IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université Joseph Fourier (Grenoble)

Activity Report 2014

Project-Team PRIMA

Perception, recognition and integration for smart environments

IN COLLABORATION WITH: Laboratoire d'Informatique de Grenoble (LIG)

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, perception and multimedia interpretation

Table of contents

Project-Team PRIMA

Keywords: Computer Vision, Interaction, Perception, Artificial Intelligence, Human-Robot Interaction, Context-Aware Computing

Creation of the Project-Team: 2003 January 01.

1. Members

Research Scientists
Sabine Coquillart [Inria, Senior Researcher, HdR]
Frédéric Devernay [Inria, Researcher]
Thierry Fraichard [Inria, Researcher, HdR]

Faculty Members
James Crowley [Team leader, INP Grenoble, Professor]
Alexandre Demeure [Univ. Grenoble I, Associate Professor]
Augustin Lux [INP Grenoble, Professor, HdR]
Patrick Reignier [INP Grenoble, Associate Professor, HdR]
Dominique Vaufreydaz [Univ. Grenoble II, Associate Professor]

Engineers
Fabien Badeig [Inria, from Mar 2014]
Raffaella Balzarini [Inria, from Apr 2014]
Morgan Bidois [Univ. Grenoble I, from Sep 2014]
Nicolas Bonnefond [Inria, AMIQUAL4HOME]
Stan Borkowski [Inria, AMIQUAL4HOME]
Jean-Rene Courtois [Univ. Grenoble I]
Mayeul de Werbier d Antigneul [Inria, AMIQUAL4HOME and EIT ICT LABS]
Alexandre Gauthier-Foichat [Inria]
Eméric Grange [Inria, from Mar 2014]
Cédric Gérard [Univ. Grenoble I, CATRENE AppsGate until May 2014]
Camille Lenoir [Univ. Grenoble I, CATRENE AppsGate]
Lucas Nacsa [Inria, Project with Schneider Electric]
Amaury Negre [CNRS]
Kouzma Petoukhov [Univ. Grenoble I, CATRENE AppsGate]
Rémi Pincent [Inria, AMIQUAL4HOME]

PhD Students
Evanthia Mavridou [Inria, until Nov 2014]
Etienne Balit [Inria]
Varun Jain [Inria, EIT ICT LABS until Dec 2014]
Dimitri Masson [Univ. Grenoble II, until Sep 2014]
Grégoire Nieto [Inria, from Oct 2014]
Rémi Paulin [Univ. Grenoble I]
Sergi Pujades-Rocamora [Inria]
Julian Quiroga [Univ. Grenoble I, until Nov 2014]
Viet Cuong Ta [Univ. Grenoble I, from Oct 2014]
Bruno Willems [Univ. Grenoble I, from Oct 2014]

Administrative Assistant
Catherine Bessiere [Inria]

Others

Ankuj Arora [Masters Student, Inria, from Feb 2014 until Sep 2014]
Pauline Andru [Inria, Intern, from Feb 2014]
Luc Ballot [Inria, ENSIMAG Student, from Feb 2014 until Jul 2014]
Ahmad Mustofa Hadi [Inria, Masters Student, from Feb 2014 until Aug 2014]
Ioanna Lampraki [Inria, Masters Student, from Feb 2014 until Aug 2014]
Maxime Portaz [Inria, Master Student]
Romain Bregier [Student Intern, from Sep 2014]
Jingtao Chen [Univ. Grenoble I, from Mar 2014]
Joelle Coutaz [Collaboratrice Externe, Univ. Grenoble I]

2. Overall Objectives

2.1. Perception, Recognition and Multimodal Interaction for Smart Spaces.

The objective of Project PRIMA is to develop the scientific and technological foundations for human environments that are capable of perceiving, acting, communicating, and interacting with people in order to provide services. The construction of such environments offers a rich set of problems related to interpretation of sensor information, learning, machine understanding, dynamic composition of components and man-machine interaction. Our goal is make progress on the theoretical foundations for perception and cognition, as well as to develop new forms of man machine interaction, by using interactive environments as a source of example problems.

An environment is a connected volume of space. An environment is said to be "perceptive" when it is capable of recognizing and describing things, people and activities within its volume. Simple forms of applications-specific perception may be constructed using a single sensor. However, to be general purpose and robust, perception must integrate information from multiple sensors and multiple modalities. Project PRIMA creates and develops machine perception techniques fusing computer vision, acoustic perception, range sensing and mechanical sensors to enable environments to perceive and understand humans and human activities.

An environment is said to be "active" when it is capable of changing its internal state. Common forms of state change include regulating ambient temperature, acoustic level and illumination. More innovative forms include context-aware presentation of information and communications, as well as services for cleaning, materials organisation and logistics. The use of multiple display surfaces coupled with location awareness offers the possibility of automatically adapting information display to fit the current activity of groups. The use of activity recognition and acoustic topic spotting offers the possibility to record a log of human to human interaction, as well as to provide relevant information without disruption. The use of steerable video projectors (with integrated visual sensing) offers the possibilities of using any surface for presentation, interaction and communication.

An environment may be considered as "interactive" when it is capable of interacting with humans using tightly coupled perception and action. Simple forms of interaction may be based on observing the manipulation of physical objects, or on visual sensing of fingers, hands or arms. Richer forms of interaction require perception and understanding of human activity and context. PRIMA has developed a novel theory for situation modeling for machine understanding of human activity, based on techniques used in Cognitive Psychology [44]. PRIMA explores multiple forms of interaction, including projected interaction widgets, observation of manipulation of objects, fusion of acoustic and visual information, and systems that model interaction context in order to predict appropriate action and services by the environment.

For the design and integration of systems for perception of humans and their actions, PRIMA has developed:

- A theoretical foundation for machine understanding of human activity using situation models.
- Robust, view invariant methods for computer vision systems using local appearance.
- A software architecture model for reactive control of multimodal perceptual systems.

The experiments in project PRIMA are oriented towards developing interactive services for smart environments. Application domains include health and activity monitoring services for assisted living, smart habitat for smart energy, context aware video recording for lectures, meetings and collaborative work, context aware services for commercial environments new forms of man-machine interaction based on perception and new forms of interactive services for education, research and entertainment. Creating interactive services requires scientific progress on a number of fundamental problems, including:

- Situation models for observing and understanding human to human interaction.
- Lifelong interactive learning.
- Robust, view-invariant image description for embedded services based on computer vision.
- New forms of multimodal human-computer interaction.
- Component-based software architectures for multimodal perception and action.
- Service-oriented software architectures for smart environments.

3. Research Program

3.1. Situation Models for Context Aware Systems and Services

Context Awareness, Smart Spaces

3.1.1. Summary

Over the last few years, the PRIMA group has pioneered the use of context aware observation of human activity in order to provide non-disruptive services. In particular, we have developed a conceptual framework for observing and modeling human activity, including human-to-human interaction, in terms of situations.

Encoding activity in situation models provides a formal representation for building systems that observe and understand human activity. Such models provide scripts of activities that tell a system what actions to expect from each individual and the appropriate behavior for the system. A situation model acts as a non-linear script for interpreting the current actions of humans, and predicting the corresponding appropriate and inappropriate actions for services. This framework organizes the observation of interaction using a hierarchy of concepts: scenario, situation, role, action and entity. Situations are organized into networks, with transition probabilities, so that possible next situations may be predicted from the current situation.

Current technology allows us to handcraft real-time systems for a specific services. The current hard challenge is to create a technology to automatically learn and adapt situation models with minimal or no disruption of human activity. An important current problem for the PRIMA group is the adaptation of Machine Learning techniques for learning situation models for describing the context of human activity.

3.1.2. Detailed Description

Context Aware Systems and Services require a model for how humans think and interact with each other and their environment. Relevant theories may be found in the field of cognitive science. Since the 1980's, Philippe Johnson-Laird and his colleagues have developed an extensive theoretical framework for human mental models [45], [46]. Johnson Laird's "situation models", provide a simple and elegant framework for predicting and explaining human abilities for spatial reasoning, game playing strategies, understanding spoken narration, understanding text and literature, social interaction and controlling behavior. While these theories are primarily used to provide models of human cognitive abilities, they are easily implemented in programmable systems [34], [33].

In Johnson-Laird's Situation Models, a situation is defined as a configuration of relations over entities. Relations are formalized as N-ary predicates such as beside or above. Entities are objects, actors, or phenomena that can be reliably observed by a perceptual system. Situation models provide a structure for organizing assemblies of entities and relations into a network of situations. For cognitive scientists, such models provide a tool to explain and predict the abilities and limitations of human perception. For machine perception systems, situation models provide the foundation for assimilation, prediction and control of perception. A situation model identifies the entities and relations that are relevant to a context, allowing the perception system to focus limited computing and sensing resources. The situation model can provide default information about the identities of entities and the configuration of relations, allowing a system to continue to operate when perception systems fail or become unreliable. The network of situations provides a mechanism to predict possible changes in entities or their relations. Finally, the situation model provides an interface between perception and human centered systems and services. On the one hand, changes in situations can provide events that drive service behavior. At the same time, the situation model can provide a default description of the environment that allows human-centered services to operate asynchronously from perceptual systems.

We have developed situation models based on the notion of a script. A theatrical script provides more than dialog for actors. A script establishes abstract characters that provide actors with a space of activity for expression of emotion. It establishes a scene within which directors can layout a stage and place characters. Situation models are based on the same principle.

A script describes an activity in terms of a scene occupied by a set of actors and props. Each actor plays a role, thus defining a set of actions, including dialog, movement and emotional expressions. An audience understands the theatrical play by recognizing the roles played by characters. In a similar manner, a user service uses the situation model to understand the actions of users. However, a theatrical script is organised as a linear sequence of scenes, while human activity involves alternatives. In our approach, the situation model is not a linear sequence, but a network of possible situations, modeled as a directed graph.

Situation models are defined using roles and relations. A role is an abstract agent or object that enables an action or activity. Entities are bound to roles based on an acceptance test. This acceptance test can be seen as a form of discriminative recognition.

