiative is based on the following main requirements:
- Reuse softwares, APIs, middlewares and frameworks frequently used in robotics (Player, CARMEN, RobotFlow, etc.)
- Adopt a rapid-prototyping approach to build complete system
- Allow distributed computing on heterogeneous platforms
- Allow concurrent use of different communication protocols, mechanisms and standards
- Accelerate user-defined developments with well defined layers, interfaces, frameworks and plugins
- Support multiple sets of concepts and abstractions

Technical constraints:
- Low level communications partially supported
- No security provided
- Incomplete documentation

4.2.13. Miro: Middleware for Autonomous Mobile Robots

Research topic: Design of a robot programming framework.
Main focus of research: Creating a new robot programming framework that allows for a rapid development of reliable and safe software on heterogeneous computer networks and supports the mixed use of several programming languages.

4.2.14. Design and Evaluation of Communication Middleware in a Distributed Humanoid Robot Architecture

Research topic: The role of a communications middleware in the distributed robot architecture for implementing tasks on humanoid robots.
Main focus of research: Design of an architecture serving various functional roles and information exchange within a distributed system, using three different communication subsystems: the Cognitive Map (CogMap), Distributed Operation via Discrete Events (DiODE), and Multimodal Communication (MC).

4.2.15. An Adaptive Middleware for Context-Sensitive Communications for Real-Time Applications in Ubiquitous Computing Environments

Research topic: An object-based adaptive middleware providing a well-defined development framework and proper runtime services in order to facilitate context-sensitive communications.

Main focus of research: A reconfigurable context-sensitive communications middleware that provides a special context-aware interface definition language for specifying context-sensitive interfaces of real-time objects, an object container framework for generating interface-specific context-analyzers, and a context-sensitive object request broker for context-sensitive object discovery and impromptu connection management.

4.2.16. Experimental Survey Results For Position, Speed And Position-Speed Modes Of Control In Systems For Remote Control Of Mobile Robots

Research topic: Experimental survey results for position, speed and position-speed modes of control in systems for remote control of mobile robots

Main focus of research: The research emphasises on test results from a survey about application of three modes of Mobile Robot (MR) remote control – position control, speed control and mixed (position-speed) control. The test setup assumes a human operator (HO) that controls one or more mobile robots using a Remote Control (RC). In most practical cases, the Speed Control mode is used for Mobile Robots, which means that HO's control interface (i.e. a joystick) position determines the speed of the controlled robot. This corresponds to the practically unlimited area of mobility of MR. The position control is used mostly for controlling Robot Manipulators (RM) which have very limited area of mobility but need precision position control. In this case, the HO's control interface determines the position of the RM.

4.2.17. A User Study of Command Strategies for Mobile Robot Teleoperation

Research topic: The article represents a user study of mobile robot teleoperation.

Main focus of research: Performance of speed, position and combined command strategies in combination with text, visual and haptic feedback information are evaluated by experiments. Two experimental tasks are designed as follows: positioning of mobile robot and navigation in complex environment. Time for task completion and motion accuracy are measured and compared for different command strategies and types of feedback. Role of haptic, text and visual feedback information in combination with described command strategies is outlined.

4.3. FP6/FP7 research

Some FP6 projects have dealt with the robotic software architecture issue. Next are summarized the most relevant, ROBOTCUB and ROSTA.

4.3.1. ROBOTCUB (ROBotic open-architecture technology for cognition, understanding and behaviours) [83]

Robot-cub scientific goals were:

- to create and open physical platform for embodied research that can be taken up and used by the research community involved in embodied cognition, and
- to advance our understanding of several key issues in cognition by exploiting this platform in the investigation of several cognitive capabilities.

To achieve these goals an embodied system was constructed able to learn how to interact with the environment through manipulation and gesture production/interpretation, and how to develop its perceptual, motor and communication skills to perform goal directed manipulation tasks.
The project studied the cognition through the implementation of a humanoid robot the size of a 3.5 year old child: the iCub [84]. This humanoid has been subsequently adopted by more than 20 laboratories worldwide.
Within the project the YARP [78] software framework (previously commented) was developed as the middleware for the iCub robot.

This project started on 2004-09-01 and finished on 2009-08-31.

4.3.2. ROSTA (Robot standards and reference architectures) [85]

The objective of RoSta was to proactively take the initiative on the definition of formal standards and the establishment of "de facto" standards in the field of robotics, especially advanced service robotics. The project was intended to take the initiative in the formulation of standards in selected key topics which are regarded to have high impact on future service robotics research and development: glossary/ontologies, robot architectures, middleware and benchmarks. The topics have been started to form the root of a whole chain of standard defining activities going beyond the specific activities of RoSta. Major RoSta-activities and results can be summarized as follows:

- Creation of a glossary/ontology for mobile manipulation and service robots
- Specification reference architecture principles for mobile manipulation and service robots.
- Specification of a middleware for mobile manipulation and service robots.
- Benchmarking.

This project started on 2007-01-01 and finished on 2008-12-31.

5. Safety

5.1. Introduction

Safety is a big issue in robotics because of the partial autonomy systems are supposed to have according to the definition. Robots systems operating autonomous in the direct environment and/or in close distance to humans have a clear potential danger for the humans involved. Actions of the robot may be unintended harmful for human or simply contra-productive. Failure of the system or a discrepancy between actions required on the basis of the system senses and what is actually required may lead to dangerous behaviour of the system. In the design of systems a different approach must be taken regarding safety for industrial robots which operate in an human-restricted zone and for healthcare robots which operate in an environment with vulnerable humans per se. Main difference to classical (industrial) applications is that user(s) are in the working area of the robot and thus most of the standard safety rules known in robotics are not applicable. Anyway and due to the lack of particular safety standards for service robots at least some of the rules from industrial robotics standards are usually considered in order to at least define a minimum standard of safety. Many systems are – for example - designed to do nothing unless specifically triggered to do something by a human operator (e.g. dead man switch). This makes, for instance, voice control potentially dangerous since it is not possible to keep on instructing the system continuously.
Another approach used sometimes is to (partly) consider rules from medical technology development (i.e. rules set in Medical Device Directive MDD and/or similar standards) even if the target system later is not meant to be certified as a medical system. Such approaches are mainly used if the application area has more medical aspects, like for robotic setups in rehabilitation or surgery. This is also connected to one other observation: if the application area of the investigated system do not have such medical aspects, safety analysis quite often is not sufficiently included in the development process. There are only few projects (like e.g. FP6-IST IROMEC – www.iromec.org) where safety analysis is fully integrated into the design and realisation process.
A new standard for (non-medical) service robots will be available soon – this standard for sure will be a much better guideline for development of assistive robot systems, like the SRS system. It is the declared goal for the SRS project to integrate regulations from this standard as much as it turns out to be possible and applicable.

SciFi and Reality

Discussion of robot safety usually also includes a reflection of the “Three Laws of Robotics” from SciFi-author Isaac Asimov [86]. He was one of the earliest thinkers to explore the implications of robots as autonomous, intelligent creatures, equal (or superior) in intelligence and abilities to their human masters. Asimov wrote a sequence of novels analyzing the difficulties that would arise if autonomous robots populated the earth. He realized that a robot might inadvertently harm itself or others, both through its actions or, at times, through its lack of action. He therefore developed a set of postulates that might prevent these problems. He postulated three laws of robotics – mainly dealing with the interaction of robots and people - and wrote a sequence of stories to illustrate the dilemmas that robots would find themselves in, and how the three laws would allow them to handle these situations. From those three laws in particular Law 1 – “A robot may not injure a human being, or, through inaction, allow a human being to come to harm” is being referred to quite often. [87]
This first “law” could be labelled with “safety”. As a result, most (all?) robot systems today are designed with (multiple) safeguards in order to minimize the likelihood that they can harm by their action. The “law” also includes a second aspect, which sometimes is not considered that much – but which is very important especially for assistive systems like SRS: the robot also should not injure

(in the broader meaning of the word) persons due to inaction. This second part of the law – do not allow harm through inaction – is quite difficult to implement. If determining how a machine's actions might affect people is difficult, trying to determine how the lack of an action might have an impact is even more so. This would be a reflective level implementation, for the robot would have to do considerable analysis and planning to determine when lack of action would lead to harm. The later phases of risk analysis and safety system design however should also address this aspect for the SRS robot setup.

5.2. Possible direct and indirect hazards of the robot to the environment

Independent from the application and from the implemented way of risk and safety management, a first step necessary is to identify possible hazards of the robot. The following grid (not claiming completeness) – also being influenced by the risk management process defined by EN ISO 14971 -- could be useful for this phase and will be also used for SRS.

Stored energies, available forces:

- Kinetic energy (when the robot or robot parts are moving)
- Potential energy (when the robot can fall down or tip)
- Power supply (batteries)
- Chemical stored energy (in case of fire)
- Radiation (even if it is unlikely that the robot uses strong radio waves or lights - or even radioactive material)
- Pressure (sound waves)
- Forces of the electro motors
- Confusion of control devices (especially in remote control phase)

Hazard to physical health:

- Parts of the robot (including accessories) could be toxic
- Parts could be sharp
- Direct physical contact between device and user
- Transfer of energy or substances between robot and user
- Cleaning of the robot

Hazard to mental health:

- Confusion of the (primary) user caused by unexpected behaviour and/or appearance of the robot
- Short term: panic
- Long term: effects by activities with robot ("robot-addiction", dependency)

Energy in the environment of the robot (outside of the robot):

- Potential energy (triggering other things to fall down or tip, making people fall)
- Kinetic energy (e.g. manipulating objects which robot then might loose)
- Other energies (not in robot) - e.g. robot throws other electrical appliances into the bath tub, trigger fire extinguisher, etc.
- Electro-magnetic radiation?
- Insufficient power supply
- Chemical energies (robot causes a fire by materials outside of the robot)
- Robot controls other devices?

Hazards resulting from the use of the system:

- Errors in transmission of remote commands and/or feedback
- Different expectations/reactions between primary user and remote operator
- Use by untrained or unpractised persons
- Reasonable unintended use
- Insufficient warning about adverse effects
- Multi-use aspects?
- Wrong measurement of environment
- Improper, inadequate or complicated HMI
- Loss of mechanical/electrical integrity (for HMI system)

5.3. Goals of hazard analysis

When the hazard analysis can be done based on realistic data of the robot, it can be determined how many people are affected by hazards of the robot and what injuries (or even how many fatalities) can be expected in worst case. This will lead to a SIL (safety integrity level).
Based on the hazards of a given real robot one can find out how the risk can be reduced below a tolerable limit. Risk reduction can be done by technical measures in the first way, or in a non - technical way, e.g. by a trained person with an emergency-off switch continuously supervising the robot.
In any case result of safety analysis suggests that the robot shall use the lowest possible energy. As an essential contribution towards this goal the robot should be constructed as light-weight as possible so that - under normal operating conditions - its mechanical energy represents no hazard at all.

Estimation of safety integrity level

IEC 61508-5 figure D2 shows a “risk graph” for determining the safety integrity level (see figure below).

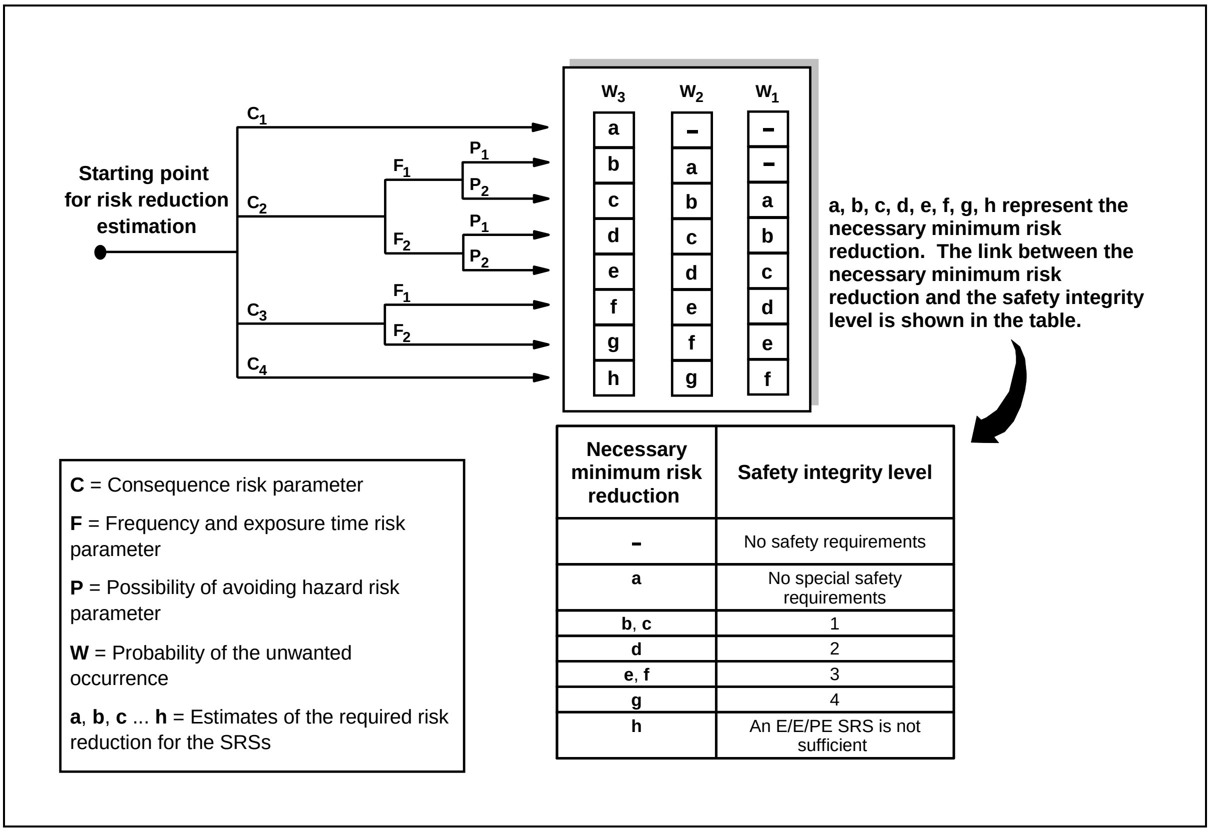


Figure: risk graph according to ICE 61508-5

- When the consequence is “minor injury” a “necessary minimum risk reduction” with “no special safety requirements” is sufficient.
- When the consequence is “serious permanent injury to one or more persons, death to one person” a “necessary minimum risk reduction” starting from “no special safety requirements” up to SIL 1,2 or even 3 is necessary, according to the exposure time in the hazardous zone, the possibility of avoiding the hazardous event and the probability of the unwanted occurrence.
- When the consequence is “death to several people” a “necessary minimum risk reduction” of SIL 2 to SIL 4 is necessary, depending on to the exposure time in the hazardous zone, the possibility of avoiding the hazardous event and the probability of the unwanted occurrence.

5.4. FP6/FP7 research

As mentioned previously literature research shows very different levels of safety analysis and risk management in research projects in the given application area – from a very superficial investigation of safety up to a fully integrated risk management in other projects. In general terms safety aspects are of more concern if the application is closer to classical “medical” applications, and are not that much elaborated for many applications in service robotic area. The “minimum standard” for many investigated projects is the application of rules and requirements from industrial robot safety regulations – as far as applicable.

For safety analysis and risk management there are different established methods available. Many projects are using methods known from the Medical Device Directive or related regulations. The process here starts with a description of the intended use of the system, followed by a detailed description of possible hazards and risks (see above). Finally, a more detailed risk analysis (FTA, FMEA, HAZOP) helps to identify the potential risks and helps to identify measures in order to keep these risks to an acceptable level.

It should me mentioned here that for the majority of the investigated projects the most mentioned and analyzed risk is the one of collision between moving parts of the robot and user/environment (e.g. [88], [89]). For SRS this risk is even more considerable, as the manipulating arm also is mounted to a moving platform, which means that there are further risky combinations from moving/non-moving arm and platform. Another important aspect related to collision/ unwanted physical interaction is for consequences of change of potential energy (e.g. part manipulated by robot falls down) and kinetic energy (e.g. part manipulated by the robot gets loosened and slips out of gripper). In [90] the authors introduce a very useful classification of *levels of physical interaction with users* as follows:

Level 0 This level has no physical interactions other than communications. Examples include timers, medication reminders, monitors, alarms and tele-links. Such devices have no robotic features, i.e. no spatial machinery, and there are no special safety concerns, other than those of any electrical or computerized home appliance. This level has recently been a popular application area for intelligent agent technology.

Level 1 Devices at this level are able to move within the user’s environment but should generally avoid physical contact with users and even avoid the user’s personal space where possible. Examples are autonomous vacuum cleaners, lawn mowers, tour guides and couriers. Extra safety considerations must be given to events such as the user accidently colliding with or falling over the robot.

Level 2 This level covers devices that can intimately share the same operating space as the user. Examples would be robot devices that fetch or retrieve objects, assist with domestic tasks such as cooking or ironing, or perform specific manipulation tasks on a desk top or work area. Some form of cooperation between user and robot is necessary for normal operation and extra safety issues include the specific hazards caused by powered kinematic devices operating near human users.

Level 3 The highest level of physical interaction occurs when the device is in frequent contact with the user during operation. Examples include rehabilitation devices that exercise the user’s limbs, powered walking aids, and motorized wheelchairs. Safety is very important because of the close-coupling with the user and failures could easily lead to injury.

In the following some examples for investigations in the area of safety for Human-Robot Interaction are given.

PHRIENDS Project:

The PHRIENDS (Physical Human-Robot interaction: depENDability and Safety) [91] project is about developing key components of the next generation of robots, including industrial robots and assist devices, designed to share the environment and to physically interact with people. The philosophy of the project proposes an integrated approach to the co-design of robots for safe physical interaction with humans, which revolutionizes the classical approach for designing industrial robots – rigid design for accuracy, active control for safety – by creating a new paradigm: design robots that are intrinsically safe, and control them to deliver performance.

PHRIENDS aims to deliver new actuator concepts and prototypes, new dependable algorithms for human-friendly robot motion planning, new control algorithms for handling safe human-robot physical interaction, and will integrate these components in meaningful subsystems for experimental testing, quantitative evaluation and optimization.

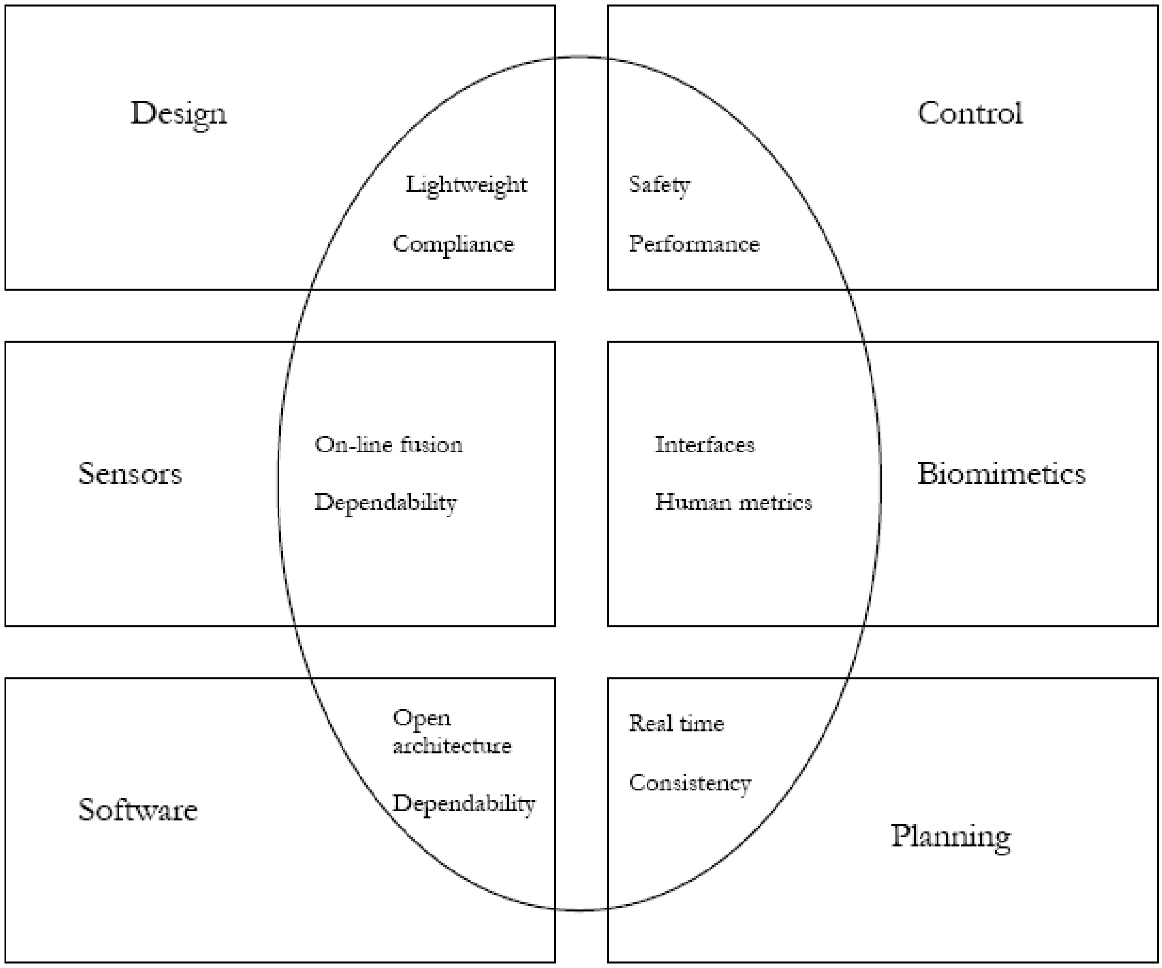


Figure 3: A „map“ of robotics for anthropic domains: main issues and superposition for pHRI, taken from [6]

Safety and dependability are the keys to a successful introduction of robots into human environments. The project claims, that very compliant transmissions may ensure safe interaction, but may be inefficient in transferring energy from actuators to the links for their fast motion. An approach to gain performance for guaranteed safety joint actuation described in the project is to allow the passive compliance of transmission to vary during the execution of tasks. The Variable Impedance Approach (VIA) is a mechanical/control co-design that allows varying rapidly and continuously during task execution the value of mechanical components such as stiffness, damping, and gear-ratio, guaranteeing low levels of injury risks and minimizing negative effect on control performance. [91] describes further that the definition of natural robot motions (e.g., for human approaching tasks) should explicitly address the issues of predictability and legibility from the human user perspective. Motion planning and control for a robot in close vicinity of human beings must not only produce non-collision trajectories. The intrinsic nature of service robotics is to deal with unstructured, time-varying environments, for which a model is hardly available. The major part of dynamicity is typically due to the unpredictable motion of a human user. Therefore, the integration of a sensor-based online reactivity component into an off-line motion plan (needed for a global analysis of the scene) seems mandatory.

Finally [91] mentions that dependability in pHRI (physical Human-Robot Interaction) during normal operation is threatened by different kinds of potential failures of unmodeled aspects in sensors, control/actuation systems, and software architecture, which may result in undesirable behaviors. Due to the critical nature of pHRI, dependability must be enforced not only for each single component, but for the whole operational robot. In all pHRI situations, safety or robot operation is essential, given the presence of humans in contact with or in the vicinity of the robots.

In this context, safety can be rephrased as ”absence of injury to humans in the robot’s environment”. Safety needs to be ensured both during nominal operation of the robot, as well in the presence of faults. In particular, it should be accepted that, in order to enforce a robot operation which is safe for the human, the completion of a programmed task may even be abandoned (this is also named survivability). The construction of a good model of humans interaction with the robot is certainly one of the main purposes of a sensory system for pHRI: vision and other proximity sensors must be able to map the position of critical actors in the scene. These sensors must be robust to changing of environmental conditions like lighting, dust and other sources of uncertainty.
For closer human-robot interaction, such as supporting people, handing over object or shaking hands –all physical events, force/torque sensors and small “tactile” sensors distributed along the structure may be needed (just like in dexterous robotic hands). Of course, it is necessary to distinguish between intentional contacts and accidental collisions with human body and hand.

MOVEMENT Project:
Aim of MOVEMENT [92] was the development of a modular robot system in order to address the three aspects of mobility. Main component of the system was a mobile robot platform which was able to automatically connect to “application modules” on demand. Different interaction types have been implemented – like fully autonomous operation (e.g. during moving to the desired application module and docking), semi-autonomic movement (including a user intention estimation component) and fully manual operation (e.g. by wheelchair-joystick).
Safety analysis in MOVEMENT has been performed strictly following the regulations from Medical Devices Directive (MDD) including a detailed FMEA for risk management. In addition, related standards for electric wheelchairs (EN 12184) and others have been analyzed and integrated into the development process.

IROMEC Project:
In IROMEC project [93] a robot system was designed which is being used as a mediator for disabled children in playful environment. Interaction and safety has been of paramount importance for this project, as the direct physical contact between robot and disabled children is one of the important features for the implemented play scenarios – but which on the other hand is a very risky situation in general. Due to the high-risk application safety analysis in IROMEC has been integrated during the entire development process. Robot prototypes have been analyzed concerning their conformity to related toy standards – a detailed HAZOP analysis was running in parallel with the design process and gave many important proposals to increase safety level of the IROMEC robot system.

5.4.1. Standards, Regulations

Due to the high complexity of the SRS system a number of standards and regulations are related to system safety and security. This chapter gives a first list of (more or less) related standards. This list does not claim completeness now and needs to be updated during the setup phase of the SRS project.

ISO 10218-1 - Robots for industrial environments -- Safety requirements -- Part 1: Robot
ISO 10218-1 specifies requirements and guidelines for the inherent safe design, protective measures, and information for use of industrial robots. It describes basic hazards associated with robots, and provides requirements to eliminate or adequately reduce the risks associated with these hazards.

ISO 10218-1 does not apply to non-industrial robots although the safety principles established in ISO 10218 may be utilized for these other robots.

ISO EN 6385 – Ergonomic principles in the design of work systems
This standard establishes ergonomic principles as basic guidelines for the design of work systems.

ISO 13849 – Safety of machinery – Safety-related parts of control systems – Part 1: General principles for design
This standard provides safety requirements and guidance on the principles for the design of safety-related parts of control systems. For these parts it specifies categories and describes characteristics of their safety functions. This includes programmable systems for all machinery and for related protective devices. It applies to all safety related parts of control systems, regardless of the type of energy used, e.g. electrical, hydraulic, pneumatic, and mechanical. It applies to all machinery applications for professional and non-professional use. It may also be applied for control systems with safety related parts used for other technical applications.

IEC 62061, Safety of machinery – Functional safety of safety-related electrical, electronic and programmable electronic control systems
This International Standard specifies requirements and makes recommendations for the design, integration and validation of safety-related electrical, electronic and programmable electronic control systems (SRECS) for machines. It is applicable to control systems used, either singly or in combination, to carry out safety functions on machines that are not portable by hand while working, including a group of machines working together in a co-ordinated manner. This standard does not cover all the requirements (e.g. guarding, non-electrical interlocking or non-electrical control) that are needed or required by other standards or regulations in order to safeguard persons from hazards. Each type of machine has unique requirements to be satisfied to provide adequate safety.

ISO 13582 – Safety of machinery – Safety distances to prevent danger zones being reached by the upper limbs
This standard gives values for safety distances to prevent danger zones being reached by the upper limbs of persons of 3 years of age and above without additional aid. The distances apply when adequate safety can be achieved by distances alone.

ISO 13854 – Safety of machinery – Minimum gaps to avoid crushing of parts of the human body
This international standard provides parameters based on values for hand/arm and approach speeds and the methodology to determine the minimum distances from sensing or actuating devices of protective equipment to a danger zone.

ISO/DIS 13857 – Safety of machinery – Safety distances to prevent danger zones being reached by upper and lower limbs (Category B)
This international standard establishes values for safety distances in both industrial and public environments to prevent machinery hazard zones being reached. The safety distances are appropriate for protective structures. It also gives information about distances to impede free access by the lower limbs. It is applicable for people of 1,4m body height and above (this includes at least the 5th percentile of persons of 14 years and older). In addition, for upper limbs only, it provides information for children older than 3 years where reaching through openings needs to be addressed. The clauses of the international standard covering lower limbs apply when access by the upper limbs is not foreseeable according to the risk assessment. The safety distances are intended to protect those persons trying to reach hazard zones under the conditions specified. This international standard does not provide safety distance information for Electro-Sensitive Protective Equipment

(ESPE). This international standard need not be applied to machinery which is covered by certain standards in which specific testing procedures for safety distances are laid down, for example using the test finger in electrical applications.

ISO 14118 – Safety of machinery – Prevention of unexpected start-up (Category B / Harmonized standards: EN 1037)
This international standard specifies designed-in means aimed at preventing unexpected machine start-up to allow safe human interventions in danger zones. This standard applies to unexpected start up from all types of energy sources.

ISO 14211 – Safety of machinery – Principles of risk assessment (Category A / Harmonized standards: EN 1050)
The primary function of this document is to describe a systematic procedure for risk assessment so that adequate and consistent safety measures can be selected. Risk assessment is an essential part of the iterative process for risk reduction which should continue until adequate safety is achieved.

IEC 60445 – Basic and safety principles for man-machine interface, marking and identification – Identification of equipment terminals and of terminations of certain designated conductors, including general rules for an alphanumeric system (Category: KA / Harmonized standards: EN 60445)
This international standard applies to the identification and marking of terminals of electrical equipment such as resistors, fuses, relays, contactors, transformers, rotating machines and, wherever applicable, to combinations of such equipment (e.g. assemblies). It also applies to the identification of terminations of certain designated conductors.

IEC 60204-1 -- Safety of machinery — Electrical equipment of machines — Part 1: General requirements
IEC 60204-1 is applicable to the electrical equipment or parts of the electrical equipment that commences at the point of connection of the supply to the electrical equipment of the machine and operate with nominal supply voltages not exceeding 1 000 V for alternating current (a.c.) and not exceeding 1 500 V for direct current (d.c.), and with nominal supply frequencies not exceeding 200 Hz.

IEC 60947-5-1, Low-voltage switchgear and control gear — Part 5-1: Control circuit devices and switching elements – Electromechanical control circuit devices
IEC 60947-5-1 applies to control circuit devices and switching elements intended for controlling, signalling, interlocking, etc., of switchgear and controlgear. It applies to control circuit devices having a rated voltage not exceeding 1 000 V a.c. (at a frequency not exceeding 1 000 Hz) or 600 V d.c.

IEC 61000-6-2, Electromagnetic compatibility (EMC) — Part 6-2: Generic standards — Immunity for industrial environments
This part of IEC 61000 for EMC immunity requirements applies to electrical and electronic apparatus intended for use in industrial environments, as described below. Immunity requirements in the frequency range 0 Hz to 400 GHz are covered. No tests need to be performed at frequencies where no requirements are specified. This standard applies to apparatus intended to be connected to a power network supplied from a high or medium voltage transformer dedicated to the supply of an installation feeding manufacturing or similar plant, and intended to operate in or in proximity to industrial locations, as described below. This standard applies also to apparatus which is battery operated and intended to be used in industrial locations.

IEC 61000-6-4, Electromagnetic compatibility (EMC) — Part 6: Generic standards — Section 4: Emission standard for industrial environments
This standard for emission requirements applies to electrical and electronic apparatus intended for use in the industrial locations (both indoor and outdoor, or in proximity to industrial power installations) for which no dedicated product or product-family emission standard exists. Disturbances in the frequency range 0 Hz to 400 GHz are covered.

A very important – and most relevant for SRS – standard will be the new standard “Robots and robotic devices — Safety requirements — Non-medical personal care robot” (ISO 13482) which is currently under development. This new international standard specifies requirements and guidelines for the inherent safe design, protective measures and information for use of non-medical personal care robots. It describes hazards associated with the use of these robots and provides requirements to eliminate, or adequately reduce, the risks associated with these hazards. The standard includes risk assessment and risk elimination/ reduction information.

6. Interaction Technology

6.1. Introduction

This section discusses relevant technologies for the interaction of the user with the SRS system. This includes hardware and user interface concepts. There are two general types of users for which the chosen interface solution must be appropriate: the elderly user in the home and the remote operator. Characteristic about the elderly user group are limited short-term memory, lower co-ordination capacity, lower sensory capability, slower ability to react, little computer experience, decreased fine motor skills, and a wide heterogeneity of user capabilities and experiences (e.g., impairments vary in degree and type between individuals and may deteriorate further over time) [94]. Regarding the remote operator, it remains to be determined which kind of user groups will be suitable for operating the robot. This decision will in part be based on the outcome of task 1.1 (user requirement assessment). Possibly users are younger relatives (e.g., son or daughter of the elderly person), professional caregivers, and a specialised remote operation service. Their characteristics have to be respected as well, for example sons and daughters may lead a busy working life. An interface that supports quick task execution or that can be used parallel to another activity (such as working in an office or taking a phone call) may be suitable here.

There are also several characteristics specific to SRS which the interface needs to support:

- **Assisted, semi-autonomous teleoperation with adaptive autonomy:** The system will try to offer autonomy ideally but fall back to semi-autonomous teleoperation whenever necessary (i.e., the user operates but is assisted by the system; high-level control), and possibly also to low-level teleoperation (conventional teleoperation without system assistance), depending on the uncertainty of the task.
- **Learning capability:** The system analyses the user interaction and tries to derive recurrent patterns which it tries to generalise and transform into knowledge which can be applied autonomously or assistively. Capability changes need to be indicated to the user and pose an interface challenge, as the interface needs to adapt to evolving robot capabilities over time [95].
- **Multi-user operation:** Since there is a local (elderly) user and a remote user, the interface needs to support both user types, including the possibility of simultaneous interaction. In combination with the learning capability, this aspect can create additional challenges. For example, the system may have learned something from the remote user, causing a capability change. This change then needs to be communicated to the local user, posing some challenges for situation awareness and the establishing of common ground [96].
- **Off-site teleoperation:** Since the remote operator may be at locations other than the house of the elderly person, the interface needs to work through the Internet, mediated through real-world networking technology such as HSPA. This is associated with delays, insufficient quality of service, and loss of data packets [97]. Ideally, affordable consumer devices would be used for teleoperation.