There is no generic algorithm capable of robustly recognizing situations from perceptual events coming from sensors. Various approaches have been explored and evaluated. Their performance is very problem and environment dependent. In order to be able to use several approaches inside the same application, it is necessary to clearly separate the specification of scenario and the implementation of the program that recognizes it, using a Model Driven Engineering approach. The transformation between a specification and its implementation must be as automatic as possible. We have explored three implementation models :

- *Synchronized petri net*. The Petri Net structure implements the temporal constraints of the initial context model (Allen operators). The synchronisation controls the Petri Net evolution based on roles and relations perception. This approach has been used for the Context Aware Video Acquisition application.
- *Fuzzy Petri Nets*. The Fuzzy Petri Net naturally expresses the smooth changes of activity states (situations) from one state to another with gradual and continuous membership function. Each fuzzy situation recognition is interpreted as a new proof of the recognition of the corresponding context. Proofs are then combined using fuzzy integrals. This approach has been used to label videos with a set of predefined scenarios (context).
- *Hidden Markov Model*. This probabilistic implementation of the situation model integrates uncertainty values that can both refer to confidence values for events and to a less rigid representation of situations and situations transitions. This approach has been used to detect interaction groups and to determinate who is interacting with whom and thus which interaction groups are formed.

Currently situation models are constructed by hand. Our challenge is to provide a technology by which situation models may be adapted and extended by explicit and implicit interaction with the user. An important aspect of taking services to the real world is an ability to adapt and extend service behaviour to accommodate individual preferences and interaction styles. Our approach is to adapt and extend an explicit model of user activity. While such adaptation requires feedback from users, it must avoid or at least minimize disruption. We are currently exploring reinforcement learning approaches to solve this problem.

With a reinforcement learning approach, the system is rewarded and punished by user reactions to system behaviours. A simplified stereotypic interaction model assures a initial behaviour. This prototypical model is adapted to each particular user in a way that maximizes its satisfaction. To minimize distraction, we are using an indirect reinforcement learning approach, in which user actions and consequences are logged, and this log is periodically used for off-line reinforcement learning to adapt and refine the context model.

Adaptations to the context model can result in changes in system behaviour. If unexpected, such changes may be disturbing for the end users. To keep user's confidence, the learned system must be able to explain its actions. We are currently exploring methods that would allow a system to explain its model of interaction. Such explanation is made possible by explicit describing context using situation models.

The PRIMA group has refined its approach to context aware observation in the development of a process for real time production of a synchronized audio-visual stream based using multiple cameras, microphones and other information sources to observe meetings and lectures. This "context aware video acquisition system" is an automatic recording system that encompasses the roles of both the cameraman and the director. The system determines the target for each camera, and selects the most appropriate camera and microphone to record the current activity at each instant of time. Determining the most appropriate camera and microphone requires a model of activities of the actors, and an understanding of the video composition rules. The model of the activities of the actors is provided by a "situation model" as described above.

In collaboration with France Telecom, we have adapted this technology to observing social activity in domestic environments. Our goal is to demonstrate new forms of services for assisted living to provide non-intrusive access to care as well to enhance informal contact with friends and family.

3.2. Service Oriented Architectures for Intelligent Environments

Software Architecture, Service Oriented Computing, Service Composition, Service Factories, Semantic Description of Functionalities

Intelligent environments are at the confluence of multiple domains of expertise. Experimenting within intelligent environments requires combining techniques for robust, autonomous perception with methods for modeling and recognition of human activity within an inherently dynamic environment. Major software engineering and architecture challenges include accomodation of a heterogeneous of devices and software, and dynamically adapting to changes human activity as well as operating conditions.

The PRIMA project explores software architectures that allow systems to be adapt to individual user preferences. Interoperability and reuse of system components is fundamental for such systems. Adopting a shared, common Service Oriented Architecture (SOA) architecture has allowed specialists from a variety of subfields to work together to build novel forms of systems and services.

In a service oriented architecture, each hardware or software component is exposed to the others as a "service". A service exposes its functionality through a well defined interface that abstracts all the implementation details and that is usually available through the network.

The most commonly known example of a service oriented architecture are the Web Services technologies that are based on web standards such as HTTP and XML. Semantic Web Services proposes to use knowledge representation methods such as ontologies to give some semantic to services functionalities. Semantic description of services makes it possible to improve the interoperability between services designed by different persons or vendors.

Taken out of the box, most SOA implementations have some "defects" preventing their adoption. Web services, due to their name, are perceived as being only for the "web" and also as having a notable performance overhead. Other implementations such as various propositions around the Java virtual machine, often requires to use a particular programming language or are not distributed. Intelligent environments involves many specialist and a hard constraint on the programming language can be a real barrier to SOA adoption.

The PRIMA project has developed OMiSCID, a middleware for service oriented architectures that addresses the particular problematics of intelligent environments. OMiSCID has emerged as an effective tool for unifying access to functionalities provided from the lowest abstraction level components (camera image acquisition, image processing) to abstract services such as activity modeling and personal assistant. OMiSCID has facilitated cooperation by experts from within the PRIMA project as well as in projects with external partners.

3.3. Robust view-invariant Computer Vision

Local Appearance, Affine Invariance, Receptive Fields

3.3.1. Summary

A long-term grand challenge in computer vision has been to develop a descriptor for image information that can be reliably used for a wide variety of computer vision tasks. Such a descriptor must capture the information in an image in a manner that is robust to changes the relative position of the camera as well as the position, pattern and spectrum of illumination.

Members of PRIMA have a long history of innovation in this area, with important results in the area of multi-resolution pyramids, scale invariant image description, appearance based object recognition and receptive field histograms published over the last 20 years. The group has most recently developed a new approach that extends scale invariant feature points for the description of elongated objects using scale invariant ridges. PRIMA has worked with ST Microelectronics to embed its multi-resolution receptive field algorithms into low-cost mobile imaging devices for video communications and mobile computing applications.

3.3.2. Detailed Description

The visual appearance of a neighbourhood can be described by a local Taylor series [48]. The coefficients of this series constitute a feature vector that compactly represents the neighbourhood appearance for indexing and matching. The set of possible local image neighbourhoods that project to the same feature vector are referred to as the "Local Jet". A key problem in computing the local jet is determining the scale at which to evaluate the image derivatives.

Lindeberg [50] has described scale invariant features based on profiles of Gaussian derivatives across scales. In particular, the profile of the Laplacian, evaluated over a range of scales at an image point, provides a local description that is "equi-variant" to changes in scale. Equi-variance means that the feature vector translates exactly with scale and can thus be used to track, index, match and recognize structures in the presence of changes in scale.

A receptive field is a local function defined over a region of an image [56]. We employ a set of receptive fields based on derivatives of the Gaussian functions as a basis for describing the local appearance. These functions resemble the receptive fields observed in the visual cortex of mammals. These receptive fields are applied to color images in which we have separated the chrominance and luminance components. Such functions are easily normalized to an intrinsic scale using the maximum of the Laplacian [50], and normalized in orientation using direction of the first derivatives [56].

The local maxima in x and y and scale of the product of a Laplacian operator with the image at a fixed position provides a "Natural interest point" [52]. Such natural interest points are salient points that may be robustly detected and used for matching. A problem with this approach is that the computational cost of determining intrinsic scale at each image position can potentially make real-time implementation unfeasible.

A vector of scale and orientation normalized Gaussian derivatives provides a characteristic vector for matching and indexing. The oriented Gaussian derivatives can easily be synthesized using the "steerability property" [39] of Gaussian derivatives. The problem is to determine the appropriate orientation. In earlier work by PRIMA members Colin de Verdiere [31], Schiele [56] and Hall [43], proposed normalising the local jet independently at each pixel to the direction of the first derivatives calculated at the intrinsic scale. This results for many view invariant image recognition tasks are described in the next section.

Key results in this area include

- Fast, video rate, calculation of scale and orientation for image description with normalized chromatic receptive fields [34].
- Robust visual features for face tracking [41], [40].
- Direct computation of time to collision over the entire visual field using rate of change of intrinsic scale [54].

We have achieved video rate calculation of scale and orientation normalized Gaussian receptive fields using an O(N) pyramid algorithm [34]. This algorithm has been used to propose an embedded system that provides real time detection and recognition of faces and objects in mobile computing devices.

Applications have been demonstrated for detection, tracking and recognition of faces as well detection of emotions and posture at video rates.

3.4. Perception for Social Interaction

Affective Computing, Perception for social interaction.

Current research on perception for interaction primarily focuses on recognition and communication of linguistic signals. However, most human-to-human interaction is non-verbal and highly dependent on social context. A technology for natural interaction requires abilities to perceive and assimilate non-verbal social signals, to understand and predict social situations, and to acquire and develop social interaction skills.

The overall goal of this research program is to provide the scientific and technological foundations for systems that observe and interact with people in a polite, socially appropriate manner. We address these objectives with research activities in three interrelated areas:

- Multimodal perception for social interactions.
- Learning models for context aware social interaction, and
- Context aware systems and services.

Our approach to each of these areas is to draw on models and theories from the cognitive and social sciences, human factors, and software architectures to develop new theories and models for computer vision and multi-modal interaction. Results will be developed, demonstrated and evaluated through the construction of systems and services for polite, socially aware interaction in the context of smart habitats.

3.4.1. Detailed Description

First part of our work on perception for social interaction has concentrated on measuring the physiological parameters of Valence, Arousal and Dominance using visual observation form environmental sensors as well as observation of facial expressions.

People express and feel emotions with their face. Because the face is both externally visible and the seat of emotional expression, facial expression of emotion plays a central role in social interaction between humans. Thus visual recognition of emotions from facial expressions is a core enabling technology for any effort to adapt systems for social interaction.