The choice for interaction technology also depends on the scenario that will be chosen for the SRS project (milestone 1b). Different tasks require different interaction approaches. For example, fetching services will require an interface that focuses on intuitive navigation through the apartment whereas cooking meals may require an interface that focuses on fine movements employing all degrees of freedom (DOF) of the robot arm. Depending on the scenario, the robot may have more or less experience and require more or less assistance by a remote operator. This level of robot autonomy also influences the choice for appropriate interaction techniques. For example, high (though not full) autonomy would suggest high-level control (e.g., touch screen point and select interaction) whereas little autonomy would require an interface that focuses on accurate translation of motion commands (e.g., a 3D mouse).

6.2. Literature research

The following summary presents eligible electronic devices along with interface approaches which may be considered for SRS user interaction. The selection emphasises off-site remote operation (by a relative, caregiver, or professional service). For the local user, SRS will largely rely on the input and output devices already employed in Care-O-bot (www.care-o-bot.de). These are touchscreen interaction through the expandable tray, speech recognition and output, robot body gestures (bowing, nodding), and basic signalling through front-faced LEDs. The current capabilities of these technologies may be expanded to better support the SRS scenario(s). It remains to be decided whether or not the elderly user will also be equipped with a device for teleoperation. Depending on the scenario, this may be appropriate. For example, the elderly user may initiate a fetch command which the robot can execute semi-autonomously. The robot would then need feedback on which object to choose. This feedback would ideally be given by the person who initiated the command, for example using a handheld touch screen and a simple point-and-select gesture.

6.2.1. Smartphones and tablet computers with (multi-)touch interfaces

Compelling about smartphones is that they are relatively low-cost consumer devices with a wide reach and people constantly carry them. This is an important aspect regarding the often required (tele-)presence of family members. The Apple iPhone seems suitable because this device is particularly intuitive to use, has a sophisticated SDK, and a large developer community.
The main advantage of tablet computers such as the Apple iPad over smartphones is the larger display. This makes a more detailed view of the environment possible and extends the usage field. Such a device may even be appropriate for operation by the elderly person in the home. Simple pointing gestures (single-touch) for object selection may assist semi-autonomous operation and have proven to be suitable for operation by elderly [98]. However, the larger size and weight also restricts usage and cannot be considered for scenarios requiring constant availability of a casual remote operator such as a relative.
[99] outline a navigation approach employing maps on a touch screen device. The user points at a location on a drawing of a room, for example the couch, which triggers robot movement to that place. This type of navigation seems well suited for teleoperation over real-world networks but may show accuracy and perspective problems if it is not complemented by other means of navigation control. [100] developed a multi-touch interface with a top-down view from a ceiling camera for robot control.
Smartphones and tablet computers may further be used as augmented displays. For example, an interesting implementation was demonstrated by [101]. Teleoperation of a vehicle was realised with the iPhone over a conventional Wi-Fi network. The interface employs a camera picture augmented with steering controls, pedals, and driver’s information. It utilises the iPhone’s built-in accelerometer for steering. The implementation shows some analogies to the required SRS teleoperation. [102] used a mixed-reality robotic interface that employs cartoon art for expressing robot states and required user interaction (e.g. for collisions). [103] enhanced the camera image of the robot’s vision with the robot’s perception of obstacles for better situation awareness. For example, objects on the ground which would hinder driving were clearly indicated to the user as computer-generated colored 3D objects. They could easily have been overseen otherwise. [104] used augmentation for path planning of a mobile robot moving in a room. The robot was supported by placing virtual nodes on a video picture of the environment. It moved from node to node. [105] provide an extensive overview of augmentation use in human-robot interaction with varying levels of autonomy.

6.2.2. Controller-free 3D gesture recognition

Human-robot interfaces are often enhanced by gesture recognition [106], [107] in order to achieve a more dialogue-oriented interface [108]. New sensing technology developed by PrimeSense and Microsoft for project Natal (www.xbox.com/en-US/live/projectnatal/) will enable gesture-based interaction without a controller in 3D space. A compact extension unit for the Xbox 360 gaming console integrates a time-of-flight camera, an RBG camera, and a 3D multi array microphone. Advanced algorithms for motion detection and multi-person body tracking will be accessible through an SDK. This technology is announced to be released in late 2010 and may be considered for SRS.

6.2.3. Accelerometer-based input devices

The Wii remote for the Nintendo Wii gaming console was the first mainstream device to rely on accelerometer control. [109] evaluated a Wii remote for robot teleoperation and compared it to joystick and button operation. They recommend to avoid the device for controlling the robot to move straight or to move with acceleration, even though it seems to be a good device for rotating the robot. Meanwhile, other, more advanced accelerometer-based devices have emerged such as Sixense TrueMotion 3D (www.sixense.com), enabling highly precise control with accuracy up to a millimetre. This device additionally relies on magnetic field tracking and tracks both hands in 3D space with 6 DOF. It is designed for use with conventional PCs, not gaming consoles. However, it needs to be determined if this class of devices may also be suitable for higher-level, semi-autonomous robot teleoperation over the Internet.

6.2.4. Speech recognition and output

Here it can be distinguished between natural language dialogue [110] and limited command sets [111] and [112]. While users often expect the former initially, it usually cannot meet their expectations and is not always appropriate [113]. Challenges are the ability to recognise spoken word at large distances, speaker independence, and pre-training [111]. Torrey [114] employed a concept called “elderspeak” where the robot speaks words slowly and loud, similar to the communication with children. This may be suitable for elderly persons with hearing impairments or cognitive decline.

6.2.5. 3D force-feedback joystick combined with a computer monitor

3D input devices with haptic feedback can offer an intuitive means of robot teleoperation. [115] carried out experiments with a teleoperated robot and a force feedback interface. They found that force feedback to the remote operator results in lower peak forces of the robot. The positive consequences of force feedback reduced task completion times, decreased peak forces and torque, and decreased cumulative forces. Similar results were obtained by [116]. [117] used force feedback to communicate the motion planning of the robot to the human remote operator.

The Novint Falcon as a low-cost consumer force-feedback 3D joystick (www.novint.com) is already in use at project partner Fraunhofer IPA. Project partner Stuttgart Media University [118] carried out user studies with the Falcon and elderly users and obtained promising results. However, controlling a 7 DOF arm (Care-O-bot) with a 3 DOF device (Falcon) can show some limitations. Other devices, some with more DOF (up to 7) exist in the professional sector, e.g. Phantom from Sensable Technologies (www.sensable.com) or Virtuose from Haption (www.haption.com). The direct, low-level control provided by such devices is however not prioritised in the SRS project and may prove problematic over real-world networks.

6.2.6. 3D mouse and computer monitor

3D mice (e.g. www.3dconnexion.com) are typically used in CAD and 3D modelling but may be considered depending on the scenario. One thing they typically lack is force feedback and the limitations are similar to the ones mentioned for the 3D joysticks.

6.2.7. Data gloves

Data gloves (e.g. www.acceleglove.com) seem more suitable for teleoperation in controlled environments (non-Internet) because the trajectory copying approach is sensitive to network stability. However they could still be considered depending on the task.

6.2.8. Combined and multi-modal approaches

Additionally, combinations of interaction techniques can be considered. For example, [119] used an approach combining joystick commands with linguistic commands for Internet robot teleoperation. Depending on the task, low-level, direct joystick control or higher-level control through spoken commands such as “turn” or “go to end” were used.

[120] suggest a framework for deictic references based around decomposition and re-synthesis of speech and gesture into a language of pointing and object labelling. A user would point at an object, for example a nut in a nut-tightening task and simultaneously issue voice commands such as “On this wheel, this nut is nut 1.” The system then links the objects pointed at to the simultaneously uttered words (e.g. “this”) to achieve spatial common ground. [105] employed deictic references in augmented reality environments.

6.3. FP6/FP7 research

6.3.1. Robot@CWE

The Robot@CWE FP6 project (www.robot-at-cwe.eu) researched robots in collaborative working environments. An HRP-2 robot was teleoperated with a PDA through the BSCW groupware (Basic Support for Cooperative Work, www.bscw.de). Their deliverable 2.1 provides an interaction model and a taxonomy of human-robot interaction. SRS project partner HP was a partner in this project. There are several publications related to teleoperation and learning which could be a useful information resource for SRS interface design, e.g. [121] [122].

6.3.2. AHUMARI

AHUMARI (www.ahumari.org), involving SRS project partner Profactor, developed augmentation based visual learning methodologies for industrial robotics. In case SRS should adopt an interface approaches involving AR, the project may be a useful resource. Another similarity is the focus on robot learning by analysing user interaction: “AHUMARI makes use of AR/VR user interaction to learn and adopt or refine strategies and for subsequent automatic re-use when new jobs and tasks are planned.” (www.ahumari.org/blog).

6.3.3. Other and non-EC-funded projects

In case SRS should adopt a scenario involving measures against cognitive decline, there are several EC-funded projects aiming to provide memory stimulation and cognitive training of the elderly population: HERMES (www.fp7-hermes.eu), Vital Mind (VM, www.vitalmind-project.eu), and Long Lasting Memories (hwww.longlastingmemories.eu). FLORENCE (no website yet) is a FP7 project under the same call as SRS which aims to develop a multi-purpose robot for ambient assistant living. The robot is going to provide care and coaching services for elderly. Like SRS, the project has a strong user focus and started in February 2010 with focus group interviews. From a robot learning perspective, SKILLS (www.skills-ip.eu) could be another relevant project. It deals with the acquisition, storing and transfer of human skill by means of multimodal interfaces. COGNIRON (www.cogniron.org) aims to adapt robot behaviour in changing situations and for various tasks so that the robot evolves and grows its capacities in close interaction with humans in an open-ended fashion. The AMIDA project (www.amiproject.org) focuses on live meetings with remote participants using augmented multi-party interaction and could be relevant in case a scenario involving videoconferencing should be adopted. DICIT (dicit.fbk.eu) addressed technical issues such as multiple speaker localization, distant-talking automatic speech recognition, mixed-initiative dialogue, and multi-modal integration. SRS could benefit from the results of this project if the interface will rely on speech recognition.

Further, some non-EC funded projects show analogies to SRS and could be of interest for interface design during the course of the SRS project: WimiCare (www.wimi-care.de), ImRoNet (www.imronet.de), and DESIRE (www.projekt-desire.de). SRS project partner Fraunhofer IPA is a consortium partner of all three of these projects.

7. Haptics

7.1. Introduction

Haptics, or haptic technology, is a tactile feedback technology that takes advantage of a user's sense of touch by applying forces, vibrations, and/or motions to the user [123]. This mechanical stimulation may be used to assist in the creation of virtual objects (objects existing only in a computer simulation), for control of such virtual objects, and for the enhancement of the remote control of machines and devices (teleoperators). It has been described as "(doing) for the sense of touch what computer graphics does for vision" [124].

7.2. Haptic loop and design

The haptic loop is implemented in the following way. First, the haptic device senses an operator input, which may be position (and its derivatives), force, muscle activity, etc. Second, the sensed input is applied to a virtual or teleoperated environment. For a virtual environment, the effect of the operator's input on virtual objects and the subsequent response to be displayed to the operator are computed based on models and a haptic rendering algorithm. In teleoperation, a manipulator that is remote in space, scale, or power attempts to track the operator's input. When the manipulator interacts with its real environment, haptic information to be relayed to the operators is recorded or estimated. Finally, actuators on the haptic device are used to physically convey touch sensations to the human operator. Based on the haptic feedback, whether through unconscious or conscious human control, or simply system dynamics, the operator input is modified. This begins another cycle of the haptic loop.

Based on the above loop, the following aspects should be considered in haptic design:

1) admittance or impedance architecture;

2) mechanisms, including measures of mechanism performance, kinematic and dynamic optimization, and grounded or ungrounded devices;

3) sensing, i.e., position sensing (by encoders) and/or force sensing (by force sensors);

4) actuation and transmission, with the requirements of low inertia, low friction, low torque ripple, back-driveability, low backlash, and higher power-to-weight ratio;

5) haptic rendering, including sensing, kinematics, collision detection, determining surface point, force calculation, kinematics, and actuation.

7.3. Applications

A variety of novel and creative applications are being developed regularly in numerous fields, including assistive technology, automotive, design, education, entertainment, human-computer interaction, manufacturing/assembly, medical simulation, micro/nanotechnology, molecular biology, prosthetics, rehabilitation, scientific visualization, space, and surgical robotics. The most common haptic device encountered by the general population is a vibration display device that provides haptic feedback while an operator plays a video game. For example, when the operator drives off the virtual road or bumps into a virtual wall, the hand controller shakes to imply driving over a rough surface or displays an impulse to represent the shock of hitting a hard surface.

The bulk of commercially available haptic devices are designed by two companies, SensAble Technologies and Immersion Corporation. SensAble has developed the Phantom line of stylus-type haptic devices. The Phantom Premium has been the most widely used haptic device in research to date. The Phantom Omni, which is an order of magnitude less expensive than the Phantom Premium, has also gained popularity among haptics and robotics researchers. Immersion has aimed at the mass market and consumer segments with a wide variety of haptics-based products, many of them involving a single degree of freedom used in, for example, video games, mobile phones and medical simulation.

Software for haptic rendering has also become widely available, through both commercial sources and research groups. Most companies that sell haptic devices also provide a standard development kit (SDK) with haptic rendering capability. In addition, not-for-profit open-source projects such as Chai3D aim to make rendering algorithms from different groups publicly available.

7.4. Academics

There exist many books on the topic of haptic technology, most of them compendiums from workshops or conferences on the subject. The representatives are [125], [126]. Also, there exist two journals that are specific to the field of haptics: *Haptics-e*, and *IEEE Transactions on Haptics*. Several conferences are specifically devoted to haptics, such as Eurohaptics and the Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems.

7.5. FP6/FP7 research

Name	Status	Research Focuses	Main Outcomes	Web address
Sound and tangible interfaces for novel product shaping - SATIN	IST-034525 Completed 2009-SEP-30	- multimodal interface based on fusion of force feedback, sound and vision for representing global and local properties of shape and material, which can be perceived statically or dynamically during exploration and modification of digital shapes performed by users through free-hand, unconstrained, robust and ergonomic interaction; - novel haptic and tangible interfaces allowing users to modify digital shapes through free-hand interaction which aims at exploiting users' dexterity and skills in physically interacting with materials.	- **Knowledge improvement** in the fields of haptic interfaces, sensors, shape modelling, metaphoric sounds, multimodal interfaces, human haptic perception of curves, in respect to state-of-the-art - Results of experiments on **human haptic perception of shapes** - **SATIN system** consisting on multimodal use of haptic strip, stereo viewing and metaphoric sounds for shape exploration and local/global modification - Prototype of **desktop haptic strip**	http://www.satin-project.eu

8. Computer Vision based Human Motion Analysis

8.1. Introduction

Computer vision is the science and technology of machines that see. “As a scientific discipline, computer vision is concerned with the theory behind artificial systems that extract information from images. The image data can take many forms, such as video sequences, views from multiple cameras, or multi-dimensional data from a medical scanner” [127]. Computer vision based motion analysis is one of the key focuses of the SRS project. It opens up the possibility of intuitively communication with robot in a highly efficient manner. In this report, computer vision technologies involved in the motion analysis will be briefly explored with specific emphasis on the high level motion understanding. Key research projects related to the topic will also be evaluated to identify potentials of relevant work already launched.

8.2. Classification of Vision Based Motion Analysis Technologies

The research topic involved in the vision based motion analysis can be classified into three categories:

- **Motion Detection and extraction:** in this step image points or regions which are relevant to further processing will be identified. Features at various levels of complexity are extracted from the image data. The final output of the step should be some set of data, i.e. a set of points or an image region which is assumed to contain a specific object.
- **Motion Tracking:** Tracking is a specific important issue in human motion analysis. It prepares data for gesture recognition. In contrast to detection and extraction, this category belongs to a higher-level computer vision problem. However, the tracking algorithms within human motion analysis usually have considerable intersection with motion segmentation during processing.
- **Motion Understanding:** This step is raised after successfully tracking the moving motions from an image sequence. Motion understanding involves action recognition and description. It will be one of the most important research areas for the future work of the SRS project.