Constructing a technology for automatic visual recognition of emotions requires solutions to a number of hard challenges. Emotions are expressed by coordinated temporal activations of 21 different facial muscles assisted by a number of additional muscles. Activations of these muscles are visible through subtle deformations in the surface structure of the face. Unfortunately, this facial structure can be masked by facial markings, makeup, facial hair, glasses and other obstructions. The exact facial geometry, as well as the coordinated expression of muscles is unique to each individual. In additions, these deformations must be observed and measured under a large variety of illumination conditions as well as a variety of observation angles. Thus the visual recognition of emotions from facial expression remains a challenging open problem in computer vision.

Despite the difficulty of this challenge, important progress has been made in the area of automatic recognition of emotions from face expressions. The systematic cataloging of facial muscle groups as facial action units by Ekman [38] has let a number of research groups to develop libraries of techniques for recognizing the elements of the FACS coding system [30]. Unfortunately, experiments with that system have revealed that the system is very sensitive to both illumination and viewing conditions, as well as the difficulty in interpreting the resulting activation levels as emotions. In particular, this approach requires a high-resolution image with a high signal-to-noise ratio obtained under strong ambient illumination. Such restrictions are not compatible with the mobile imaging system used on tablet computers and mobile phones that are the target of this effort.

As an alternative to detecting activation of facial action units by tracking individual face muscles, we propose to measure physiological parameters that underlie emotions with a global approach. Most human emotions can be expressed as trajectories in a three dimensional space whose features are the physiological parameters of Pleasure-Displeasure, Arousal-Passivity and Dominance-Submission. These three physiological parameters can be measured in a variety of manners including on-body accelerometers, prosody, heart-rate, head movement and global face expression.

In our work, we address the recognition of social behaviours multimodal information. These are unconscious inmate cognitive processes that are vital to human communication and interaction. Recognition of social behaviours enables anticipation and improves the quality of interaction between humans. Among social behaviours, we have focused on engagement, the expression of intention for interaction. During the engagement phase, many non-verbal signals are used to communicate the intention to engage to the partner [58]. These include posture, gaze, spatial information, gestures, and vocal cues.

For example, within the context of frail or elderly people at home, a companion robot must also be able to detect the engagement of humans in order to adapt their responses during interaction with humans to increase their acceptability. Classical approaches for engagement with robots use spatial information such as human position and speed, human-robot distance and the angle of arrival. Our believe is that uni-modal methods may be suitable for static display [59] and robots in wide space area [49] but not for home environments. In an apartment, relative spatial information of people and robot are not as discriminative as in an open space. Passing by the robot in a corridor should not lead to an engagement detection, and possible socially inappropriate behaviour by the robot.

In our experiments, we used a kompai robot from Robosoft [29]. As an alternative to wearable physiological sensors (such as pulse bracelet Cardiocam, etc.) we integrate multimodal features using a Kinect sensor (see figure 1). In addition of the spatial cues from the laser telemeter, one can use new multimodal features based on persons and skeletons tracking, sound localization, etc. Some of these new features are inspired from results in cognitive science domain [55].

Figure 1. On the left image, one can see the telemeter range in red, the foot detection (blue spot) and the angle view from the Kinect (in green). the middle and right image represent RGB camera in depth view from the Kinect.

Our multimodal approach has been confronted to a robot centered dataset for multimodal social signal processing recorded in a home-like environment [36]. The evaluation on our corpus highlights its robustness and validates use of such technique in real environment. Experimental validation shows that the use of multimodal sensors gives better results than only spatial features (50% of error reduction). Our experimentations also confirm results from [55]: relative shoulder rotation, speed and facing visage are among crucial features for engagement detection.

3.5. End User control of Smart Environments

End users programming, smart home, smart environment

Pervasive computing promises unprecedented empowerment from the flexible and robust combination of software services with the physical world. Software researchers assimilate this promise as system autonomy where users are conveniently kept out of the loop. Their hypothesis is that services, such as music playback and calendars, are developed by service providers and pre-assembled by software designers to form new service frontends. Their scientific challenge is then to develop secure, multiscale, multi-layered, virtualized infrastructures that guarantee service front-end continuity. Although service continuity is desirable in many circumstances, end users, with this interpretation of ubiquitous computing, are doomed to behave as mere consumers, just like with conventional desktop computing.

Another interpretation of the promises of ubiquitous computing, is the empowerment of end users with tools that allow them to create and reshape their own interactive spaces. Our hypothesis is that end users are willing to shape their own interactive spaces by coupling smart artifacts, building imaginative new functionality that were not anticipated by system designers. A number of tools and techniques have been developed to support this view such as CAMP [57] or iCAP [37].

We are investigating an End-User Programming (EUP) approach to give the control back to the inhabitants. In our vision, smart homes will be incrementally equiped with sensors, actuators and services by inhabitants themselves. Our research programm therefore focus on tools and languages to enable inhabitants in activities related to EUP for Smart Homes :

- Installation and maintenance of devices and services. This may imply having facilities to attribute names.
- Visualizing and controling of the Smart Habitat.
- Programming and testing. This imply one or more programming languages and programming environment which could rely on the previous point. The programming language is especially important. Indeed, in the context of the Smart Homes, End-User Programms are most likely to be routines in the sens of [35] than procedure in the sense of traditionnal programming languages.
- Detecting and solving conflicts related to contradictory programs or goals.

4. New Software and Platforms

4.1. OMiSCID Middleware for Distributed Multimodal Perception

Participants: Amaury Negre, Patrick Reignier, Dominique Vaufreydaz [correspondant].

Middleware, Distributed perceptual systems

OMiSCID is lightweight middleware for dynamic integration of perceptual services in interactive environments. This middleware abstracts network communications and provides service introspection and discovery using DNS-SD (DNS-based Service Discovery). Services can declare simplex or duplex communication channels and variables. The middleware supports the low-latency, high-bandwidth communications required in interactive perceptual applications. It is designed to allow independently developed perceptual components to be integrated to construct user services. Thus our system has been designed to be cross-language, cross-platform, and easy to learn. It provides low latency communications suitable for audio and visual perception for interactive services.

OMiSCID has been designed to be easy to learn in order to stimulate software reuse in research teams and is revealing to have a high adoption rate. To maximize this adoption and have it usable in projects involving external partners, the OMiSCID middleware has been released under an open source licence. To maximize its target audience, OMiSCID is available from a wide variety of programming languages: C++, Java, Python and Matlab. A website containing information and documentations about OMiSCID has been set up to improve the visibility and promote the use of this middleware.

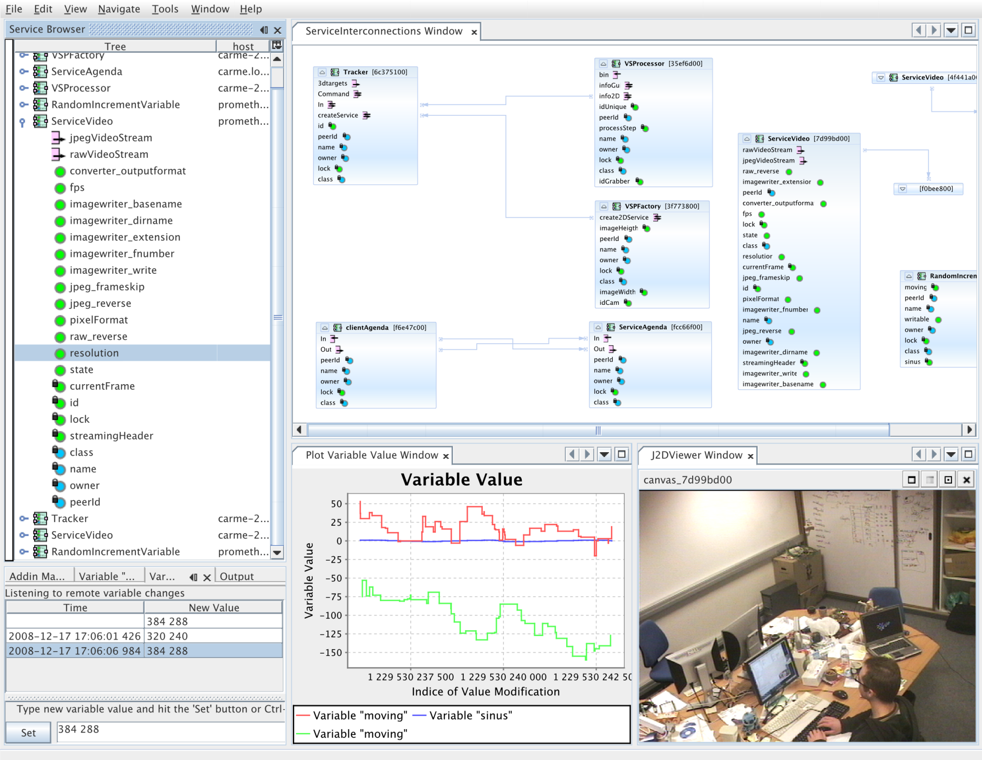

Figure 2. OMiSCID GUI showing a list of running services and some modules for service interconnections, variable plotting, live video stream display and variable control

4.2. Pal Middleware

Participants: Amaury Negre, Dominique Vaufreydaz [correspondant].

Middleware, Distributed perceptual systems, Robotic Operating System (ROS), IPL PAL

A part of our efforts in the PAL project has been put toward developing a solution that would ease the integration of our multi-partners' software components.

The design of PAL Middleware responds to a requirement that within the PAL project, each partner is responsible for maintaining 1) its software heritage 2) its resources 3) its competences and fields of research and expertise; 4) current practices in terms of programming language, (c/c++, Java, Python), computing platforms (OSx, Linux, Windows, Android, etc.) and interconnect software components (OSGi, OMiSCID, MPI, PVM, etc.); and 5) its particular needs and constraints.

For it to be widely accepted, the PAL middleware must be designed to be ecologic and pragmatic. Ecologic in the sense that the solution does not perturb the ecology of each ecosystem, pragmatic in the sense that setting up this solution did not require an heavy development effort, also because PAL and is required to reuse existing software solutions.

For developing PALGate we introduced a novel concept: software gate. Unlike software components/services which can be instantiated, a software gate is only a concept, it is defined as an ecologic and hermetic interface between different ecosystems. A software gate is characterized by the subset of functionalities it exposes to other gates, where the functionalities it exposes are provided by the software components/services of its belonging ecosystem. A software gate is hermetic in the sense that only a selected subset of functionalities of an ecosystem are exposed but also because it propagates only filtered information exposed by other gates into its ecosystem. The last characteristic of a software gate is that it makes explicit to other gates the communication mechanisms it uses.

While a software gate is only conceptual, the PAL middleware is an implementation of a gate oriented middleware. The PAL Middleware uses ROS to support the basic communication between gates. Within PALGate, each ecosystem is associated to only one software gate. Practically, PAL middleware 1) is a ROS stack containing gates definition 2) is a set of conventions (e.g. stack organization, package/node/topic/service names, namespaces, etc.) 3) it provides dedicated tools to ease the integration and its usage by partners. A software gate in PAL is a ROS package containing definition of ROS types (i.e. msgs and srvs types), but also exposed ROS communication channels (i.e. topics and RPCs).

With this architecture each partner has to provide the PAL middleware with a package containing the definition of its gate. Then in order a) to expose functionalities out of their ecosystem and b) to propagate information into their ecosystem, each partner must create ROS nodes. These ROS nodes let each partner interface their ecosystem through ROS topics and ROS services without having to change anything about their architecture. For instance if a partner is using Java and OSGi, it can create nodes in ROS Java that will expose/register functionalities through ROS services, publish/subscribe information using ROS topics.

4.3. EmoPRAMAD

Participant: Dominique Vaufreydaz [correspondant].

Affective computing,

Within the Pramad project, we want to offer a full affective loop between the companion robot and the elderly people at home. This affective loop is necessary within the context of everyday interaction of elderly and the companion robot. A part of this loop is to make the robot express emotions in response to the emotional state of the user. To do that, we need to test our working hypothesis about the visual representation of emotions with the 3D face of robot. EmoPRAMAD is an evaluation tool designed to conduct comparative studies between human faces and the 3D faces expressing a defined set of emotions.

The evaluation conducted though EmoPRAMAD concerns both unimodal (facial only) and bimodal conditions (facial/sound). The emotions set is composed of 4 basic emotions (joy, fear, anger, sadness) and a neutral state. While experimenting, the software collects several parameters in order to evaluate more than correctness of the answers: time to respond, length of mouse moves, etc. Experimentation is still in progress at Inria in Grenoble, University Pierre and Marie Currie and Broca Hospital in Paris. A set of 235 participants from 14 to 88 years old was already recorded.

4.4. Detection and Tracking of Pedestrians in INRETS Intelligent Urban Spaces Platform

Participants: Claudine Combe, James Crowley [correspondant], Lukas Rummelhard.

Visual detection and tracking of pedestrians, Intelligent Urban Space

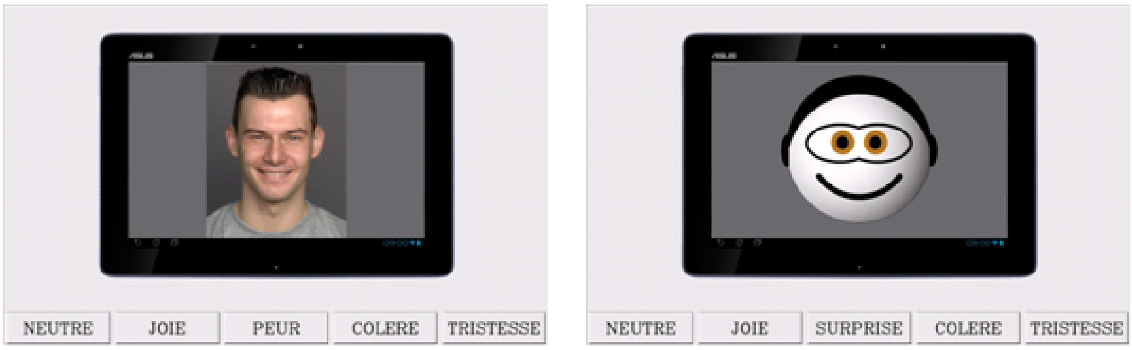

Figure 3. EmoPRAMAD interfaces with a human face and a 3D face from our virtual agent.

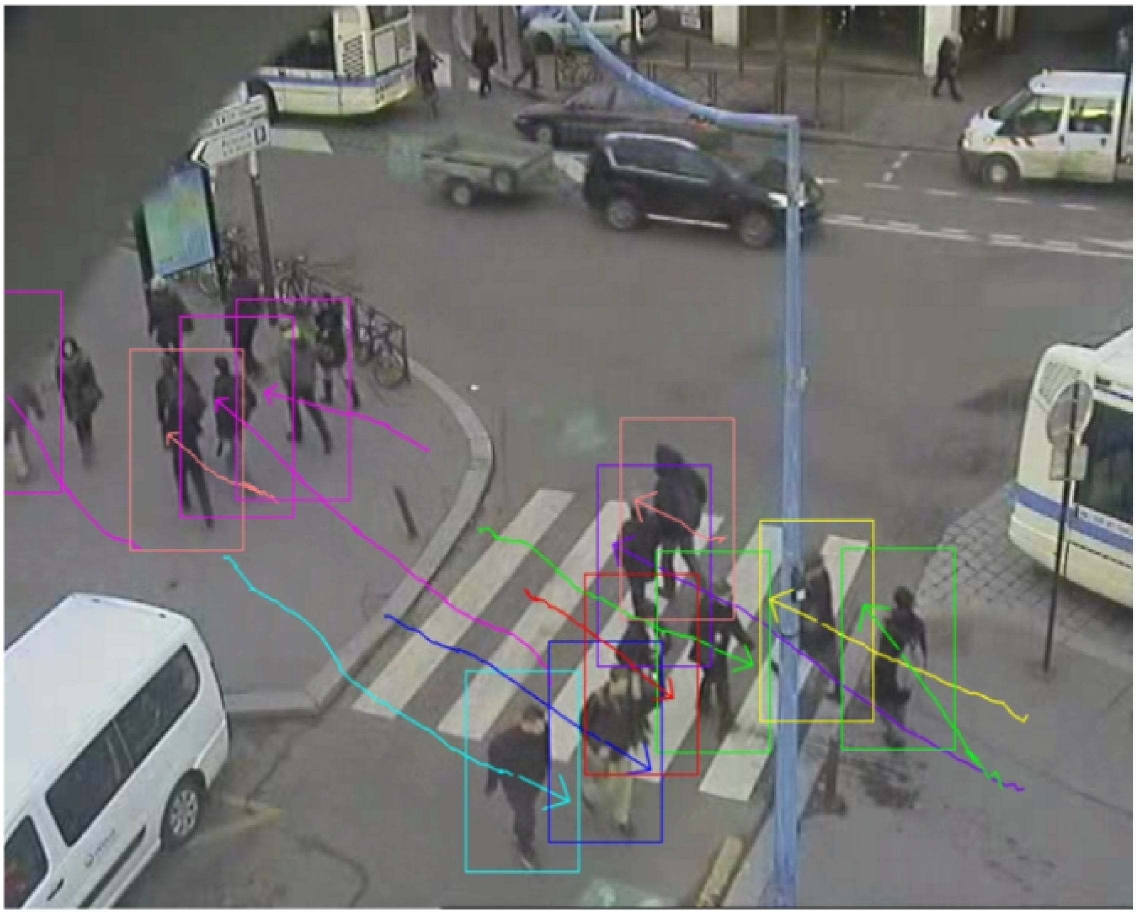

Figure 4. Cipebus: pedestrian tracking system.

The project ANR-07-TSFA-009-01 CIPEBUS ("Carrefour Intelligent - Pole d'Echange - Bus) has been proposed by INRETS-IFSTTAR, in collaboration with Inria, Citilog, Fareco, and the city of Versaille. The Objective of the CIPEBUS project is to develop an experimental platform for observing activity in a network of urban streets in order to experiment with techniques for optimizing circulation by context aware control of traffic lights.

Within CipeBus, Inria has developed a real time multi-camera computer vision system to detect and track people using a network of surveillance cameras. The CipeBus combines real time pedestrian detection with 2D and 3D Bayesian tracking to record the current position and trajectory of pedestrians in an urban environment under natural view conditions. The system extends the sliding window approach to use a half-octave Gaussian Pyramid to explore hypotheses of pedestrians at different positions and scales. A cascade classifier is used to determine the probability that a pedestrian can be found at a particular position and scale. Detected pedestrians are then tracked using a particle filter.

The resulting software system has been installed and tested at the INRETS CipeBus platform and is currently used for experiments in controlling the traffic lights to optimize the flow of pedertrians and public transportation while minimizing the delay imposed on private automobiles.

4.5. Multisensor observation of human activity for integrated energy and comfort management

Participants: Claudine Combe, James Crowley [correspondant], Lucas Nacsa, Amaury Negre, Lukas Rummelhard.

multimodal tracking of human activity

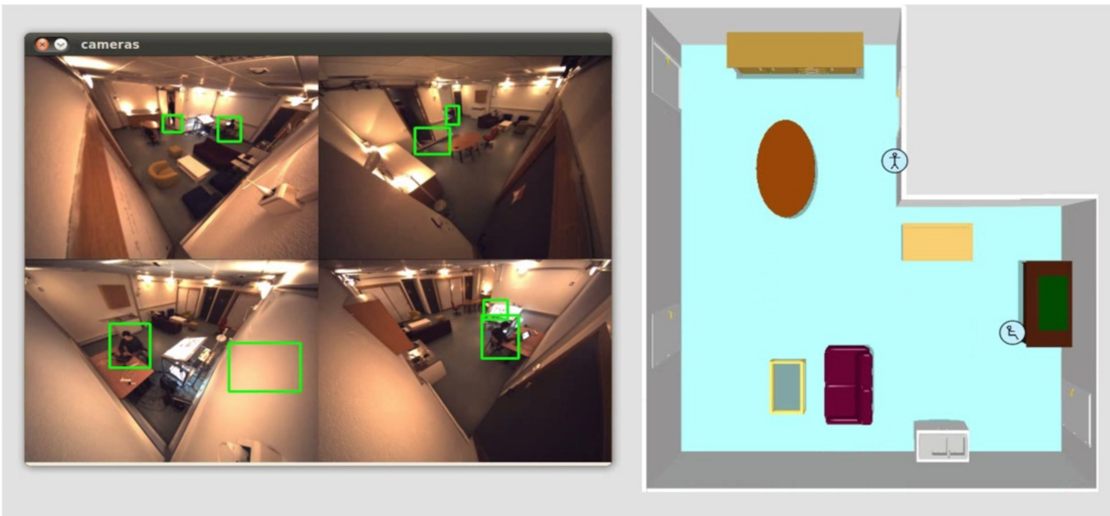

Figure 5. The 3D tracker integrates observations from multiple sensors

As part of Inria's contribution of ICTLabs Action TSES - Smart Energy Systems, we have constructed a system that integrates information from multiple environmental sensor to detect and track people in indoor environments. This system, constructed as part of activity 11831 Open SES Experience Labs for Prosumers and New Services, has been released to ICTLabs partners in June 2012. It has also been used for construction of a smart spaces testbed at Schneider Electric.