A classification of the research topic of computer vision involved in the process is listed as follows:

Detection & Extraction	Motion Segmentation	Background subtraction	Sensitive to dynamic scene changes
		Statistical methods	Employing colours or edges as statistical quantities; popular due to its robustness to noises, shadow, changes of lighting conditions, etc.
		Temporal differencing	Using pixel-wise difference between two or three consecutive frames; adaptive to dynamic environments but possibly generating holes inside moving entities
		Optical flow	Detecting change regions using their flow vectors which describe coherent motion of points or features between image frames; competent to detect independently moving objects, but computationally complex and sensitive to noise
	Object Extraction	Shape-based extraction	Localized interest points such as corners, blobs or points. More precise by temporal consistency constraints
		Motion-based extraction	Periodic property of non-rigid articulated human motion; achieved by time-frequency analysis, optical flow, also combination with the above shape-based method
Tracking	Model-based tracking	Stick figure	Approximate a human body as a combination of line segments linked by joints,
		2-D contour	Human body segments are analogous to 2-D ribbons or blobs, but restricted to the camera’s angle

		Volumetric models	More complex 3-D volumetric models, the better results may be expected but may require more parameters and lead to more expensive computation during the matching process.
	Region-based tracking		Identifying a connected region associated with each moving object and tracking it over time by a cross-correlation measure. Reasonably working result, but not good with long shadows and in congested situations.
	Active-contour-based tracking		Representing the bounding contour and dynamically updating it over time; Reducing computational complexity but requiring good initial condition.
	Feature-based tracking		Not tracking the objects as a whole but their sub-features such as points and lines
Motion understanding	General techniques	Dynamic time warping	Template-based dynamic programming matching technique
		Hidden Markov models or Neural Network	Networks with memory, superior to DTW in processing un-segmented successive data
	Action recognition	Template matching	First converting an image sequence into a static shape pattern, and then comparing it to pre-stored action prototypes during recognition; Low computational complexity and simple implementation, but sensitive to noise, the time interval of the movements, and the viewpoint dependent.
		State-space approaches	Defining each static gesture as a key state and regarding any motion sequence as a tour through various key states; It may involve complex iterative computation.
	Semantic description		Detailed in the following sessions.

Table 1 Classification of Vision Based Motion Analysis Technologies

The interest in this area has led to a large number of research papers. The state of art is best summarised by the following table of key review papers in the area.

Year	Author	Focus
1995	Cedras and Shah [128]	Motion detection & extraction
1996	Ju [129]	Motion understanding
1998	Aggarwal et al. [130]	Articulated and elastic non-rigid motion
1999	Aggarwal and Cai [131]	Motion extraction
1999	Gavrila [132]	Motion understanding
2001	Moeslund and Granum [133]	Motion tracking and understanding
2003	Wang et al. [134]	Motion detection, tracking, and understanding
2003	Buxton [135]	Motion understanding
2004	Hu et al. [136]	Surveillance
2004	Aggarwal and Park [137]	Motion understanding
2006	T.B. Moeslund et al. [138]	Motion tracking and understanding

Table 2 key review papers for vision based motion analysis technologies

When an observed frame or sequence is available, human motion understanding becomes a representation problem or more specifically a classification problem.

8.3. Semantic human motion representation

Semantic description is the last step and the highest level in human motion analysis [134], whose implementation is based on preprocessed human detection, tracking and action recognition. Its purpose is to reasonably choose a group of motion words or short expressions or simple sentences to obtain the mapping relationships between human behaviours in image sequences and the natural language (NL). Due to its advantage in human-computer interaction, this methodology has recently received considerable attention [139-144].

In the following chapters, we will firstly review various taxonomies to summary the main methodologies for semantic description; and then several hierarchical approaches with high practicability for SRS project will be highlighted; finally, some existing problems in this domain will be explored.

8.3.1. Taxonomies

8.3.1.1. STP/HLSP/LP

In [139], the semantic representations are further divided as Spatiotemporal Predicates (STP) to deal with metric-temporal relations, High-Level Semantic Predicates (HLSP) to express semantic relations among entities resulting from applying situational models over STP, and Linguistic Predicates (LP) to represent linguistic-oriented knowledge for NL generation and understanding. Among them, STP allows inference of higher-level predicates upon asserted facts, but is limited to metric-temporal reasoning; HLSP is the linguistic-oriented highest level of interpretation to process behavioural models (contextual and intentional) but with the limitation of domain-dependent and target-oriented; LP implies linguistic models (syntax, morphology, alignment, etc.) and facilitates to convert between logic and NL but language-dependent, that is, each LP requires distinct thematic arguments depending on the language and situation. Thereby, in real applications especially in hierarchical representation [141-143], they work together and perform advantages in different layers, which will be stated in detail later.

8.3.1.2. Bottom-up/top-down

Bottom-up method means learning behaviour patterns from observation. It might introduce inaccuracy from the vision system and the ambiguity of possible interpretations due to a semantic gap [144]. To cope with this drawback, some recent contributions have proposed the use of ontologies to restrict the domain like “surveillance on urban intersections” or “elderly health monitoring” [145]. Ontologies are argued to be “particularly useful in the application field of NL understanding” [140] which describe a set of concepts and their relations and relate quantitative data with semantic concepts.

The integration of these two schemes is proposed in some works [140] [144]. The structure of the proposed system in [140] is based on a modular architecture which allows both top-down and bottom-up flows of information. Concretely, it evaluates the information at a certain cognitive stage, and generates new descriptions oriented to modules at higher levels. This is the bottom-up data flow. Then, it sends high-level productions and inferences back in a reactive manner to support low-level processing and guide mechanisms for evaluation. This is the top-down data flow. It is worth pointing out that traditional top-down approaches do neither evolve over time nor learn from new observations, thereby being affected by any modification in the scenario or the incorporation of new behaviours; but here the improvement always updates the history of asserted facts.

8.3.1.3. Statistical models/formalised reasoning

Besides the above taxonomies, Hu et al. [136] categorise behaviour description methods into statistical models and formalised reasoning. A representative statistical model is the Bayesian network model [146]. It interprets certain events and behaviours by analysis of time sequences and statistical modelling. However, it rests on lower-level recognition based on motion concepts, and does not yet involve high-level concepts, such as events and scenarios, and the relationships between these concepts which need high-level reasoning based on a large amount of prior knowledge. On the contrary, formalised reasoning [141, 142] uses symbol systems to represent behaviour patterns and reasoning methods such as predication logic to recognise and classify events. The most significant disadvantage of the formalised reasoning methods is that they cannot handle uncertainty of events [147].

8.3.1.4. Verb-based/situation-based

Verb-based semantic description denotes that the preprocessing work is for selecting a most suitable verb to represent an event and selecting other syntactic elements by surrounding and to supplement the verb. Undoubtedly, verb (motion) is the centre in the common syntactical agent-motion-target triplets. The representatives are shown in [141, 142]. In contrast, [139, 140] apply situations instead of verbs as basic elements for an ontological categorisation of occurrences, where the chosen list of entities include agents as those which can spontaneously act to change a situation; objects as static elements of the scene; locations; and also a set of abstract descriptors which permit to add fuzzy modifiers to the conditions related to the entities. Each situation is related with a set of required entities as above. The main advantage of this approach, as argued by the authors, is an independency of the particularities of verbs from a concrete NL, thus facilitating addition of multiple languages.

The above summaries are tabulated in Table 1. Note that, they are just classified in different ways; actually, however, one application can simultaneously belong to different categories, like top-down and formalised reasoning, as shown in the following detailed review of some hierarchical methods.

Taxonomy	Categorising focus	Category	Description	Advantage	Disadvantage	Works
STP / HLSP / LP	Relations between "events"	Spatiotemporal Predicates (STP)	Metric-temporal (basic relations)	Allows inference of higher-level predicates upon asserted facts	Limited to metric-temporal reasoning	[148]
		High-Level Semantic Predicates (HLSP)	Thematic roles (inferential role semantics)	Linguistic-oriented, highest level of interpretation	Domain-dependent and target-oriented	[139]
		Linguistic Predicates (LP)	Linguistic-oriented knowledge (NL semantics)	Facilitates to convert between logic and NL	Language-dependent	[149]
Bottom-up / top-down	Modeling process and learning capability	Bottom-up	Learning behaviour patterns from observation	Learning from new observations	Introducing inaccuracy from the vision system and the ambiguity of possible interpretations	[140]
		Top-down (ontology)	Predefined behaviour pattern by experts in a restricted domain	“Particularly useful in the application field of NL understanding”	Neither evolving over time nor learning from new observations	[141-143, 145]
Statistical models / formalised reasoning	Model certainty	Statistical models	Interpreting certain events and behaviours by analysis of time sequences and statistical modelling	Increasing popularity, able to handle uncertainty of events	Resting on lower-level recognition	[146]

		Formalised reasoning	Using symbol systems to represent behaviour patterns and reasoning methods	High-level reasoning	Cannot handle uncertainty of events	[141, 142]
Verb-based / situation-based	Basic elements in semantic representation	Verb-based	selecting a most suitable verb to represent an event and selecting other syntactic elements by surrounding and to supplement the verb	Mature and popular in real applications	Language-dependent	[141, 142]
		Situation-based	Each situation is related with a set of required entities thus structuring the domain knowledge more comprehensively.	Facilitating addition of multiple languages	Not essential in normal single-language application environment	[139, 140]

Table 3 Summary of different taxonomies of semantic human motion representation

8.3.2. Methodologies

8.3.2.1. Grammars

Language is a basic mechanism used by humans to define and describe video events. It is, therefore, intuitive that formal notions of language, as defined by grammar models [150], would be natural to model the inherently semantic properties of video events. Formally, a grammar model consists of three components: a set of terminals, a set of nonterminals, and a set of production rules. Besides the basic deterministic semantic grammar models, extensions include stochastic grammars allowing probabilities to be associated with each production rule, and attribute grammars which formally associate conditions with each production rule.

Learning of semantic event models including grammar models is a challenging problem. Although several works have explored the problem of automatically learning a grammar model for video event representation, the event description and recognition in semantic terms afforded by grammar approaches can, generally, only be achieved through manual specification of the model using expert domain knowledge.

8.3.2.2. Petri Nets (PN)

The PN formalism allows such a graphical representation of the event model and can be used to naturally model semantic relations that often occur in video events including nonsequential (and sequential) temporal relations, spatial and logical composition, hierarchy, concurrency, and partial ordering. Some extensions include timed transitions to allow representation of duration constraints, stochastic timed transitions for dealing with uncertainty within duration constraints, and associating probabilities with tokens to cope with uncertain observations.

An additional advantage of the PN event model is its ability to deal with “incomplete” events. Unlike other models, PNs are able to give a semantically meaningful snapshot of the video input at any time. This ability can be used to give a prediction on the next state or provide a likelihood of reaching a particular event of interest [151]. One drawback of the PN model is that the semantic nature makes learning these models from training data infeasible/ill defined. Another disadvantage of PN event models is their deterministic nature.

8.3.2.3. Constraint satisfaction

Another approach is to pose the problem of recognition as one of constraint satisfaction. Early work in constraint recognition introduced the notion of chronicles, which are undirected constraint graphs describing the temporal constraints of atomic subevents. The event recognition task in these approaches is reduced to mapping the set of constraint to a temporal constraint network and determining whether the abstracted video sequence satisfies these constraints. In addition to temporal constraints, more recent work incorporates semantic knowledge about temporal constraints pertaining to the properties of objects participating in the scene. Description logics [152] offer a very rich framework for representing video events including compositional hierarchy specification as well as semantic relationships.

The advantage of this approach is that the constraints can be formulated as an ontology for a particular event domain and reused in different applications.

8.3.2.4. Logic approaches

Logic-based event models have been introduced for video event understanding recently. In this type of event model, knowledge about an event domain is specified as a set of logic predicates. A particular event is recognized using logical inference techniques such as resolution. These techniques are not tractable in general, but are useful as long as the number of predicates, inference rules, and groundings (usually corresponding to the number of objects in the video sequence) are kept low.

Initial work applies the first-order logic framework of Prolog to recognition in the event domain of parking lot surveillance [153], and some extensions include multivalued logics and Markov logics.

8.3.3. Hierarchical human behaviour description

Natural language is too complex to comprehend for an intelligent system or a robot, which can understand semantic terms only if the NL is decomposed into basic (atomic) elements that can be represented by relations of digital data called “feature value”. For instance, in the sentence “Event 1 happens earlier than Event 2”, “earlier” means the starting time of Event 1 is less than that of Event 2 in number, exhibited as a mathematical relationship. On the other hand, conversely, the hierarchical description systems integrate the basic or low-level computer-translated language elements into high-level events and scenarios according to the spatial/temporal/logical constraints to make language concise and precise for human comprehension. Therefore, the hierarchical representation can be seen as a practical human-computer interface in human behaviour description. Its related algorithms and applications are reported in [39, 141-143, 154, 155]. The first three are reviewed here.

In the initial algorithm of Kojima et al.[141], they extract a head region of a human, as a portion of the whole body, from each image frame, and its three-dimensional pose and position are estimated using a model-based approach. Next, the trajectory of these parameters is divided into the segments of monotonous movement. The conceptual features for each segment, such as degrees of change of pose and position and that of relative of distance from other objects in the surroundings, are evaluated. Meanwhile, a most suitable verb is selected and other syntactic elements are supplied. Finally, the natural language text for interpreting human behaviours is generated using machine translation technology. Despite no words like “hierarchy” or “multi-level” appearing in the text, such feature embodies in the extraction of conceptual features of a behaviour illustrated by the analysis of concept of the verb “to approach” by its four conceptual features.

In its improved version [142], the hierarchical feature becomes explicit. First, a hierarchical state-transition diagram (STD) is constructed respectively for body actions, head actions and hand actions. STD marks durative action verb and instantaneous action verb correspondent to a state and state transition respectively. Next, case frames are generated using the STD. Then, three case frames from each body part are composed into unity case frame of whole body action, not merging directly but obeying three specific rules.

Nevatia et al. [143] define an event ontology that allows natural representation of complex spatio-temporal events common in the physical world by a composition of simpler events. The events are abstracted into three hierarchies. Primitive events are defined directly from the mobile object properties. Single-thread composite events are a number of primitive events with temporal sequencing. Multi-thread composite events are a number of single-thread events with temporal/spatial/logical relationships. The hierarchical event representation naturally leads to a language description of the events. An Event Recognition Language (ERL) is then defined allowing the users to define the events of interest conveniently without interacting with the low level processing in the program.

8.3.4. Existing imperfections

NL human behaviour description is a challenging research subject. At present, it is still restricted to simple and specific action patterns. Therefore, research on semantic description of human behaviours in complex unconstrained scenes still remains open.

Concerning the relationships between behaviours and semantic concepts, the key problems include the modeling of semantic concepts of motions, and the automatic learning of semantic concepts of behaviors. As for semantic recognition and natural language description of object behaviours, organising recognised concepts and further representing object behaviours in brief and clear NL is one of the ultimate goals. In addition, the synchronous description, i.e., giving the description before behaviour finishes (during the behaviour is progressing), is also a challenge. New incremental description method should be designed which is able to predict object behaviors.

As stated in [144], the use of ontologies seems to be an initial bridge to cope with the semantic gap between the qualitative concepts and the quantitative observations; however, it is still a challenging question for future work how to join top-down and bottom-up approaches in a way that all the advantages are kept and the drawbacks are minimised.