This software, named MultiSensor activity tracker, integrates information from multiple environmental sensors to keep track of the location and activity of people in a smart environment. This model is designed to be used by a home energy broker that would work in conjunction with a smart grid to manage the energy consumption of home appliances, balancing the needs of inhabitants with opportunities for savings offered by electricity rates. This database will also be used for by advisor services that will offer advice to inhabitants on the consequences to energy consumption and energy cost that could potentially result from changes to lifestyle or home energy use.

Work in this task draws from earlier result from a number of development projects at Inria. In the ANR Casper project Inria created Bayesian tracking system for human activity using a voxel based occupancy grid. Within the INRA ADT PAL project, Inria is creating methods for plug and play installation of visual and acoustic sensors for tracking human activity within indoor environments.

While a voxel based Bayesian tracker has served well for a number of applications, a number of limitations have been observed. For example, under certain circumstances, the sensor data can provide contradictory or ambiguous data about the location and activities of people. Resolving such cases required the Bayesian tracker to choose between a numbers of competing hypotheses, potentially resulting in errors. Several members of our group have argued that an alternative integration approach based on the use of a Particle filter would solve these problems and provide a more reliable tracking system. This task has been undertaken to evaluate this hypothesis. The system configured and optimized for detecting and tracking people within rooms using multiple calibrated cameras. The system currently uses corner mounted cartesian cameras, ceiling mounted cameras with wide angle lenses and panoramic cameras placed on tables. Cameras may be connected and disconnected while the component is running, but they must be pre-calibrated to a common room reference frame. We are currently experimenting with techniques for Bayesian estimation of camera parameters for auto-calibration. Cameras may be connected dynamically.

The original system 3DBT has been declared with the APP "Agence pour la Protection des Programmes" under the Interdeposit Digital number IDDN.FR.001.490023.000.S.P.2006.000.10000. A revised declaration for the latest version of the system is currently being prepared.

4.6. Stereo Viewfinder

Participants: Frédéric Devernay [correspondant], Loic Lefort, Elise Mansilla, Sergi Pujades-Rocamora.

Stereoscopy, Auto-calibration, Real-time video processing, Feature matching

This software has been filed with the APP "Agence pour la Protection des Programmes" under the Interdeposit Digital number IDDN.FR.001.370083.000.S.P.2007.000.10000

4.7. Visual Emotion Recognition for Health and Well Being.

Participants: James Crowley [correspondant], Varun Jain, Sergi Pujades-Rocamora.

Visual Emotion Recognition

People express and feel emotions with their face. Because the face is the both externally visible and the seat of emotional expression, facial expression of emotion plays a central role in social interaction between humans. Thus visual recognition of emotions from facial expressions is a core enabling technology for any effort to adapt ICT to improve Health and Wellbeing.

Constructing a technology for automatic visual recognition of emotions requires solutions to a number of hard challenges. Emotions are expressed by coordinated temporal activations of 21 different facial muscles assisted by a number of additional muscles. Activations of these muscles are visible through subtle deformations in the surface structure of the face. Unfortunately, this facial structure can be masked by facial markings, makeup, facial hair, glasses and other obstructions. The exact facial geometry, as well as the coordinated expression of muscles is unique to each individual. In additions, these deformations must be observed and measured under a large variety of illumination conditions as well as a variety of observation angles. Thus the visual recognition of emotions from facial expression remains a challenging open problem in computer vision.

Despite the difficulty of this challenge, important progress has been made in the area of automatic recognition of emotions from face expressions. The systematic cataloging of facial muscle groups as facial action units by Ekman [38] has let a number of research groups to develop libraries of techniques for recognizing the elements of the FACS coding system [30]. Unfortunately, experiments with that system have revealed that the system is very sensitive to both illumination and viewing conditions, as well as the difficulty in interpreting the resulting activation levels as emotions. In particular, this approach requires a high-resolution image with a high signal-to-noise ratio obtained under strong ambient illumination. Such restrictions are not compatible with the mobile imaging system used on tablet computers and mobile phones that are the target of this effort.

As an alternative to detecting activation of facial action units by tracking individual face muscles, we propose to measure physiological parameters that underlie emotions with a global approach. Most human emotions can be expressed as trajectories in a three dimensional space whose features are the physiological parameters of Pleasure-Displeasure, Arousal-Passivity and Dominance-Submission. These three physiological parameters can be measured in a variety of manners including on-body accelerometers, prosody, heart-rate, head movement and global face expression.

The PRIMA Group at Inria has developed robust fast algorithms for detection and recognition of human faces suitable for use in embedded visual systems for mobile devices and telephones. The objective of the work described in this report is to employ these techniques to construct a software system for measuring the physiological parameters commonly associated with emotions that can be embedded in mobile computing devices such as cell phones and tablets.

A revised software package has recently been released to our ICTlab partners for face detection, face tracking, gender and age estimation, and orientation estimation, as part of ICTlabs Smart Spaces action line. This software has been declared with the APP "Agence pour la Protection des Programmes" under the Interdeposit Digital number IDDN.FR.001.370003.000.S.P.2007.000.21000.

A software library, named PrimaCV has been designed, debugged and tested, and released to ICTLabs partners for real time image acquisition, robust invariant multi-scale image description, highly optimized face detection, and face tracking. This software has been substantially modified so as to run on an mobile computing device using the Tegra 3 GPU.

4.8. AppsGate - Smart Home Application Gateway

Participants: Alexandre Demeure, James Crowley [correspondant], Emeric Grange, Cedric Gerard, Camille Lenoir, Kouzma Petoukhov.

Smart Home Applications Gateway

PRIMA has participated in the development of the AppsGate Home Application Gateway Architecture. The AppsGate architecture is based on the HMI Middleware developed in cooperation with the IIHM and Adele groups of the UMR Laboratoire Informatique de Grenoble (LIG). The HMI Middleware is designed to facilitate the development of end-user applications on top of the core software components described in the sections above, while ensuring service continuity and usability. The key features of the HMI Middleware include:

- Integration of sensors and actuators managed by a variety of protocols, and provision of a uniform abstraction for these devices as component-oriented-services,
- Integration of Web services made available on the cloud by a variety of web service providers, and provision of a uniform abstraction for these services as component-oriented-services,
- Communication between the HMI middleware and client applications - typically, user interfaces for controlling and programming the smart home, that run on high-end devices such as smartphones, tablets, and TVs.

As part of the Appsgate middleware, we have developed SPOK, an End User Development Environment, that enables inhabitants to control and program their smart Homes via a web interface. The current version of SPOK includes an editor for editing programs using a pseudo-natural language and an interpreter. A multi-syntax editor as well as additional services such as a debugger and a simulator are currently under development.

4.9. a SmartEnergy Serious Game

Participant: Patrick Reignier.

This ongoing serious game is the result of a collaboration with Ayesha Kashif (LIG), Stephane Ploix (G-Scop) and Julie Dugdale (LIG). It has been developped as part of the Grenoble INP SmartEnergy project.

Inhabitants play a key role in buildings global energy consumption but it is difficult to involve them in energy management. Our objective is to make energy consumption visible by simulating inside a serious game the energy impact of inhabitants behaviours. A serious game is curently under development, coupling a 3D virtual environment and a building energy simulator. The 3D virtual environment is based on the JMonkey 3D engine. New houses can be easily imported using SweetHome 3D and Blender. The building energy simulator is EnergyPlus. The 3D engine and the energy engine are coupled using the Functional Mock-up Interface (FMI) standard. Using this standard will allow to easily switch between existing building energy simulators

5. New Results

5.1. Highlights of the Year

On March 14, 2014, James Crowley was named Chevalier de l'Ordre national du Mérite.

On August 2014, the paper "Human-Robot Motion: an Attention-Based Navigation Approach" [14] by Thierry Fraichard, Remi Paulin & Patrick Reignier has been nominated for the best paper award at the IEEE Int. Symp. on Robot and Human Interactive Communication (RO-MAN 2014), Edinburgh (UK).

On December 2014, Patrick Reignier was a member of the EDF grand jury for smart energy
BEST PAPER AWARD :
[14] **Human-Robot Motion: An Attention-Based Navigation Approach in IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN)**. T. FRAICHARD, R. PAULIN, P. REIGNIER.

5.2. Attention-Based Navigation

Participants: Thierry Fraichard, Remi Paulin, Patrick Reignier..

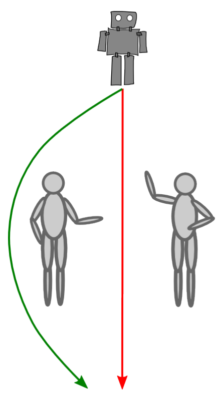

Figure 6. People are not pieces of furniture! Motion in red is definitely shorter but it is not appropriate.