8.4. FP6/FP7 research

Name	Status	Research Focuses	Main Outcomes for SRS	Web address
Cognitive Vision Systems (**CogVis**)	**IST-29375** Completed 2003-DEC-31	The CogVis consortium has been brought together to study the design of Cognitive Vision Systems. In this context a "cognitive vision system" is defined as a *system* that uses visual information to achieve: • Recognition and Categorisation of Objects, Structures and Events • Learning and Adaptation • Memory and Representation of Knowledge • Control and Attention	**Recognition & categorisation** CogVis: The Object and Event Database; Appearance and Contour Based Methods for Object Categorization Algorithms for robust subspace recognition; Algorithm for categorisation using subspace approach; **Interpretation** Conceptual Framework for High-level Vision; Generalising visual experiences associated with observed actions sequence; Recognising intentions in scenes with actions; A qualitative spatial inference service for a visual agent;	https://cogvis.nada.kth.se
Marker-less real-time tracking for augmented reality image synthesis (**MATRIS**)	**IST-002013** Completed 2007-JAN-31	The project is developing a hybrid camera tracking system, using vision and inertial sensors. The vision part tracks naturally occurring features in the scene, whose 3D positions are known. Vision-based tracking algorithms can give good absolute accuracy, but are hard to run at full video frame rate, particularly when tracking natural features. They can also fail when insufficient features are visible. Conversely, an Inertial Measurement Unit (IMU) can provide very rapid and robust measurements, but these tend to drift over time. By fusing information from both sources, accurate and stable measurements of camera pose can be obtained at a high rate.	**Real-time camera pose from feature tracking and sensor fusion** For scenes modelled using planar patches, these patches are tracked 'live' as the camera moves. A sensor fusion process, based on an Extended Kalman Filter, uses the image coordinates of the tracked features together with the output of the accelerometers, gyroscopes and magnetometers in the inertial measurement unit, to compute position and orientation.	http://www.ist-matris.org
Cognitive systems using perception-action learning (**COSPAL**)	**IST-004176** Completed	A new system architecture and new learning strategies for artificial cognitive systems. The novelty of the approach lies in the interaction of continuous and symbolic perception and action, which results in robust and stable motor and sensorial capabilities of the system	Learning Responses to Visual Stimuli; Real-Time Visual Recognition of Objects and Scenes Using P-Channel Matching; Category-Level Object Recognition;	http://www.cospal.org

Action recognition and tracking based on time-of-flight sensors (**ARTTS**)	**IST-034107** Completed 2009-SEP-30	Development of algorithms and prototype systems to solve open computer-vision problems and enable new applications that involve multimodal interfaces and the observation of people and their actions. Unlike a conventional video camera, a TOF camera delivers not only an intensity image but also a range map that contains a distance measurement at each pixel, obtained by measuring the time required by light to reach the object and return to the camera (time-of-flight principle).	**3D-TOF Database** The ARTTS 3D-TOF database contains datasets for typical application scenarios. Specifically, the database contains the following datasets: * Face detection dataset, containing around 1300 images with labelled faces and 3600 images not containing faces * Gesture dataset, containing 9 different simple gestures performed by different users * Head orientation dataset, containing heads of 13 different users in different orientations to the camera * Head distance dataset, containing heads of 10 different users in different orientations both at 60 and at 90 cm distance from the camera * A dataset of faces at different ranges from the camera for testing the scale-invariance of features **ARTTS Toolbox** The ARTTS toolbox contains algorithms for 3D-TOF image data. Specifically, the toolbox is divided into three parts each of which provides the listed functionality: * Signal processing toolbox for: Image saving and loading; Image acquisition; Visualization; Foreground object segmentation; Signal improvement; Computation of image features * Object tracking toolbox for: Labelling of body pose, faces, and facial features; Facial feature tracking; SOM-based pose tracking; Multiple person tracking; Gait analysis; * Action recognition toolbox for: Labelling of actions in TOF image sequences; Range flow computation; Intrusion detection based on histogram measures; Action recognition from trajectories of key points; Action recognition using motion history images	http://www.artts.eu
Multi-sensory autonomous cognitive systems interacting with dynamic environments for perceiving and learning affordances (**MACS**)	**IST-004381** Completed 2007-NOV-30	The MACS project aims at developing affordance-based control as a new architecture paradigm for embedded cognitive systems. and allows purposive behaviour. Moreover, the system is able to extend itself through exploration of the environment and has therefore the potential to solve problems of a large range of complexity levels.	A computational model for visual learning of affordance-like cues; Surface detection in 3D range data for mobile manipulation;	http://www.macs-eu.org
Perception On Purpose (**POP**)	**IST-027268** Completed 2008-DEC-31	Development of a fundamentally new approach, perception on purpose, which is based on 5 principles. First, visual and auditory information should be integrated in both space and time. Second, active exploration of the environment is required to improve the audiovisual signal-to-noise ratio. Third, the enormous potential sensory requirements of the entire input array should be rendered manageable by multi-modal models of attentional processes. Fourth, bottom-up perception should be stabilized by top-down cognitive function and lead to purposeful action. Finally, all parts of the system should be underpinned by rigorous mathematical theory, from physical models of low-level binocular and binaural sensory processing to trainable probabilistic models of audiovisual scenes.	**3D surface estimation from dense stereoscopic matching** An algorithm that provides disparities in accordance with the surface properties of the scene **Audiovisual detection and localization of multiple speakers** Detecting and localising objects in a scene that are both seen and heard. **POPEYE: An audiovisual robotic platform** Platform for the development of audio-visual algorithms that study and exploit the most important types of motion involved in human vision and hearing https://gforge.inria.fr/ **JAMF: An attention model framework** A flexible modelling environment which allows non-experts to develop models of sensory processing and attention yet do not compromise on efficiency of the code. http://jamf.eu/jamf/	http://perception.inrialpes.fr/POP
Vision and Chemi-resistor equipped Web-connected	**IST-045541** Completed	The robots will be installed with onboard TV/IR cameras, LADAR and other sensors to enhance scene reconstruction, as well as a	**Stereo Vision** Stereo Vision for Robotic Applications in the Presence of Non-ideal Lighting Conditions;	https://view-finder-project.eu

finding Robots (**VIEW-FINDER**)	2009-NOV-30	wide array of chemical sensors. The data will be sent to the base station for processing and presented to the command of the operation combined with information originating from a web of sources. The information can also be forwarded to the relevant forces dealing with the crisis (e.g. fire fighters, rescue workers and police).	Real-Time Obstacle Avoidance Algorithms Using Stereo Vision; Review of stereo matching algorithms for 3D vision; **GPGPU** NVIDIA CUDA Application in the Cognitive Supervision and Control of the Multi Robot System;	
Augmented multi-party interaction with distance access (**AMIDA**)	**IST-033812** Completed 2009-DEC-31	Live meetings with remote participants, using low-cost commodity sensors (such as web cams and cheaper microphones), and targeting the development of advanced videoconferencing systems featuring new functionalities such as (1) filtering, searching and browsing, (2) remote monitoring, (3) interactive accelerated playback, (4) meeting support, and (5) shared context and presence.	**Visual Identification** Fast Illumination Invariant Face Detection using Haar Local Binary Pattern Features Face Recognition using Bayesian Networks to combine intensity and colour Information **Gestures and Actions** Example-based pose recovery with a CSP classifier to recognize human actions using recovered poses.	http://www.amiproject.org
Developing versatile and robust perception using sonar systems that integrate active sensing, morphology and behaviour (**CHIROPING**)	**ICT- 215370** To be completed 2011-JAN-31	The project will implement and evaluate two demonstration sensors built as biomimetic models of an insect gleaning and a water-trawling bat species respectively. It will use a classic biomimetic methodology, involving close collaboration between bat ethologists and roboticists. It will proceed by identifying and measuring the relevant acoustic and morphological parameters of a few carefully selected bat species, reconstructing from that the bat's acoustic experience as it flies through natural hunting tasks.	Acoustic modelling techniques for thin deforming objects; New visual methods for acquisition of small complex convoluted shapes;	http://www.companionable.net
Heterogeneous 3-D perception across visual fragments (**EYESHOTS**)	**ICT-217077** To be completed 2011-FEB-28	The goal of EYESHOTS is to investigate the interplay existing between vision and motion control, and to study how to exploit this interaction to achieve knowledge of the surrounding environment that allows a robot to act properly. Robot perception can be flexibly integrated with its own actions and the understanding of planned actions of humans in a shared workspace. The research relies upon the assumption that a complete and operative cognition of visual space can be achieved only through active exploration of it:	Real-time phase-based optical flow on the GPU; Virtual reality simulator for active stereo vision systems;	http://www.eyeshots.it

		the natural effectors of this cognition are the eyes and the arms.		
Planetary robotics vision ground processing (**RROVISG**)	**SPA-218814** To be completed 2011-DEC-31	PRoVisG will build a unified European framework for Robotic Vision Ground Processing. State-of-art computer vision technology will be collected inside and outside Europe to better exploit the image data gathered during future robotic space missions to the Moon and the Planets. This will lead to a significant enhancement of the scientific, technologic and educational outcome of such missions.	3D Vision for Panoramic Camera of ESA ExoMars mission 2016; Aerobots Navigation & Mapping;	http://www.provisg.eu

Table 4 Summary of key research projects on vision based human motion analysis

9. Environment perception

9.1. Introduction

Interaction with the environment is a key feature in mobile robotics. In order to be able to act in an unknown or only partially known dynamic environment, perception of the surroundings of the robot is of huge importance. Based on the commonly used sensor systems, adequate methods for environment perception have to be found and applied.

Most of the sensors like laser range finders or time-of-flight cameras acquire data that can be used to construct a 3D point cloud of the environment. In combination with color cameras, colored point clouds can be obtained. Because both the sensors and the robot localization are bound to errors (e.g. noise) algorithms that are able to build a consistent point cloud from different sensor readings are a major topic in environment modelling. Based on that point cloud map, features that describe the environment can be extracted. Using this geometrical information, it is possible to model semantic maps of the surroundings of a robot.

The following research focuses on point cloud registration, geometric feature extraction and semantic mapping.

9.2. Point cloud registration

As base for a semantic environment map, a consistent point cloud has to be built out of subsequent sensor readings. There are several approaches to deal with that problem. The iterative closest point algorithm (ICP) is used by [156] and [157] in order to register point clouds from laser range finder scans. Whereas [157] solely relies on raw data to align the clouds, [156] extracts geometric features and [158] uses them to generate an initial guess of the point cloud transformation.

There are also approaches to use visual features and ICP for point cloud registration. In [159] information from a laser range finder and a color camera is combined in order to generate 3D features. Afterwards, an ICP algorithm based on the Mahalanobis distance is applied. Another variant is scan registration using Normal Distributions Transform [160] which models point distributions as Gaussian mixture-models in color-space.

Some approaches like [161] rely on calibration of time-of-flight sensors in order to compensate the measurement errors. Moreover RANSAC is used to match surfaces between two sensor readings.

A method developed at IPA is based on the fastSLAM algorithm [162]. 2D feature points are extracted from colour images and assigned 3D coordinates. The fastSLAM filter than builds a consistent feature map and corrects the erroneous robot position. Based on that information a coloured point cloud of the whole scene can be constructed.

9.3. Feature extraction

If a consistent point cloud of the environment exists, features can be extracted in order to recognize objects of interest.

First of all, visual features extracted from color camera images can be used to identify regions of interest or assign the environment model to a knowledge database. There are various approaches to visual features. For example, SURF [163] provides a feature point detector based on the Hessian matrix and a descriptor that it rotation- and scale-invariant. The work in [164] proposes a different detector and descriptor that is more efficient with respect to calculation costs than SURF.

2D features cannot be applied to the tasks of extracting geometric shapes from point clouds. For that, 3D features are a promising approach. Rusu et. al. proposed persistent point feature histograms in [158]. They extract features from 3D point clouds according to surface curvature and normal vectors. Those features are robust w. r. t. scaling and rotation and were recently improved regarding calculation speed, as described in [165].

9.4. Semantic mapping

Semantic maps that identify and label objects in a point cloud are crucial for environment perception and scene interpretation. Various research activities have been set up in this field. For example, [166] introduced an algorithm to identify types and places from laser range data. A graph-based approach to semantic mapping is presented in [167]. Recently, Rusu et. al. introduced a method to identify furniture and objects in a kitchen environment [168]. However, none of those approaches provides a general method for the construction of semantic maps considering different robots with different sensors and various types of possible environments.

In the field of scene interpretation a lot of theoretical work has taken place. A cognitive architecture for self-conciousness of a robot has been proposed in [169]. It yields promising results but suffers from computational requirements that prohibit the application on a robot. Gravot et. al. focus on motion planning based on scene information [170] and present some simulation results but do not apply their method on a real robot.

10. Mobile Manipulation

10.1.Mobile Manipulation Path Planning

A mobile manipulator consisting of a mobile base, movable in a plane, and a manipulator, has the advantage of a nearly infinite workspace. When trying to find collision free path's the big configuration space, due to high degrees of freedom, have implications on the used path planners. Planners currently used in mobile manipulation can be divided in static planners and dynamic planners, depending on the usage of static or dynamic environment models in the planning.

The most common static planning algorithms are Probabilistic Roadmaps [171], Rapidly-exploring random trees (RRT) [172] and Probabilistic Cell Decomposition (PCD) [173]. The Probabilistic Roadmap approach creates random sample configurations in the configuration space, checks them for collision freedom and connects them, if the distance between the configurations is below a pre set threshold. The RRT algorithm extends the configuration tree into the unexplored configuration space. Especially when using two trees from start and goal of the path search this algorithm is very effective when having high degrees of freedom. The PCD algorithm is a cell based random sampling of the configuration space. If a collision could occur in a cell, the state of the cell will be set to occupied. A path can then be found through the non occupied cells.

If the environment changes due to dynamic obstacles re-planning has to be done. Since planning is a very time consuming task, re-planning with the above described algorithm will not be practical. In [174] a framework is described to use potential fields to adapt already planned paths. An extension to use potential fields with probabilistic roadmaps is further described in [175]. The Adaptive Dynamic PRM path planner described in [176] uses the anytime D* algorithm to quickly find a initial roadmap and afterwards dynamically improves the roadmap. An algorithm for improving probabilistic roadmaps is the multi-query PRM [177] planner. It detects invalid edges due to collision and removes them from the roadmap in an optimized way.

10.2.Manipulator Motion Analysis

To analyze motion of mobile manipulators different approaches can be used. Mobile manipulator control algorithms use different metrics to resolve the redundancy introduced in such systems. These metrics can also be used to describe motions. In [178] different manipulability measures are used to move the redundant degrees of freedom to a optimized position. This manipulability describes the distance of the joints of the robot from a singular configuration. Another metric used in mobile manipulator control are kinematic constraints like distance from joint angles, distance from obstacles and non-holonomic constraints. A different approach to measure the quality of the configuration of a robot is the capability map described in [179]. The workspace of the manipulator will be mapped a priori and the position of the manipulator within this map can be tracked during movement. Another way of tracking the movement of a mobile manipulator is the usage of motion tracking technologies developed in the last years. These techniques mainly focus on tracking humanoid motion but may be adapted to be used with robots. A good overview on camera based human motion analysis is given in [133, 138]. Tracking and pose estimation Recent work [180] tries to use Gaussian processes to interpret motions of robots or environments.

11. Sensor Fusion

11.1.Introduction

Sensor fusion is the combination of sensor data from different sources with the aim to create information that exceeds the quality of each individual source. In terms of quality one usually relates to accuracy, completeness or confidence. The data sources considered in this project are two color cameras used for stereoscopic vision and one time-of-flight sensor that directly delivers 3-D range data. The project aims at combining both modalities to create highly accurate 3-D point clouds with associated color information. In the following the, the characteristics of the two sensor modalities are described, beginning with the 3-D time-of-flight sensor.

Time-of-flight cameras emit modulated near infra-red light to illuminate a given scene. The reflection of the modulated light is collected in a CMOS matrix. By comparing the returning signal to the camera's source modulation, the phase shift is measured which is a linear function of distance to the reflecting surface. Using the described procedure, the time-of-flight sensor is able to operate in real time at about 30 Hz. It creates dense point clouds, however with a limit spatial resolution. As the measurement principle assumes perfectly sinusoidal signals, which is not achievable in reality, the measured distance is subject to noise. It comprises about 1% of the measured distance. Also the measurement principle is biased as a function of object albedo, resulting in poor performance on textured scenes. A prominent example is the distance measurement of a checkerboard, where the balck squares seem significantly closer to the camera than the white squares. Additionally, the quality of the measured intensity image is low.

Stereo vision estimates depth through triangulation on point correspondences across image pairs and the knowledge of the cameras intrinsic and extrinsic parameters. On most textured scenes, stereo is able to provides high resolution point clouds. However, in the absence of features, the system fails to measure depth. Due to the different viewing angles of the two cameras, the system is also prone to occlusions. Additionally, low frequency distortions often disturb the feature association, leading to false depth measurements. Current state-of-the-art stereo matching algorithms achieve accurate dense depth maps only when using global optimization algorithms, needing at least several seconds computation time. Only local correlation based methods are fast enough for real time applications, at the cost of less accuracy and sparse depth maps.

To combine the advantages of several sensor modalities, the following levels of sensor fusion are distinguished ([181], [182])

- Level 0 - Data Alignment
- Level 1 - Entity Assessment
- Level 2 - Situation Assessment
- Level 3 - Impact Assessment
- Level 4 - Process Refinement (i.e. sensor management)
- Level 5 - User Refinement

Level 0 processes do not need to make any assumptions about the cause of the sensor data. They are concerned with the structure of the data set meaning that feature extraction is performed based on the signal characteristics i.e. the extraction of point clouds or line features
Level 1 processes are directly related to entities in that they estimate their identity, location, dynamics or other attributes.
Level 2 processes are concerned with the relationship among entities i.e. a human is associated with two feet, two arms an upper body and a head. Using inference, the presence of one entity is used to estimate another.
Level 3 processes deal with the prediction of a cost function given the current state and objective. It is used for planning to predict next actions based on the current cost function.
Level 4 processes encompass the adaption of the fusion process in relation to resource management by incorporating feedback to support the desired objectives i.e. selective processing of data and usage of sensors.
Level 5 processes are related to knowledge management. They are concerned with the assignment of different access rights to the data and user adaptive data visualization to support the cognitive decision making of humans

In this project sensor fusion is mainly conducted on Level 0 to operate on a most general level and allow a broad application of the sensor fusion results, also for other applications.