The domain of service-robots is growing fast and has become the focus of many researchers and industrials alike. Their application areas been extremely broad, from logistics to handicap assistance. A large proportion of such robots are expected to share humans' living space and thus must be endowed with navigation capabilities that exceed the standard requirements pertaining to autonomous navigation such as motion safety. In a human populated environment, optimality does not boil down to minimising resources such as time or distance travelled anymore, the robot motion must abide by social/cultural rules and be **appropriate**, *e.g.* Fig. 6.

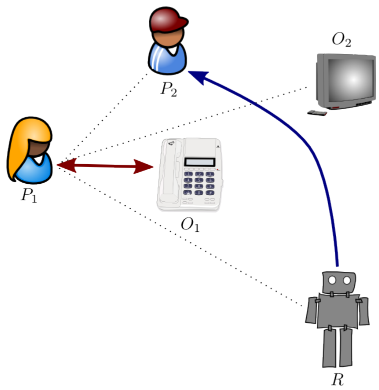

Figure 7. Attention vs activity: although P_1's current activity is being on the phone, part of her attention may be directed towards P_2, the TV set or the robot R. Suppose now that R moves towards P_2 in a way that hides the TV from P_1. Such a behavior would not be appropriate should P_1 be actually paying attention to the TV.

Most of the approaches proposed so far relies upon the definition of so-called **social spaces**, *i.e.* regions in the environment that, for different reasons, the persons consider as psychologically theirs. Such social spaces are primarily characterized using either the position of the person, *e.g.* "Personal space" [42], or the activity it is currently engaged in, *e.g.* "Interaction Space" [47] and "Activity Space" [51]. The most common approach is then to define costmaps on such social spaces: the higher the cost, the less desirable it is for the robot to be at the corresponding position. The costmaps are ultimately used for motion planning and navigation purposes. Such approaches are interesting however their spatial nature (being inside or outside the space) make them less suitable when facing more complicated situations, *e.g.* Fig. 7. To overcome those limits, we suggest using the psychological concept of **attention**, which plays a central role when humans navigate around each other. Besides lifting the limits of social spaces, this concept brings a new degree of control over the motion of the robot, namely the invasive and distracting character of the robot motion, which have so far proven hard to tackle with the conventional tools such as social spaces. Beside leading appropriate motion, attention-based navigation enable interaction through motion by predicting the quantity of attention the human will give to the robot.

Building upon a computational model of attention that was earlier proposed in [53]. we have developed the novel concept of **attention field**. The attention field is straightforward to define: it is a measure of the amount of attention that a given person would allocate to the robot, should the robot be in a given position/state. It is mapping from the state space of the robot to $\mathbb{R}$. The attention field can serve as an attention predictor that can be used to predict potential attentional situations. This knowledge can in turn be used to decide what the robot should do in the future depending on its current task.

Let us illustrate this on a simple scenario featuring a person, a TV and a robot (denoted by P_1, O_1 and R in Fig. 8-left). The person is currently watching the TV: this is his current *activity*. This activity relates to his *intention* and is modeled by the yellow vector $\overrightarrow{I}$ in Fig. 8-left that is directed from the person to the TV.

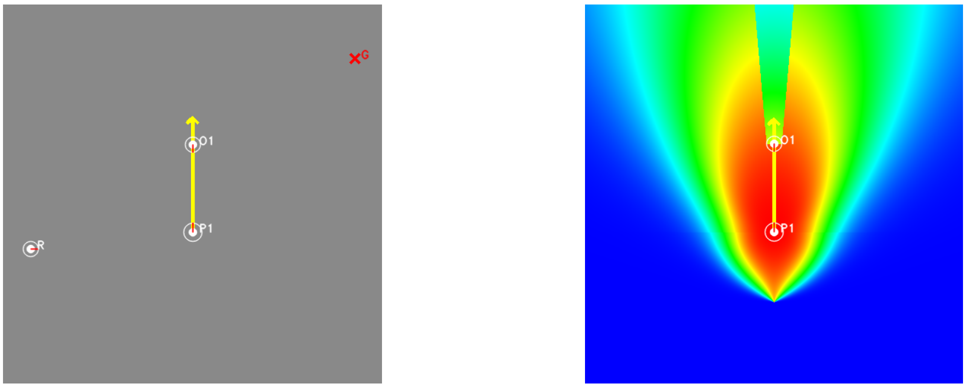

Figure 8. Person-TV-Robot scenario (left), Attention field for the person P_1 (right).

Fig. 8-right depicts the attention field for the person; it is a mapping from $\mathbb{R}^2$ to $\mathbb{R}$ that gives the amount of attention that the person is paying to the robot when it is at a given position (x, y). Fig. 8-right should be interpreted as follows: the warmer the color, the higher the amount of attention given by the person to the robot. It integrates both the visual and auditory perception capabilities of P_1.

In 2014, we have furthered the development of the concept of attention field and demonstrated different ways to use its attention prediction capability on various scenarios. The main results obtained have been reported in a conference article that has been nominated for the Best Paper Award [14]. Work is ongoing to quantify the social "goodness" of the paths provided by our approach, to further the use of the concept of attention on more challenging and dynamic scenarios and to offer an approach to fill the gap between appropriate motion and interaction through motion.

5.3. SPOK: End User Programming for Smart Homes

Participant: Alexandre Demeure.

As part of the CATRENE project AppsGate, we have developed SPOK, an End User Development Environment, that enables inhabitants to control and program their smart Homes via a web interface. The current version of SPOK includes an editor for editing programs using a pseudo-natural language and an interpreter. A multi-syntax editor as well as additional services such as a debugger and a simulator are expected for the second version.

A multi-syntax editor will allow users to build syntactically correct programs using the syntax that is most appropriate to them or by using a combination of them. These syntaxes include pseudo-natural language (i.e. a constrained natural language) and graphical iconic syntax (as exemplified by Scratch [Maloney et al. 2010]). The interaction techniques used to enter programs may be menu-based, free typing, as well as by demonstration in the physical home or by the way of the simulator. The simulator is the dual digital representation of the real home. It is intended to serve also as a debugger for testing and correcting end-user programs.

Whatever syntax used by end-users, programs are translated into syntactic abstract trees whose leaves reference services provided by the Core HMI and/or by the Extended HMI Middleware. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input.

In order to support a dynamically extensible grammar as well as to provide end-users with feedforward at the user interface of the editor, the grammar used by the editor is split into 2 parts: the root grammar and the device specific grammars. The root grammar specifies the generic structures of an end-user program: loops, conditions, etc. The device specific grammars are separated from the root grammar to be able to dynamically build the final grammar to be compliant with what is currently installed and detected by the AppsGate server. Each device type brings with it its own events, status and actions. These grammatical elements are injected into the root grammar when generating the parser and for compiling end-user programs.

The language used by end-users to express their programs is a pseudo-natural language using the rule-based programming paradigm. The left hand side of a rule is composed of events and conditions, and the right hand side specifies the actions to be taken when the left hand-side is true or becomes true. A program may include several rules that can be executed either in parallel or sequentially. Once entered, programs are translated into syntactic abstract trees. The interpreter, executes end-user programs, using the corresponding syntactic abstract trees as input. SPOK is implemented as a mix of OSGi and ApAM components where ApAM is in turn a middleware that runs on top of OSGi.

5.4. Qualitative approaches for building energy management

Participant: Patrick Reignier.

Reducing housing energy costs is a major challenge of the 21st century. In the near future, the main issue for building construction is the thermal insulation, but in the longer term, the issues are those of renewable energy (solar, wind, etc.) and smart buildings. Home automation system basically consists of household appliances linked via a communication network allowing interactions for control purposes. Thanks to this network, a load management mechanism can be carried out: it is called distributed control. An optimal home energy management system is still a goal to aim for, because lots of aspects are still not completely fulfilled. Most of the energy systems respect only the energy needs, but they don’t tackle the user needs or satisfaction. Energy systems also have a lack when it comes to the dynamicity of the environments (the system ability to adapt).The problem is similar for the existing HMI (Human User Interface) of those Home Automation Systems where only experts can understand the data coming from the sensors and most important, the energy plan coming from management system (How? and Why?). The goal of this study is to propose a house energy model that can be both used to predict at some level energy evolution and that can be understood by the end user. The house energy model is based on Fuzzy Cognitive Maps representing cause-effects relations. It is first designed by an expert and then automatically tuned to a particular house using machine learning approaches. Preliminary experiments have been done this year using the Predis datasets.

6. Bilateral Contracts and Grants with Industry

6.1. Infrared Visual Sensors

PRIMA has worked with Schneider Electric on embedded image analysis algorithms for a new generation of far-infrared visual sensors. The objective is to develop an integrated visual sensor with very low power consumption. Such systems can be used to estimate temperature in different parts of a room, as well as to provide information about human presence and human activity.

7. Partnerships and Cooperations

7.1. European Initiatives

7.1.1. CATRENE Project AppsGate - Smart Home Application Gateway

Duration: june 2012 - June 2015

Coordinator: ST Microelectronics

Other partners: Pace, Technicolor, NXP, Myriad France SAS, 4MOD Technology, HI-IBERIA Ingenieria y Proyectos, ADD Semiconductor, Video Stream Network, SoftKinetic, Optrima, Fraunhofer, Vsonix, Evalan, University UJF/LIG, and Institut Telecom.

The Prima Project team has worked with 15 other partners to develop a new generation of set-top box for smart home applications. In close collaboration with ST Microelectronics and Immotronics, Prima has developed the core middleware components for plug and play integration of smart home devices for distributed smart home services, as well as interactive tools for End User Development of Smart Home services.

AppsGate has developed an Open Platform to provide integrated home applications to the consumer mass market. The set-top box is the primary point of entry into the digital home for television services including cable TV, satellite TV, and IPTV. AppsGate will transform the set-box into a residential gateway, capable of delivering multiple services to the home, including video, voice and data. PRIMA is involved in designing End User Development tools dedicated for the Smart Home

7.1.2. ICTLabs Smart Energy Systems Activity 11831 Making Energy Visible

Participants: Sabine Coquillart, James Crowley [correspondant], Patrick Reignier, Mayeul de Werbier d Antigneul.

Smart energy Systems, Smart Grids

Within Activity 11831 Open SES Experience Labs, PRIMA is responsible for the A1405 "Making Energy Visible" within the Smart Energy Systems action line of ICTlabs. The objective of this task is to design, implement and evaluate tools for online and offline interaction with energy usage. The group works with Immotronics to define an open middleware that will enable logging, aggregation and interactive visualization of data and information on energy consumption and on environmental parameters that define comfort. The open middleware will include an SQL Data base, web socket and an xml markup language to define a common naming scheme, tools for assigning location in both space (geometry coordinates) and place (functional places), as well as data aggregation tools.

The open middleware will serve as a common software platform that will be used for the Inria Rapid Deployment mini-kit as well as for data acquisition by other partners using other sensors. Univ Bologna will provide (sell) copies of their new energy measurement meter for integration into the system. Univ of Bologna, Fraunhofer, Fortis and Inria will construct tools for offline and online visualization. The system will be deployed and evaluated by social scientists at the living lab of Politechnico Turin. Turin will specify requirements for visualisation of energy and comfort for smart grid applications, and perform user studies and evaluations on the resulting systems.

Visualisation includes on-line and offline exploration, as well as tools for html web interface, Mobile augmented reality tools, apps for display on mobile devices, 3D visual interaction, and immersive interaction with an oculus Rift.

7.2. National Initiatives

7.2.1. EquipEx AmiQual4Home - Ambient Intelligence for Quality of Life

Participants: Stan Borkowski, Sabine Coquillart, Joelle Coutaz, James Crowley [correspondant], Alexandre Demeure, Thierry Fraichard, Amaury Negre, Patrick Reignier, Dominique Vaufreydaz, Nicolas Bonnefond, Rémi Pincent, Mayeul de Werbier d Antigneul, Rémi Barraquand, David Lombard.

Ambient Intelligence, Equipment d’Excellence, Investissement d’Avenir

The AmiQual4Home Innovation Factory is an open research facility for innovation and experimentation with human-centered services based on the use of large-scale deployment of interconnected digital devices capable of perception, action, interaction and communication. The Innovation Factory is composed of a collection of workshops for rapid creation of prototypes, surrounded by a collection of living labs and supported by a industrial innovation and transfer service. Creation of the Innovation Factory has been made possible by a 2.140 Million Euro grant from French National programme "Investissement d’avenir", together with substantial contributions of resources by Grenoble INP, Univ Joseph Fourier, UPMF, CNRS, Schneider Electric and the Commune of Montbonnot. The objective is to provide the academic and industrial communities with an open platform to enable research on design, integration and evaluation of systems and services for smart habitats.

The AmiQual4Home Innovation Factory is a unique combination of three different innovation instruments: (1) Workshops for rapid prototyping of devices that embed perception, action, interaction and communication in ordinary objects based on the MIT FabLab model, (2) Facilities for real-world test and evaluation of devices and services organised as open Living Labs, (3) Resources for assisting students, researchers, entrepreneurs and industrial partners in creating new economic activities. The proposed research facility will enable scientific research on these problems while also enabling design and evaluation of new forms of products and services with local industry.

The core of the AmiQual4Home Innovation Factory is a Creativity Lab composed of a collection of five workshops for the rapid prototyping of devices that integrate perception, action, interaction and communications into ordinary objects. The Creativity Lab is surrounded by a collection of six Living Labs for experimentation and evaluation in real world conditions. The combination of fabrication facilities and living labs will enable students, researchers, engineers, and entrepreneurs to experiment in co-creation and evaluation. The AmiQual4Home Innovation Factory will also include an innovation and transfer service to enable students, researchers and local entrepreneurs to create and grow new commercial activities based on the confluence of digital technologies with ordinary objects. The AmiQual4Home Innovation Factory will also provide an infrastructure for participation in education, innovation and research activities of the European Institute of Technology (EIT) KIC ICTLabs.

The AmiQual4Home Innovation Factory will enable a unique new form of coordinated ICT-SHS research that is not currently possible in France, by bringing together expertise from ICT and SHS to better understand human and social behaviour and to develop and evaluate novel systems and services for societal challenges. The confrontation of solutions from these different disciplines in a set of application domains (energy, comfort, cost of living, mobility, well-being) is expected to lead to the emergence of a common, generic foundation for Ambient Intelligence that can then be applied to other domains and locations. The initial multidisciplinary consortium will progressively develop interdisciplinary expertise with new concepts, theories, tools and methods for Ambient Intelligence.

The potential impact of such a technology, commonly referred to as "Ambient Intelligence", has been documented by the working groups of the French Ministry of Research (MESR) [32] as well as the SNRI (Stratégie Nationale de la Recherche et de l'Innovation).

In 2013 our efforts were focused on specifying the requirements for major components of the project, and on finalising contractual issues with ANR. We defined the layout and arrangement of the Creativity Lab workshops, we started the specification of the instrumentation needed for the Living Labs, and developed a first version of a set of easy-deployable wireless sensors for infield data acquisition, that we call the Rapid Deployment Minikit. A set of CNC machines was purchased, including a Fused Filament Fabrication 3D printer, a CO2 Laser cutter and engraver, and a CNC mill.

Following the kickoff meeting of the project held in October 2013, we received positive feedback and interest form local industry and research institutions, and several national project proposals submitted in 2013 identified AmiQual4Home as an important ressource.

7.2.2. FUI PRAMAD

Participants: Claudine Combe, Lucas Nacsa, Maxime Portaz, Amaury Negre, Dominique Vaufreydaz [correspondant].

Pramad is a collaborative project about *Plateforme Robotique dÁssistance et de Maintien à Domicile*. There are seven partners:

- R&D/industry: Orange Labs (project leader) and Covéa Tech (insurance company),

- Small companies: Interaction games (game designer, note that Wizardbox, the original partner was bought by Interaction games) and Robosoft (robot).
- Academic labs: Inria/PRIMA, ISIR (Paris VI) and Hôpital Broca (Paris).

The objectives of this project are to design and evaluate robot companion technologies to maintain frail people at home. Working with its partners, PRIMA research topics are:

- social interaction,
- robotic assistance,
- serious game for frailty evaluation and cognitive stimulation.

7.2.3. Inria Project-Teams PAL

Participants: Rémi Barraquand, Thierry Fraichard, Patrick Reignier, Amaury Negre, Dominique Vaufreydaz [correspondant].

The 12 Inria Project-Teams (IPT) participating in a Large-scale initiative action Personally Assisted Living (PAL http://pal.inria.fr) propose to work together to develop technologies and services to improve the autonomy and quality of life for elderly and fragile persons. The goal of this program is to unite these groups around an experimental infrastructure, designed to enable collaborative experimentations.

PAL is organized arround 12 IPT:

Demar, E-Motion, Flowers, Hephaistos, Lagadic, Lagadic-Sophia, Maia, Madynes, Phoenix, Prima, Stars and Reves.

The role of PRIMA within this project is to develop new algorithms mainly along two research axes:

- assessing frailty degree of the elderly,
- social interaction.

7.3. International Research Visitors

7.3.1. Visits to International Teams

7.3.1.1. Sabbatical programme

Fraichard Thierry, Sabbatical Visit to BIU, ISRAEL, from May 2014 - May 2015

7.3.1.2. Research stays abroad

Varun Jain, 6 month visit to Carnegie-Mellon University, Pittsburgh PA, on a Region Rhone-Alpes ExploraDoc Grant from January 2014 to July 2014

8. Dissemination

8.1. Promoting Scientific Activities

8.1.1. Scientific events organisation

8.1.1.1. Member of the organizing committee

- Sabine Coquillart is member of the organizing committee and co-chair of the Program Committee for GRAPP 2014 Lisbon, Portugal.
- Sabine Coquillart is member of the organizing committee for IEEE VR 2015, Arles, France

8.1.1.2. Member of the conference program committee

- James Crowley: Ubicomp 2014.
- Sabine Coquillart: 3DCVE 2014, GRAPP 2014, IEEE 3DUI 2014, IEEE VR 2014, ISVC 2014, SVR 2014-13-12-11-10, VRST 2014-13, WSCG 2014.

- Thierry Fraichard: Associate Editor for ICRA 2014 and IROS 2014.

8.1.1.3. Reviewer

- James Crowley: Ubicomp 2014, ICPR 2014. CVPR 2014, IE 2014, ICRA 2014.
- Sabine Coquillart: 3DCVE 2014, GRAPP 2014, IEEE 3DUI 2014, IEEE VR 2014, ISVC 2014, SVR 2014-13-12-11-10, VRST 2014-13, WSCG 2014.
- Thierry Fraichard: IROS 2014.
- Patrick Reignier: IUI 2014, ROMAN 2014, UCAmI 2014,
- Dominique Vaufreydaz: MTEL 2014, UBICOMM 2014, ROMAN 2014], Workshop on Assistance and Service Robotics in a Human Environment at IROS 2014, WUSPE2013

8.1.2. Journal

8.1.2.1. Member of the editorial board

- Dominique Vaufreydaz: International Journal On Advances in Internet Technology 2014, Journal of Robotics and Autonomous Systems, special issue: Assistive Robotics

8.2. Teaching - Supervision - Juries

8.2.1. Teaching

8.2.1.1. James Crowley

- Computer Vision, Course 24h EqTD, M2 year, Master of Science in Informatics at Grenoble,
- Intelligent Systems, Cours 54h EqTD, ENSIMAG.

James Crowley is Director Master of Science in Informatics at Grenoble.

8.2.1.2. Sabine Coquillart

- Sabine Coquillart teaches a course on Virtual Reality and 3D User Interfaces for the GVR Master 2R, 2013-2014.
- Sabine Coquillart teaches a one day course on 3D User Interfaces and Augmented Reality for the Numerical Modeling and Virtual Reality Master 2 in Laval, 2013-2014.