11.2.Literature research

Zhu et al. calculate in [183] a depth probability distribution for both sensor modalities and combines them using Markov Random Fields (MAP-MRF). The time-of-flight sensor is calibrated using a 4-D look-up table to map the measured intensity and 3-D coordinates to the ground truth distance given by the stereo camera pair. The Stereo cameras and time-of-flight sensor are calibrated to each other. To calculate the most likely disparity given the sensor measurements the data term of a global optimization algorithm for stereo vision is multiplied by a term describing the euclidean distance between the corresponding 3-D coordinates from stereo and time-of-flight sensor. In [184] Zhu et al. extend the fusion technology to the temporal domain by using a dynamic Markov Random Field to infer depth from both spatial and temporal neighbors. Like in their previous paper they calculate the most likely disparity given the sensor measurements, by augmenting the data term of a global optimization algorithm for stereo vision with a term describing the eucledian distance between the corresponding 3-D coordinates from stereo and time-of-flight sensor. To achieve temporal dependencies between different frames taken at different time, they propose a layered structure where each layer is itself a MRF and connections between the layers describe temporal dependencies. An additional smoothness term of the global optimization function is introduced to describe temporal smoothness.
Gudmundsson et al perform in [185] sensor fusion by calculating disparity estimates for stereo vision from the 3-D time-of-flight sensor. This constrains the stereo algorithm on a per pixel basis, resulting in more accurate disparity maps. All cameras are stereo calibrated to each other by down-scaling the color images to resolution of the time-of-flight sensor. The two color sensors are stereo calibrated to each other using their original image size. For each pixel of the time-of-flight sensor the corresponding left and right pixels of the color images are determined using the calculated homographies from stereo calibration. After up-scaling the color image to its original size, this gives an initial disparity estimate for the stereo camera pair used as a per pixel constraint for stereo matching.
Hahne et al. use in [186] a graph cut approach, to initialize the domain of the volumetric grid with the depth information from the low resolution time-of-flight camera to cut down computation time and increase accuracy of the depth estimation. Initially, the two color

cameras and the time-of-flight sensor are calibrated to each other to get their extrinsic and intrinsic parameters. To compute the desired depth map, the voxels of a 400x300x100 grid are associated with their corresponding depth values from the time-of-flight sensor. Those voxels, not present in the time-of-flight image are not considered during the graph cut procedure. During graph cut, the data term is extended by a term describing the difference between the assigned distance of the voxel and the measured distance from the time-of-flight sensor. Additionally, the smoothing term is extended to incorporate discontinuities in the time-of-flight depth image. The surface with minimizing the energy function is calculated using the standard graph cut algorithm. The required computation time is around a few minutes.
In [187], Hahne et al. combine a time-of-flight sensor and a stereo rig by using data from a time-of-flight sensor to limit the disparity search range for the stereo rig on demand. The proposed method is real-time capable and applies adaptive disparity search ranges on the stereo images based on the measured distance with the time-of-flight sensor to enable more reliable range values from stereo at depth discontinuities. The time-of-flight sensor and the stereo rig are extrinsically calibrated to each other. Prior to calculating disparity guesses from the time-of-flight sensor, each range value from the sensor is assigned a confidence value. The dominating cause for wrong depth values are surfaces with bad reflection properties, like dark areas. This information is available through the amplitude information of the sensor. Therefore, Hahne et al. apply a 3x3 median filter to preprocess the amplitude image and threshold the amplitude values to reject unreliable depth values. Rejected values are interpolated with neighboring reliable range values and further improved by applying the information from the stereo rig. To generate a dense depth estimate from stereo, the preprocessed time-of-flight values are used to generate a piecewise bilinear mesh. The mesh is transformed relative to the stereo rig's coordinate system and the intersection of a viewing ray from the stereo rig with the surface defines an initial disparity guess and a range according to the associated confidence value. Regions with a confidence value below a specified threshold are further processed using a standard correlation based stereo algorithm on the stereo images. For pixels with a valid confidence value, the disparity guess from the time-of-flight sensor is directly used to calculate the corresponding range for the stereo image.
Bartczak and Koch introduce in [188] a cost function for each pixel of a high resolution color image where the minima of the function's per pixel value corresponds to the optimal depth. At first the low resolution time-of-flight image is warped to fit the size of the high resolution stereo depth map by creating a triangle mesh from the 3-D time-of-flight data and re-projecting the mesh on the stereo image. Wrong 3-D information originating from the different viewpoints and depth discontinuities are detected by comparing a triangles normal with the stereo rig's line of view. Triangles with normals close to 90 degree are removed. Then, the authors define a cost function for each pixel over its local neighborhood that incorporates the squared distance between the depth measured by the time-of-flight sensor and the proposed depth as well as the color consistency of the left and right stereo image patch given the proposed distance. The depth value with the smallest cost is selected for each pixel.
Yang et al. present in [189] a post processing step to enhance the low resolution time-of-flight image, using one or two high resolution color images. Initially, the time-of-flight data is up-sampled to the size of the high resolution color images and provides the basis for a probability distribution of depth. From the Probability distribution a coarse 3-D cost volume is derived, which is enhanced during an iterative refinement process. The cost function calculates the squared difference between the selected and the proposed depth for each pixel. Then a bilateral filter is iteratively applied to each slice of the cost volume to smooth it while preserving edges, resulting in a new cost volume. The bilateral filter applies the assumption that neighboring pixels with similar color are likely to have similar depth. The depth with

minimal cost is selected for each pixel. To reduce discontinuities originating from the quantization effects in depth, a further sub-pixel estimation is applied after the minimal cost depth values have been selected. For each depth candidate its depth value is interpolated between the next larger and next lower possible range using a quadric function.

12. Machine Learning

12.1.Aim

To conduct a literature review of the current available robotics cognition technologies and an analysis of the techniques with respect of cognitive capabilities development for SRS.

12.2.Methodology

FP6/FP7 completed and on-going projects and wider literatures have been reviewed against key cognitive features of SRS, in particular, the following aspects:

1. Robot leaning, including imitation-based learning
2. Skill/knowledge generalisation and hierarchy
3. Mind-actions mapping.

Totally 18 FP6 projects and 58 FP7 projects which have the element of cognition were identified and studied. Research outcomes from these projects have been collected from the projects' websites and research papers published by the projects teams.

Research papers have also been reviewed.

12.3.Discovery

12.3.1. Robot learning

Learning plays an important role in the development of motivation-behaviour mapping, and skills and knowledge. GNOSYS aimed to develop architecture for cognitive agents, which integrates the "cycle of perception, knowledge acquisition, abstraction and reasoning. The abstraction element of the architecture would be "responsible for creating and organising a hierarchy of concepts". However, no paper is found from its website regarding the abstract mechanism [190]. SKILLS project focused on learning skills by the development of mappings between motion and forces [191-193]. ITALK project focuses on individual and social learning [194].

Imitation learning is a commonly used robot learning approach. The imitation based learning uses social cues such as pointing and gazing to indicate what the user intended to do next [195-197]. The user first teaches a robot by demonstrating gestures, for example, pointing to and gazing an object, to the robot. These gestures serve as social cues of his interest on the object. Then the robot imitates the gestures for the user's approval. This imitation process enables the robot to recognise the user's intention when it captures the same gestures. Experiments carried out in [197] can be described as below: During a first phase of the interaction, the designer demonstrated a gesture in front of a robot. The robot then observed the designer's gesture. Joint angles trajectories are collected from a motion sensor. The second phase was begun when the robot collected the different movements of user. The robot compared the gesture it collected with the gesture stored earlier and finds the cues of them. Then the robot pointed at an object that the user most likely to be interested. The robot then turned to user for evolution of its selection. The designer signals to the robot whether the same object has been selected by nodding/shaking his/her head. In the imitation learning, a Hidden Markov Model (HMM) with full covariance matrix is used to extract the characteristics of different gestures which are used later to recognise gestures from the user. The characteristic of a gesture is expressed by transition across the state of the HMM. Using such a model requires the estimation of a large set of parameters. An Expectation-Maximisation (EM) algorithm is used to estimate the HMM parameters. The estimation starts from initial estimates and converges to a local maximum of a likelihood function. It first performs a rough clustering. Next, EM is carried out to estimate a Gaussian Mixture Model (GMM). Finally, the transitions across the states are encoded in a HMM created with the GMM state distribution.

The most direct way to let the robot to understand the users' motivation is conversation. Hassch et al. [198] developed a Bielefeld Robot Companion (BIRON) which is a robot who accompanies to a human. It consists of cameras, microphones, laser range finder, speech recognition system, and other components. This robot is able to understand its users' intention through oral instructions and observation of the user's sight.

Tapus and Mataric [199] proposed a reinforcement learning based approach to robot behaviour adaptation. The aim of this approach is to develop a robotic system capable of adapting its behaviours according to the user's personality, preference, and profile in order to provide an engaging and motivating customised protocol.

Parental action demonstration to infant has also been studied for robot learning [200]. This is because in order to learn how to perform the task, robots have to appropriately select the relevant information while paying no or less attention to irrelevant and in action demonstration to infants, parents significantly modify their body movement to maintain the infants' attention and support their processing of the action, which leads to their better understanding of the action. Parental modification such as suppression and addition of their body movement might make the relevant features more salient than irrelevant. Such important aspects highlighted by parental actions can be detected by bottom-up visual attention. Even if robots do not know what are relevant or circumstantial, parental modifications would draw the robots' attention to the important aspects of the action. The attention model is able to detect outstanding locations in a scene in terms of primitive features. That is, the model does not require any a priori knowledge about the context nor the environment, but enables robots to detect likely important locations.

12.3.2. Skill/knowledge generalisation

Although handful projects claimed skill/knowledge abstraction as one of the objectives of, related research can only be found in the publications of two projects. PASCAL 2 developed a Support Vector Machines (SVM) based feature selection method. In this study features are ranked using K-Nearest-Neighbours combined with a forward selection. The result of the feature selection is tested on SVM to select the optimal number of features. This method was tested with the Outguess steganographic software and 14 features are selected while keeping the same classification performances, which confirms that the selected features are efficient for a wide variety of embedding rates [201]. IM-CLeVeR project aims to develop a cumulatively learning mechanism to develop hierarchical recursive architectures for organising behaviours, based on the divisibility of ongoing behaviour into discrete tasks, which are comprised of subtask sequences [202].

Skill/knowledge abstraction or generalisation is the procedure of picking up the common features among a group of similar skills or pieces of knowledge. Quantitatively modelling similarity is a real challenge. Approaches based on very different assumptions about how similarity should be represented have been reported [203].

If several attributes are observed for each sample, each of these attributes is called a *variable*. These variables can be used as co-ordinates to construct a *vector space*. Similarity measures can then be defined based on geometric representations in the space. Giving two objects, X and Y, that can be represented as two points in a space, the similarity between the two objects can then be considered as functions of the distance, dp(X, Y), in this space. Using Shepard's law, the similarity between X and Y is modelled as k(X, Y), an exponential function of the distance dp(X, Y). Deviating from Shepard's original formulation, the distance is nowadays often modified by taking it to the power of q, in order to give the model more flexibility. K(X, Y) is often referred as a similarity kernel. With p chosen to be 2, the similarity kernel is called a Gaussian kernel. The Gaussian kernel is extremely popular in machine learning. With p chosen to be 1, the similarity kernel is sometimes called a Laplacian kernel. The most widely used distance functions are the Minkowski distances such as $L_m(\mathbf{x},\mathbf{y}) = \left(\sum_{i=1}^{p} |x_i - y_i|^m \right)^{1/m}$ with ($m \geq 1$). These include the Euclidean distance (L_2), and the city-block or absolute value distance (L_1) [204].

Category learning also plays important role in generalisation, as similarity cannot guarantee a good generalisation though there is a tight relationship between similarity and generalization in machine learning. Category learning is often phrased as density estimation. Imagine two classes in which exemplars from each category are drawn from a probability density function that completely determines the distribution of features within each category. If a learner knew the distribution of features within a category, he or she could examine the features of a new stimulus and assign it to the category with the highest likelihood of having generated this pattern of features. Hence, learning to categorize could mean learning the distribution of features [205, 206].

12.3.3. Mind-actions mapping

Understanding of human behaviour is one of the objectives of PROMETHEUS project. The project has developed a combined Bayesian-Markovian approach for behaviour recognition. Markov chains are supplemented with additional low-level features taken from multiple sources, which are then combined in an efficient manner using Bayesian networks. The technique proposed in this paper uses a Bayesian network per Markov state. The Bayesian networks represent primitive events, in behavioural recognition terms, and are used to drive the transitional probability between Markov states. Outputs from the Bayesian networks are used to decide whether to transition to the next state or remain in the current state. Bayes-Markov chains reduce the complex problem by separating large networks into self-contained primitive events represented by small Bayes nets. Small Bayes nets represents primitive events can be reused in different behaviours. In Bayes- Markov chains, the task of understanding and interpreting information falls to the Bayesian networks, as no hidden model is used. The creation of a Markov model that incorporates all the behaviours possible for a given application can be very complex. To solve this problem, the concept of Bayes-Markov threads is used. Bayes-Markov threads run in a simultaneous manner, are updated as required, and share resources [207].

CO-FRIEND developed a novel architecture for high-level scene interpretation. The interpretation process based on the rules generated from the OWL ontology. First the primitive aggregates are added to the working memory of the Jess engine. Then the agenda (list of activations) of the Jess engine is analysed. If the agenda is not empty, the command is given to run the engine. The rules fire and add new facts, representing instances of higher level aggregates, to the working memory. Continuing this in a loop, more and more aggregates are instantiated, consistent with the corresponding conceptual constraints [208].

12.4. Assessment

The drawback of imitation learning is that it only allows a robot to respond to pre-defined social cues. Due to this reason, the robot actually forces its user to remember to use the same social cues given a certain circumstances when expressing his intentions, otherwise, the robot will not able to recognise. As SRS is proposed for assisting elderly who can find difficulties to remember social cues and associated circumstances, the effectiveness of the robot which is trained using imitation-based learning may be jeopardised. In addition, SRS is expected to be able to accumulate its skills and knowledge along side with being manipulated, an on-line learning mechanism is needed. Imitation learning reported in [195-197], on the other hand, is often used as off-line learning.

The key to the success of the reinforcement learning approach given in [199] is the award function. This function, on the other hand, also limits the application of this approach because its definition depends on how a robot's behaviour is parameterised. For fulfilling different tasks, this function may have to be defined differently.

Both category learning and similarity measure are useful in SRS project in terms of skill generalisation. The identification of features is essential to the definition of a vector space on which similarity measures can be defined. Category learning might be used for identifying features of skills from tasks remotely performed by people who manipulate a robot. An online approach, however, will be needed in SRS project to allow the system to carry out the category learning automatically.

SRS is a remotely controllable and semi-autonomous robotic system. Its property of autonomy could be achieved in the way that first the system finds out what tasks people who remotely manipulate the robot want to achieve and then uses its skills to fulfil the identified tasks. This cognitive capability means that the robot can reason about the mind of people who remotely control the robot via the first few actions they take, which has to be based on mappings from the mind to the actions.

Human behaviour understanding is important for building up this cognitive capability. The combined approach proposed in PROMETHEUS project is basically statistics-based and therefore relies on a large number of samples to train its Bayesian decision tree. The rule-based method developed in CO-FRIEND project also needs a large number of samples to build up its rule base. Although they can be combined with on-line learning to accumulatively improve the Bayesian decision tree and the rule base, in cases where the large numbers of samples are not available in the first place, interpretations developed using these approaches become not reliable. Both projects have not worked out solutions to the problem.

12.5. Conclusion

From psychological point of view, skill acquisition contains three stages, namely, cognitive stage, associative stage and autonomous stage. The semi-autonomous property of the SRS may also need to be implemented by following this procedure. Therefore, an SRS is able to identify its user's *intentions* (as well as task) according to the first few actions she/he takes and then to take over the fulfilment of the task using its own skills. It is important to enable the robot to identify its user's intention. Main-behaviours mappings and reasoning based on the mappings are essential to the intention identification. Equally important is generalised skills the robot possesses. Literatures reviewed show directions in both aspects to certain degrees but cannot used directly to develop semi-autonomous property for the SRS. Hybrid and innovative approaches will be needed to develop.

Reference:

1. Miskelly, F.G., *Assistive technology in elderly care.* Age and Ageing, 2001. **30**(6): p. 455-458.
2. Rejeski, W.J., et al., *Measuring disability in older adults: The International Classification System of Functioning, Disability and Health (ICF) framework.* Geriatrics and Gerontology International, 2008. **8**(1): p. 48-54.
3. *Assistive Products for Persons with Disability*, in *Classification and Terminology*. 2007, Switzerland: ISO Publications: International Organization for Standardization
4. (2007) *The Global Aging Experience Project – Ethnographic Research 2006*.
5. *The universal declaration of human rights*. 1948, United Nations New York.
6. *International Covenant on Economic, Social and Cultural Rights*. 1966, United Nations New York.
7. *International Covenant on Civil and Political Rights*. 1966, United Nations New York.
8. Sheridan, T., *Telerobotics, Automation, and Human Supervisory Control*. 1992, Cambridge, MA: MIT Press.
9. Hirzinger, G., *Multisensory shared autonomy and tele-sensor programming - Key issues in space robotics.* Robotics and Autonomous Systems, 1993. **11**(3-4): p. 141-162.
10. Brunner, B., K. Arbter, and G. Hirzinger. *Task directed programming of sensor based robots*. in *IEEE/RSJ/GI International Conference on Intelligent Robots and Systems*. 1994. Munich, Ger: IEEE.
11. Ortmaier, T., et al., *Motion estimation in beating heart surgery.* IEEE Transactions on Biomedical Engineering, 2005. **52**(10): p. 1729-1740.
12. Burke, J.L., et al., *Final report for the DARPA/NSF interdisciplinary study on human-robot interaction.* IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 2004. **34**(2): p. 103-112.
13. Nunes, R.J.C. *Home Automation - A Step Towards Better Energy Management*. in *International Conference on Renewable Energies and Power Quality*. 2003.
14. Sommer, T.J., *Telemedicine: A useful and necessary tool to improving quality of healthcare in the European Union.* Computer Methods and Programs in Biomedicine, 1995. **48**(1-2): p. 73-77.
15. Celler, B.G., et al. *Instrumentation system for the remote monitoring of changes in functional health status of the elderly at home*. in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. 1994. Baltimore, MD, USA: IEEE.
16. Tu, M.C., et al. *Fundamentals and design of smart home middleware*. in *Proceedings of the 2009 International Joint Conference on Computational Sciences and Optimization, CSO 2009*. 2009. Sanya, Hainan.
17. Farella, E., M. Falavigna, and B. Ricco, *Aware and smart environments: The Casattenta project.* Microelectronics Journal, 2010.
18. Hynes, M., et al. *Monitoring of activity levels of the elderly in home and community environments using off the shelf cellular handsets*. in *ICCE 2010 - 2010 Digest of Technical Papers International Conference on Consumer Electronics*. 2010. Las Vegas, NV.
19. Atallah, L., et al., *Real-time activity classification using ambient and wearable sensors.* IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(6): p. 1031-1039.

20. Anderson, D., et al., *Modeling human activity from voxel person using fuzzy logic*. IEEE Transactions on Fuzzy Systems, 2009. **17**(1): p. 39-49.
21. Zouba, N., F. Bremond, and M. Thonnat. *Multisensor fusion for monitoring elderly activities at home*. in *6th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS 2009*. 2009. Genova.
22. Kim, H.H., et al., *Resident location-recognition algorithm using a Bayesian classifier in the PIR sensor-based indoor location-aware system*. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 2009. **39**(2): p. 240-245.
23. Gaddam, A., S.C. Mukhopadhyay, and G.S. Gupta. *Smart home for elderly care using optimized number of wireless sensors*. in *Codec - 2009 - 4th International Conference on Computers and Devices for Communication*. 2009. Kolkata.
24. Fleury, A., et al. *Sound and speech detection and classification in a Health Smart Home*. in *Proceedings of the 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS'08 - "Personalized Healthcare through Technology"*. 2008. Vancouver, BC.
25. Estudillo-Valderrama, M.Á., et al., *Design and implementation of a distributed fall detection system-personal server*. IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(6): p. 874-881.
26. Noury, N., et al. *Remote follow up of health through the monitoring of electrical activities on the residential power line - Preliminary results of an experimentation*. in *2009 11th IEEE International Conference on e-Health Networking, Applications and Services, Healthcom 2009*. 2009. Sydney.
27. Chen, Y.H., et al. *Preference model assisted activity recognition learning in a smart home environment*. in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009*. 2009. St. Louis, MO.
28. Ali, F.M., et al. *Combined Fuzzy State Q-learning algorithm to predict context aware user activity under uncertainty in assistive environment*. in *Proc. 9th ACIS Int. Conf. Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, SNPD 2008 and 2nd Int. Workshop on Advanced Internet Technology and Applications*. 2008. Phuket.
29. Reguiga, K.S.B., N. Aitsaadi, and G. Pujolle. *An adaptive control algorithm in intelligent home networks*. in *2009 1st International Conference on Communications and Networking, ComNet 2009*. 2009. Hammamet.
30. Papamatthaiakis, G., G.C. Polyzos, and G. Xylomenos. *Monitoring and modeling simple everyday activities of the elderly at home*. in *7th IEEE Consumer Communications and Networking Conference, CCNC 2010*. 2010. Las Vegas, NV.
31. Lim, J.E., et al., *A context-aware fitness guide system for exercise optimization in U-health*. IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(3): p. 370-379.
32. Tamura, T., et al., *A wearable airbag to prevent fall injuries*. IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(6): p. 910-914.
33. Apiletti, D., et al., *Real-time analysis of physiological data to support medical applications*. IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(3): p. 313-321.
34. Park, S. and S. Jayaraman, *Smart textile-based wearable biomedical systems: A transition plan for research to reality*. IEEE Transactions on Information Technology in Biomedicine, 2010. **14**(1): p. 86-92.

35. Zakrzewski, M., et al. *Utilization of wireless sensor network for health monitoring in home environment*. in *Proceedings - 2009 IEEE International Symposium on Industrial Embedded Systems, SIES 2009*. 2009. Lausanne.
36. Adami, A.M., et al., *Detection of movement in bed using unobtrusive load cell sensors.* IEEE Transactions on Information Technology in Biomedicine, 2010. **14**(2): p. 481-490.
37. Shin, S.H., T. Hashimoto, and S. Hatano, *Automatic detection system for cough sounds as a symptom of abnormal health condition.* IEEE Transactions on Information Technology in Biomedicine, 2009. **13**(4): p. 486-493.
38. Uzunay, Y. and K. Bicakci. *SHA: A secure voice activated smart home for quadriplegia patients*. in *Proceedings - 2007 IEEE International Conference on Bioinformaticsand Biomedicine Workshops, BIBMW*. 2007. San Jose, CA.
39. Park, S. and J.K. Aggarwal, *A hierarchical Bayesian network for event recognition of human actions and interactions.* Multimedia Systems, 2004. **10**(2): p. 164-179.
40. Bonino, D., et al. *Technology independent interoperation of domotic devices through rules*. in *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*. 2009. Kyoto.
41. Zamora-Izquierdo, M.A., J. Santa, and A.F. Gómez-Skarmeta, *Integral and Networked Home Automation Solution towards Indoor Ambient Intelligence*, in *IEEE Pervasive Computing*. 2010.
42. Balasubramanian, K. and A. Cellatoglu, *Improvements in home automation strategies for designing apparatus for efficient smart home.* IEEE Transactions on Consumer Electronics, 2008. **54**(4): p. 1681-1687.
43. Hsu, C.C. and L.Z. Wang. *A smart home resource management system for multiple inhabitants by agent conceding negotiation*. in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. 2008. Singapore.
44. Hong, K.S., H.J. Kim, and C. Lee. *Automated grocery ordering systems for smart home*. in *Proceedings of Future Generation Communication and Networking, FGCN 2007*. 2007. Jeju Island.
45. Intille, S.S., *Designing a home of the future.* IEEE Pervasive Computing, 2002. **1**(2): p. 76-82.
46. Fang, A., et al. *The realization of intelligent home by zigbee wireless network technology*. in *Proceedings of the 2009 Pacific-Asia Conference on Circuits, Communications and System, PACCS 2009*. 2009. Chengdu.
47. Tsou, Y.P., et al. *Building a remote supervisory control network system for smart home applications*. in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. 2007. Taipei.
48. Amoretti, M., et al. *E-Inclusive Videoconference Services in Ambient Assisted Living Environments*. in *3rd European Conference on Ambient Intelligence*. 2009.
49. Emparanza, J.P., et al. *Ambient telephony: Designing a communication system for enhancing social presence in home mediated communication.* in *Proceedings - 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, ACII 2009*. 2009. Amsterdam.
50. Hsu, J.M. and I.R. Chang. *Design a virtual object representing human-machine interaction for music playback control in smart home*. in *Proceedings of the International Conference on Complex, Intelligent and Software Intensive Systems, CISIS 2009*. 2009. Fukuoka.
51. Jaijongrak, V.R., S. Chantasuban, and S. Thiemjarus. *Towards a BSN-based gesture interface for intelligent home applications*. in *ICCAS-SICE 2009 - ICROS-SICE International Joint Conference 2009*. 2009. Fukuoka.

52. Son, Y.S. and K.D. Moon. *Home energy management system based on power line communication*. in *ICCE 2010 - 2010 Digest of Technical Papers International Conference on Consumer Electronics*. 2010. Las Vegas, NV.
53. Li, B., P. Hathaipontaluk, and S. Luo. *Intelligent oven in smart home environment*. in *ICRCCS 2009 - 2009 International Conference on Research Challenges in Computer Science*. 2009. Shanghai.
54. Pishva, D. and K. Takeda, *Product-based security model for smart home appliances.* IEEE Aerospace and Electronic Systems Magazine, 2008. **23**(10): p. 32-41.
55. Wang, C.D., X.L. Mo, and H.B. Wang. *An intelligent home middleware system based on context- Awareness*. in *5th International Conference on Natural Computation, ICNC 2009*. 2009. Tianjian.
56. Filipescu, A., et al. *Control of mobile platforms as robotic assistants for elderly*. in *Proceedings of 2009 7th Asian Control Conference, ASCC 2009*. 2009. Hong Kong.
57. Yu, W., et al., *Design and implementation of a ubiquitous robotic space.* IEEE Transactions on Automation Science and Engineering, 2009. **6**(4): p. 633-640.
58. Li, H., L. Kong, and P. Wu. *Detecting Abnormal State of Elderly for Service Robot with H-FCM*. in *IEEE International Conference on Automation and Logistics Shenyang*. 2009. China.
59. Tsai, C.C., et al. *Human-robot interaction of an active mobile robotic assistant in intelligent space environments*. in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*. 2009. San Antonio, TX.
60. Prenzel, O., J. Feuser, and A. GrÃ¤ser. *Rehabilitation robot in intelligent home environment - Software architecture and implementation of a distributed system*. in *Proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics*. 2005. Chicago, IL.
61. Papadopoulos, N., et al. *A connected home platform and development framework for smart home control applications*. in *IEEE International Conference on Industrial Informatics (INDIN)*. 2009. Cardiff.
62. Gouin-Vallerand, C. and S. Giroux. *Managing and deployment of applications with OSGi in the context of Smart Homes*. in *3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob 2007*. 2007. White Plains, NY.
63. Helal, S., et al. *Enabling Location-Aware Pervasive Computing Applications for the Elderly*. in *First IEEE International Conference on Pervasive Computing and Communications* 2003.
64. Baker, C.R., et al. *ZUMA: A platform for smart-home environments*. in *IET Conference Publications*. 2006. Athens.
65. Perry, M., et al., *Multimodal and ubiquitous computing systems: Supporting independent-living older users.* IEEE Transactions on Information Technology in Biomedicine, 2004. **8**(3): p. 258-270.
66. Li, C., Y. Liao, and Y.H. Hung. *A study on the different recreational demands of intelligent environment technology among the three elderly groups*. in *2009 11th IEEE International Conference on e-Health Networking, Applications and Services, Healthcom 2009*. 2009. Sydney.
67. Benoit, O., et al. *User-centered activity management system for elderly people: Empowering older people with interactive technologies to manage their activities at the retirement home*. in *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare - Pervasive Health 2009, PCTHealth 2009*. 2009. London.

68. Krumm, J., et al. *Steve Shafer, Multi-Camera Multi-Person Tracking for EasyLiving*. in *The Third IEEE International Workshop on Visual Surveillance*. 2000.
69. Böhlen, M., *Second order ambient intelligence.* Journal of Ambient Intelligence and Smart Environments 2009. **1**: p. 1-5.
70. CeNSE. *Earth calling: Turn off the lights!* 2010; Available from: http://www.hpl.hp.com/news/2009/oct-dec/cense.html.
71. *Open Jaus – Jaus Robotics Software Development Kit (SDK).* 2010; Available from: http://www.openjaus.com/.
72. *ORCA Robotics*. 2010; Available from: http://orca-robotics.sourceforge.net.
73. *The OROCOS Project*. 2010; Available from: http://www.orocos.org.
74. *Documentation – ROS Wiki*. 2010; Available from: http://www.ros.org/wiki.
75. *Player Project*. 2010; Available from: http://playerstage.sourceforge.net.
76. *Microsoft Robotics*. 2010; Available from: http://msdn.microsoft.com/en-us/robotics/default.aspx.
77. *CLARAty*. 2010; Available from: http://claraty.jpl.nasa.gov/man/overview/index.php.
78. *YARP: Welcome to YARP*. 2010; Available from: http://eris.liralab.it/yarpdoc/index.html.
79. *CARMEN*. 2010; Available from: http://carmen.sourceforge.net.
80. *MOOS*. 2010; Available from: http://www.robots.ox.ac.uk/~mobile/MOOS/wiki/pmwiki.php.
81. *Robocomp*. 2010; Available from: http://sourceforge.net/apps/mediawiki/robocomp/index.php?title=Main_Page.
82. *MARIE*. 2010; Available from: http://marie.sourceforge.net/wiki/index.php/MARIE.
83. *RobotCub – An open framework for research in embodied cognition*. 2010; Available from: http://www.robotcub.org.
84. *iCub*. 2010; Available from: http://www.icub.org.
85. *RoSta: Project*. 2010; Available from: http://www.robot-standards.org/index.php?id=8.
86. Asimov, I., *The Foundation Trilogy: Foundation, Foundation and Empire; Second Foundation - The Stars, Like Dust, The Naked Sun, I Robot* 1982, New York Octopus/Heinemann.
87. Clarke, R., *Asimov's laws of robotics: implications for information technology.* Computer, 1993. **26**(12): p. 53-61.
88. Haddadin, S., A. Albu-SchäCurrency Signffer, and G. Hirzinger, *Requirements for safe robots: Measurements, analysis and new insights.* International Journal of Robotics Research, 2009. **28**(11-12): p. 1507-1527.
89. De Santis, A., et al., *Human-robot interaction control using force and vision*, in *Lecture Notes in Control and Information Sciences*, C. Bonivento, C. Rossi, and L. Marconi, Editors. 2007. p. 51-70.
90. Meng, Q. and M.H. Lee, *Design issues for assistive robotics for the elderly.* Advanced Engineering Informatics, 2006. **20**(2): p. 171-186.
91. Alami, R., et al., *Safe and Dependable Physical Human-Robot Interaction in Anthropic Domains: State of the Art and Challenges*, in *IEEE/RSJ Conference on Intelligent Robots and Systems*. 2006: Beijing China.
92. *MOVEMENT Project Homepage*. 2010; Available from: http://www.is.tuwien.ac.at/fortec/reha.e/projects/movement/
93. *IROMEC Project Homepage*. 2010; Available from: http://www.iromec.org.
94. Dong, H., S. Keates, and P.J. Clarkson. *Accommodating older users' functional capabilities*. in *Workshop on a New Research Agenda for Older Adults*. 2002. London, UK.

95. Scholtz, J. *Theory and evaluation of human robot interactions*. in *System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on*. 2003.
96. Kiesler, S. *Fostering common ground in human-robot interaction*. in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*. 2005. Nashville, TN.
97. Arcara, P. and C. Melchiorri, *Control schemes for teleoperation with time delay: A comparative study*. Robotics and Autonomous Systems, 2002. **38**(1): p. 49-64.
98. Burmester, M., *Optimierung der Erlern- und Benutzbarkeit von Benutzungsschnittstellen interaktiver Hausgeräte auf der Basis der speziellen Anforderungen älterer Menschen*. 2001, Düsseldorf: VDI Verlag.
99. Ryu, H. and W. Lee. *Where you point is where the robot is*. in *ACM International Conference Proceeding Series*. 2006. Christchurch.
100. Kato, J., et al. *Multi-touch interface for controlling multiple mobile robots*. in *Conference on Human Factors in Computing Systems - Proceedings*. 2009. Boston, MA.
101. *Freie Universität Berlin, Autonomous Car Project*. 2010; Available from: http://robotics.mi.fu-berlin.de/pmwiki/pmwiki.php.
102. Young, J.E., M. Xin, and E. Sharlin. *Robot expressionism through cartooning*. in *HRI 2007 - Proceedings of the 2007 ACM/IEEE Conference on Human-Robot Interaction - Robot as Team Member*. 2007. Arlington, VA.
103. Ricks, B., C.W. Nielsen, and M.A. Goodrich. *Ecological displays for robot interaction: a new perspective*. in *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. 2004.
104. Giesler, B., et al. *Sharing skills: Using augmented reality for human-Robot collaboration*. in *Proceedings of SPIE - The International Society for Optical Engineering*. 2004. San Jose, CA.
105. Green, S.A., et al., *Augmented Reality for Human-Robot Collaboration*, in *Human-Robot Interaction* N. Sarkar, Editor. 2007, Itech Education and Publishing: Vienna. p. 65-94.
106. Dillmann, R., et al. *Interactive Natural Programming of Robots: Introductory Overview*. in *2nd IARP/IEEE-RAS Joint Workshop on Technical Challenge for Dependable Robots in Human Environments*. 2002. Toulouse, France.
107. Ehrenmann, M., et al. *Programming service tasks in household environments by human demonstration*. in *Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on*. 2002.
108. Fong, T. and I. Nourbakhsh, *Interaction challenges in human-robot space exploration*. interactions, 2005. **12**(2): p. 42-45.
109. Song, T.H., et al. *A Study on usability of human-robot interaction using a mobile computer and a human interface device*. in *ACM International Conference Proceeding Series*. 2007. Singapore.
110. Spiliotopoulos, D., I. Androutsopoulos, and C.D. Spyropoulos. *Human-Robot Interaction Based On Spoken Natural Language Dialogue*. in *European Workshop on Service and Humanoid Robots* 2001.
111. Roy, N., et al. *Towards Personal Service Robots for the Elderly*. in *Workshop on Interactive Robots and Entertainment (WIRE)*. 2000. Pittsburgh, Pennsylvania, USA.
112. Parlitz, C., et al., *Intuitive human-machine-interaction and implementation on a household robot companion*, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2007: Beijing. p. 922-929.

113. Kulyukin, V.A. *On natural language dialogue with assistive robots*. in *HRI 2006: Proceedings of the 2006 ACM Conference on Human-Robot Interaction*. 2006. Salt Lake City, Utah.
114. Torrey, C., S.R. Fussell, and S. Kiesler. *Appropriate accommodations: Speech technologies and the needs of older adults*. in *AAAI Fall Symposium - Technical Report*. 2005. Arlington, VA.
115. Glassmire, J., et al. *Cooperative manipulation between humans and teleoperated agents*. in *Proceedings - 12th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, HAPTICS*. 2004. Chicago, IL.
116. Horiguchi, Y., T. Sawaragi, and G. Akashi. *Naturalistic human-robot collaboration based upon mixed-initiative interactions in teleoperating environment*. in *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*. 2000.
117. Fernández, V., et al. *Active human-mobile manipulator cooperation through intention recognition*. in *Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on*. 2001.
118. Schätzlein, F., et al. *Evaluation of interaction techniques for AAL in virtual reality with senior citizens*. in *3rd European Conference on Ambient Intelligence (AmI09)*. 2009. Salzburg, Austria.
119. Wang, M. and J.N.K. Liu, *Interactive control for Internet-based mobile robot teleoperation.* Robotics and Autonomous Systems, 2005. **52**(2-3): p. 160-179.
120. Brooks, A.G. and C. Breazeal, *Working with robots and objects: revisiting deictic reference for achieving spatial common ground*, in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 2006, ACM: Salt Lake City, Utah, USA.
121. Weiss, A., et al. *Teaching a humanoid: A user study on learning by demonstration with HOAP-3*. in *Robot and Human Interactive Communication, 2009. RO-MAN 2009. The 18th IEEE International Symposium on*. 2009.
122. Weiss, A., et al., *Autonomous vs. tele-operated: how people perceive human-robot collaboration with hrp-2*, in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. 2009, ACM: La Jolla, California, USA.
123. Monkman, G.J., *An electrorheological tactile display.* Presence: Teleoperators & Virtual Environments, 1992. **1**: p. 10.
124. Klein, D., et al., *Modelling the response of a tactile array using electrorheological fluids.* Journal of Physics D: Applied Physics, 2004. **37**(5): p. 794-803.
125. Burdea, G.C., *Force and Touch Feedback for Virtual Reality*. 1 ed. 1996, New York: Wiley-Interscience.
126. Lin, M.C., *Haptic Rendering: Foundations, Algorithms and Applications*, ed. M.O. Ming C. Lin. 2008, London: A K Peters.
127. *WIKI Computer Vision*. 2010; http://en.wikipedia.org/wiki/Computer_vision].
128. Cédras, C. and M. Shah, *Motion-based recognition a survey.* Image and Vision Computing, 1995. **13**(2): p. 129-155.
129. Ju, S., *Human Motion Estimation and Recognition (Depth Oral Report).* Technical report, University of Toronto, 1996.
130. Aggarwal, J.K., et al., *Nonrigid Motion Analysis: Articulated and Elastic Motion.* Computer Vision and Image Understanding, 1998. **70**(2): p. 142-156.
131. Aggarwal, J. and Q. Cai, *Human motion analysis: A review.* Human motion analysis: A review. Comput Vis Image Understand, 1999: p. 295-304.
132. Gavrila, D.M., *The visual analysis of human movement: A survey.* Computer Vision and Image Understanding, 1999. **73**(1): p. 82-98.

133. Moeslund, T.B. and E. Granum, *A survey of computer vision-based human motion capture*. Computer Vision and Image Understanding, 2001. **81**(3): p. 231-268.
134. Wang, L., W. Hu, and T. Tan, *Recent developments in human motion analysis*. Pattern Recognition, 2003. **36**(3): p. 585-601.
135. Buxton, H., *Learning and understanding dynamic scene activity: A review*. Image and Vision Computing, 2003. **21**(1): p. 125-136.
136. Hu, W., et al., *A survey on visual surveillance of object motion and behaviors*. IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews, 2004. **34**(3): p. 334-352.
137. Aggarwal, J.K. and S. Park. *Human motion: Modeling and recognition of actions and interactions*. in *Proceedings - 2nd International Symposium on 3D Data Processing, Visualization, and Transmission. 3DPVT 2004*. 2004. Thessaloniki.
138. Moeslund, T., A. Hilton, and V. Kruger, *A survey of advances in vision-based human motion capture and analysis*. Computer Image and Vision Understanding, 2006. **103**(2-3): p. 90-126.
139. Fernández, C. and J. Gonzàlez, *Ontology for semantic integration in a cognitive surveillance system*, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2007. p. 260-263.
140. Fernández, C., et al., *Interpretation of complex situations in a semantic-based surveillance framework*. Signal Processing: Image Communication, 2008. **23**(7): p. 554-569.
141. Kojima, A., et al., *Generating natural language description of human behavior from video images*. Proceedings - International Conference on Pattern Recognition, 2000. **15**(4): p. 728-731.
142. Kojima, A., T. Tamura, and K. Fukunaga, *Natural language description of human activities from video images based on concept hierarchy of actions*. International Journal of Computer Vision, 2002. **50**(2): p. 171-184.
143. Nevatia, R., T. Zhao, and S. Hongeng, *Hierarchical language-based representation of events in video streams*. IEEE Conf. on Computer Vision and Pattern Recognition, 2003. **4**: p. 39.
144. Baiget, P., et al., *Observing human behavior in image sequences: the video-hermeneutics challenge*, in *Computer Vision: Advances in Research & Development (CVCRD'2008)*. 2008: Bellaterra, Spain.
145. Cao, Y., L. Tao, and G. Xu. *An Event-driven Context Model in Elderly Health Monitoring*. in *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing (UIC-ATC '09)* 2009.
146. Remagnino, P., T. Tan, and K. Baker, *Multi-agent visual surveillance of dynamic scenes*. Image and Vision Computing, 1998. **16**(8): p. 529-532.
147. Liu, Z.Q., et al., *Dynamic image sequence analysis using fuzzy measures*. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2001. **31**(4): p. 557-572.
148. Karl, S., fer, and B. Christoph, *"F-Limette" fuzzy logic programming integrating metric temporal extensions*. Journal of Symbolic Computation, 1996. **22**(5-6): p. 725-727.
149. Kamp, H. and U. Reyle, *From Discourse to Logic*. 1993, London: Kluwer Academic Publishers.
150. Alfred, V.A. and D.U. Jeffrey, *The theory of parsing, translation, and compiling*. 1972, Englewood Cliffs, NJ: Prentice-Hall, Inc.

151. Borzin, A., E. Rivlin, and M. Rudzsky. *Surveillance event interpretation using generalized stochastic Petri Nets*. in *8th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2007)*. 2007.
152. Hartz, J. and B. Neumann. *Learning a knowledge base of ontological concepts for high-level scene interpretation*. in *Proceedings of 6th International Conference on Machine Learning and Applications (ICMLA 2007)*. 2007.
153. Shet, V.D., D. Harwood, and L.S. Davis. *VidMAP: Video monitoring of activity with prolog*. in *Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS 2005)*. 2005.
154. Ryoo, M.S. and J.K. Aggarwal, *Semantic representation and recognition of continued and recursive human activities.* International Journal of Computer Vision, 2009. **82**(1): p. 1-24.
155. Ivanov, Y.A. and A.F. Bobick, *Recognition of visual activities and interactions by stochastic parsing.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. **22**(8): p. 852-872.
156. Rusu, R.B., et al. *Aligning point cloud views using persistent feature histograms*. in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*. 2008. Nice.
157. Nuchter, A., H. Surmann, and J. Hertzberg. *Automatic model refinement for 3D reconstruction with mobile robots*. in *3-D Digital Imaging and Modeling, 2003. 3DIM 2003. Proceedings. Fourth International Conference on*. 2003.
158. Rusu, R.B., et al. *Persistent Point Feature Histograms for 3D Point Clouds*. in *the 21st IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2008.
159. Andreasson, H. and A. Lilienthal. *Vision Aided 3D Laser Scanner Based Registration*. in *3rd European Conference on Mobile Robots*. 2007.
160. Huhle, B., et al. *Registration of colored 3D point clouds with a Kernel-based extension to the normal distributions transform*. in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. 2008.
161. Mure-Dubois, J. and H. Hügli, *Fusion of Time of Flight Camera Point Clouds*, in *Workshop on Multi-camera and Multi-modal Sensor Fusion Algorithms and Applications - M2SFA2 2008*. 2008: Marseille, France.
162. Montemerlo, M., et al. *FastSLAM: A factored solution to the simultaneous localization and mapping problem*. in *Proceedings of the National Conference on Artificial Intelligence*. 2002. Edmonton, Alta.
163. Bay, H., et al., *Speeded-Up Robust Features (SURF).* Computer Vision and Image Understanding, 2008. **110**(3): p. 346-359.
164. Ozuysal, M., et al., *Fast Keypoint Recognition Using Random Ferns.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2010. **32**(3): p. 448-461.
165. Rusu, R.B., N. Blodow, and M. Beetz. *Fast Point Feature Histograms (FPFH) for 3D registration*. in *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*. 2009.
166. Triebel, R., O. Mozos, and W. Burgard. ***Relational Learning in Mobile Robotics: An Application to Semantic Labeling of Objects in 2D and 3D Environment Maps***. in *31th Annual Conference of the Gesellschaft für Klassifikation, 2003*. 2003.
167. Vasudevan, S., et al., *Cognitive maps for mobile robots-an object based approach.* Robotics and Autonomous Systems, 2007. **55**(5): p. 359-371.
168. Rusu, R.B., et al., *Towards 3D Point cloud based object maps for household environments.* Robotics and Autonomous Systems, 2008. **56**(11): p. 927-941.

169. Chella, A., M. Frixione, and S. Gaglio, *A cognitive architecture for robot self-consciousness.* Artificial Intelligence in Medicine, 2008. **44**(2): p. 147-154.
170. Gravot, F., et al. *Cooking for humanoid robot, a task that needs symbolic and geometric reasonings*. in *Proceedings - IEEE International Conference on Robotics and Automation*. 2006. Orlando, FL.
171. Kavraki, L.E., et al., *Probabilistic roadmaps for path planning in high-dimensional configuration spaces.* Robotics and Automation, IEEE Transactions on, 1996. **12**(4): p. 566-580.
172. La Valle, S., *Rapidly-exploring random trees: A new tool for path planning*, in *Technical Report TR 98-11*. 1998.
173. Lingelbach, F. *Path planning using probabilistic cell decomposition*. in *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*. 2004.
174. Brock, O. and O. Khatib, *Elastic strips: A framework for motion generation in human environments.* International Journal of Robotics Research, 2002. **21**(12): p. 1031-1052.
175. Yang, Y. and O. Brock. *Elastic roadmaps: Globally task-consistent motion for autonomous mobile manipulation in dynamic environments*. in *Robotics: Science and Systems*. 2006. Philadelphia, USA.
176. Belghith, K., et al. *Anytime dynamic path-planning with flexible probabilistic roadmaps*. in *Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on*. 2006.
177. Jaillet, L. and T. Simeon. *A PRM-based motion planner for dynamically changing environments*. in *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. 2004.
178. Bayle, B., J.Y. Fourquet, and M. Renaud. *Using manipulability with nonholonomic mobile manipulators*. in *International Conference on Field and Service Robotics*. 2001.
179. Zacharias, F., C. Borst, and G. Hirzinger. *Capturing robot workspace structure: representing robot capabilities*. in *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*. 2007.
180. Sturm, J., et al., *Learning kinematic models for articulated objects*, in *Proceedings of the 21st international jont conference on Artifical intelligence*. 2009, Morgan Kaufmann Publishers Inc.: Pasadena, California, USA.
181. Steinberg, A.N., C.L. Bowman, and F.E. White, *Revisions to the JDL data fusion model.* Proceedings of SPIE - The International Society for Optical Engineering, 1999. **3719**: p. 430-441.
182. Bedworth, M. and J. O'Brien, *Omnibus model: a new model of data fusion?* IEEE Aerospace and Electronic Systems Magazine, 2000. **15**(4): p. 30-36.
183. Zhu, J., et al. *Fusion of time-of-flight depth and stereo for high accuracy depth maps*. in *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*. 2008. Anchorage, AK.
184. Zhu, J., et al., *Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010. **32**(5): p. 899-909.
185. Gudmundsson, S.A., H. Aanaes, and R. Larsen, *Fusion of stereo vision and Time-Of-Flight imaging for improved 3D estimation.* Int. J. Intell. Syst. Technol. Appl., 2008. **5**(3/4): p. 425-433.
186. Hahne, U. and M. Alexa, *Combining Time-of-Flight Depth and Stereo Images without Accurate Extrinsic Calibration.* Int. J. Intell. Syst. Technol. Appl., 2008. **5**(3/4): p. 325-333.

187. Hahne, U. and M. Alexa, *Depth Imaging by Combining Time-of-Flight and On-Demand Stereo*, in *Proceedings of the DAGM 2009 Workshop on Dynamic 3D Imaging*. 2009, Springer-Verlag: Jena, Germany.
188. Bartczak, B. and R. Koch, *Dense Depth Maps from Low Resolution Time-of-Flight Depth and High Resolution Color Views*, in *Proceedings of the 5th International Symposium on Advances in Visual Computing: Part II*. 2009, Springer-Verlag: Las Vegas, Nevada.
189. Yang, Q., et al. *Spatial-Depth Super Resolution for Range Images*. in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*. 2007.
190. *GNOSYS Vision*. 2010; Available from: http://www.ics.forth.gr/gnosys/main/main_frames.html
191. *SKILLS*. 2010; Available from: http://www.skills-ip.eu.
192. Ruffaldi, E., et al., *Vibrotactile Perception Assessment for a Rowing Training System.* World Haptics 2009: Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, Proceedings, 2009: p. 350-355.
193. Filippeschi, A., et al., *Dynamic models of team rowing for a virtual environment rowing training system.* The International Journal of Virtual Reality, 2009. **8**(4): p. 19-26.
194. *Italk*. 2010; Available from: http://www.italkproject.org/.
195. Dillmann, R., *Teaching and learning of robot tasks via observation of human performance.* Robotics and Autonomous Systems, 2004. **47**(2-3): p. 109-116.
196. Breazeal, C., et al., *Learning From and About Others: Towards Using Imitation to Bootstrap the Social Understanding of Others by Robots.* Artif. Life, 2005. **11**(1-2): p. 31-62.
197. Calinon, S. and A. Billard, *Teaching a Humanoid Robot to Recognize and Reproduce Social Cues.* The 15th IEEE International Symposium on Robot and Human Interactive Communication, 2006: p. 346-351
198. Hassch, A., et al., *BIRON-The bielefeld robot companion.* Workshop on Advances in Service Robotics, 2004: p. 27-32.
199. Tapus, A. and M. Mtaraic, *Towards active learning for socially assistive robots.* Robotics Challenges for Machine Learning Workshop, The 21st Annual Conference on Natural Information Proceeding Systems, 2007.
200. Nagai, Y. and K. Rohlfing, *Computational Analysis of Motionese Toward Scaffolding Robot Action Learning.* IEEE Transactions on Autonomous Mental Development, 2009. **1**(1): p. 44-54.
201. Miche, Y., et al., *A feature selection methodology for steganalysis.* Multimedia Content Representation, Classification and Security, 2006. **4105**: p. 49-56.
202. Botvinick, M.M., Y. Niv, and A.C. Barto, *Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective.* Cognition, 2009. **113**(3): p. 262-280.
203. Navarro, D.J., *Representing simulus similarity*. 2002, University of Adelaide, Dept. of Psychology: Adelaide.
204. Borg, I. and P.J.F. Groenen, *Modern Multidimensional Scaling: Theory and Applications (Second Edition)*. Vol. New York. 2005: Springer. 614.
205. Jakel, F., B. Scholkopf, and F.A. Wichmann, *A tutorial on kernel methods for categorization.* Journal of Mathematical Psychology, 2007. **51**(6): p. 343-358.
206. Jakel, F., B. Scholkopf, and F.A. Wichmann, *Generalization and similarity in exemplar models of categorization: Insights from machine learning.* Psychonomic Bulletin & Review, 2008. **15**(2): p. 256-271.

207. Carter, N., D. Young, and J. Ferryman, *A combined Bayesian Markovian approach for behaviour recognition.* 18th International Conference on Pattern Recognition, Vol 1, Proceedings, 2006: p. 761-764.

208. Bohlken, W. and B. Neumann, *Generation of Rules from Ontologies for High-Level Scene Interpretation*, in *Proceedings of the 2009 International Symposium on Rule Interchange and Applications*. 2009, Springer-Verlag: Las Vegas, Nevada. p. 93-107.