8.2.1.3. Thierry Fraichard

Master: Thierry Fraichard, Introduction to Perception and Robotics, 23h eqTD, M1 MOSIG, Univ. of Grenoble, France.

Master: Thierry Fraichard, Motion in Dynamic Workspaces, 2h eqTD, Computer SCience Master, Bar Ilan Univ., Israel.

Co-responsibility of the Graphic, Vision and Robotics track of the international MOSIG Master programme.

8.2.1.4. Patrick Reignier

Master: Patrick Reignier, Projet GÃ©nie Logiciel, 55h eqTD, M1, ENSIMAG/Grenoble INP, France.

Master : Patrick Reignier, Developpement d'applications communicantes, 18h eqTD, M2, ENSIMAG/Grenoble - INP, France

Master : Patrick Reignier, Introduction aux applications rÃ©parties, 18h eqTD, M2, ENSIMAG/Grenoble - INP, France

Master: Patrick Reignier, Programmation Internet, 18h eqTD, M1, ENSIMAG/Grenoble INP, France

Master : Patrick Reignier, Algorithmique, 50h eq TD, M1, ENSIMAG/Grenoble INP, France

Licence: Patrick Reignier, Projet C, 20h eqTD, L3, ENSIMAG/Grenoble INP, France.

Patrick Reignier is supervising the industrial part of the "formation en apprentissage" of the ENSIMAG engineering school.

8.2.2. Supervision

8.2.2.1. PhD defended in 2014

- Julian Quiroga, "Scene Flow from RGBD Images", Th√®se doctorale de l'Universit√© de Grenoble, 7 November 2014, Directed by Frederic Devernay and James L. Crowley
- Evanthia Mavridou, "Robust Image Description Using the Laplacian Profile and Radial Fourier Transform", Th√®se doctorale de l'Universit√© de Grenoble, 25 November 2014, directed by James L. Crowley and Augustin Lux.
- Dimitri Mason, "Inspirez ! Explorez ! Soutien √† la cr√©ativit√© en conception d'interfaces homme-machine", Th√®se doctorale de l'Universit√© de Grenoble, 25 Sept 2014, Directed by Alexandre Demeure and Gaelle Calvary.

8.2.2.2. Current Doctoral Students

- Varun Jain, Perception of Human Emotions, University of Grenoble, expected March 2015, James Crowley
- Sergi Pujades-Rocamora, Modeles de cameras et algorithmes pour la creation de contenu video 3D, expected fall 2015, Remi Ronfard (HDR) et Frederic Devernay
- Etienne Balit, Multimodalite et interaction sociale, University of Grenoble, expected Fall 2016, Patrick Reignier (Professor), Dominique Vaufreydaz.
- Viet Cuong Ta, Multiple Users localization in public large-scale space, University of Grenoble, expected Fall 2016, Eric Castelli (HDR, Mica laboratory Hanoi, Vietnam), Dominique Vaufreydaz.
- Remi Paulin, Human-Robot Motion, University of Grenoble, expected Fall 2016, Thierry Fraichard (HDR).
- Chen Jingtao, Doctoral student of University of Grenoble, expected 2017, directed by S. Coquillart Gr√©goire Nieto, Doctoral student of University of Grenoble, expected 2017, directed by F. Devernay Romain Br√©gier, Doctoral student of University of Grenoble, expected 2017, directed by F. Devernay, J. L. Crowley Alberto Quintero Delgado, Doctoral student of University of Grenoble, expected 2017, directed by F. Devernay

8.2.3. Juries

Thierry Fraichard served as an jury member for the PhD of Jean Gregoire from Ecole Nationale Superieure des Mines de Paris, "Priority-Based Coordination of Mobile Robots", and for the PhD of Helene Vorobieva from Universite d'Evry, "Conception et methode de validation des lois de controle pour des systemes de conduite automatisee du vehicule".

James L Crowley served on the following doctoral and Habilitation Juries:

- Président, Thesis Jury of Arturo ESCOBEDO, Thèse Doctorale Université de Grenoble, Oct 2014
- Président, Thesis Jury of Simon COURTEMANCHE, Thèse Doctorale Université de Grenoble, Oct 2014
- Président, Thesis Jury of Yann LAURILLAU, Habilitation, Université de Grenoble, Nov 2014
- Président, Thesis Jury of Vineet GANDHI, Thèse Doctorale Université de Grenoble, Dec 2014

Patrick Reignier served on the following doctoral Juries:

- Member doctoral jury of Aysha Kashif, Thèse Doctorale Université de Grenoble, Jan 2014
- Reporter doctoral jury of Olfa Mabrouk University Paris Est Créteil, Nov 2014

Sabine Coquillard servec on the following doctoral juries:

- Member doctoral jury of Thi Thong Huyen NGUYEN, INSA Rennes. Nov. 2014
- Member doctoral jury of Reporter doctoral jury of Jérôme Ardouin, INSA Rennes, Dec. 2014

8.3. Popularization

- James Crowley participated in the Radio emission Science Publique presented by Michel Alberganti on Radio France Culture on Friday 21 February from 14h00 to 15h00. Theme of the program: "Should we fear intelligent robots?"
- Patrick Reignier was interviewed by a journalist of Science et Vie for a paper on Artificial Intelligence

9. Bibliography

Major publications by the team in recent years

[1] O. BRDICZKA, J. CROWLEY, P. REIGNIER. *Learning Situation Models in a Smart Home*, in "IEEE Transactions on System, Man and Cybernetics, part B", 2009, vol. 39, n^{o} 1, http://hal.inria.fr/hal-00665319/en

[2] O. BRDICZKA, M. LANGET, J. MAISONNASSE, J. L. CROWLEY. *Detecting human behavior models from multimodal observation in a smart home*, in "Automation Science and Engineering, IEEE Transactions on", 2009, vol. 6, n^{o} 4, pp. 588–597

[3] O. BRDICZKA, J. MAISONNASSE, P. REIGNIER, J. CROWLEY. *Detecting Small Group Activities from Multimodal Observations*, in "International Journal of Applied Intelligence", 2009, vol. 30, n^{o} 1, pp. 47–57, http://www-prima.imag.fr/jlc/papers/articleAI_brdiczka12Mars07.pdf

[4] J. COUTAZ, J. CROWLEY, S. DOBSON, D. GARLAN. *Context is key*, in "Communications of the ACM", March 2005, vol. 48, n^{o} 3, pp. 49-53

[5] J. CROWLEY, O. BRDICZKA, P. REIGNIER. *Learning Situation Models for Understanding Activity*, in "The 5th International Conference on Development and Learning 2006 (ICDL06)", Bloomington, Il., USA, June 2006

[6] J. CROWLEY, J. COUTAZ, F. BÉRARD. *Things that See: Machine Perception for Human Computer Interaction*, in "Communications of the A.C.M.", March 2000, vol. 43, n^{o} 3, pp. 54-64

[7] J. CROWLEY, J. COUTAZ, G. REY, P. REIGNIER. *Using Context to Structure Perceptual Processes for Observing Activity*, in "UBICOMP", Sweden, September 2002

[8] N. GOURIER, D. HALL, J. L. CROWLEY. *Facial features detection robust to pose, illumination and identity*, in "Systems, Man and Cybernetics, 2004 IEEE International Conference on", IEEE, 2004, vol. 1, pp. 617–622

[9] N. GOURIER, J. MAISONNASSE, D. HALL, J. L. CROWLEY. *Head pose estimation on low resolution images*, in "Multimodal Technologies for Perception of Humans", Springer Berlin Heidelberg, 2007, pp. 270–280

[10] D. HALL. *Automatic parameter regulation of perceptual systems*, in "Image and Vision Computing", August 2006, vol. 24, n^{o} 8, pp. 870–881, http://dx.doi.org/10.1016/j.imavis.2006.02.011

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] J. QUIROGA. *Scene Flow Estimation from RGBD Images*, Grenoble University, November 2014, https://hal.archives-ouvertes.fr/tel-01097763

Articles in International Peer-Reviewed Journals

[12] J. QUIROGA, F. DEVERNAY, J. L. CROWLEY. *Local scene flow by tracking in intensity and depth*, in "Journal of Visual Communication and Image Representation", January 2014, vol. 25, n^o 1, pp. 98-107 [*DOI :* 10.1016/J.JVCIR.2013.03.018], https://hal.inria.fr/hal-00817011

International Conferences with Proceedings

[13] S. BOURAINE, T. FRAICHARD, O. AZOUAOUI, H. SALHI. *Passively Safe Partial Motion Planning for Mobile Robots with Limited Field-of-Views in Unknown Dynamic Environments*, in "IEEE Int. Conf. on Robotics and Automation (ICRA)", Hong Kong, Hong Kong SAR China, June 2014, https://hal.inria.fr/hal-01018463

[14] *Best Paper*
T. FRAICHARD, R. PAULIN, P. REIGNIER. *Human-Robot Motion: An Attention-Based Navigation Approach*, in "IEEE Int. Symp. on Robot and Human Interactive Communication (ROMAN)", Edinburgh, United Kingdom, August 2014, https://hal.inria.fr/hal-01018471.

[15] V. JAIN, J. L. CROWLEY, A. DEY, A. LUX. *Depression Estimation Using Audiovisual Features and Fisher Vector Encoding*, in "ACM Multimedia 2014", Orlando, FL, United States, November 2014 [*DOI :* 10.1145/2661806.2661817], https://hal.inria.fr/hal-01081358

[16] V. JAIN, J. L. CROWLEY, A. LUX. *Local Binary Patterns Calculated Over Gaussian Derivative Images*, in "ICPR 2014 - 22nd International Conference on Pattern Recognition", Stockholm, Sweden, IEEE Computer Society, August 2014 [*DOI :* 10.1109/ICPR.2014.683], https://hal.inria.fr/hal-01061099

[17] E. MAVRIDOU, J. L. CROWLEY, A. LUX. *Multiscale Shape Description with Laplacian Profile and Fourier Transform*, in "ICIAR - International Conference on Image Analysis and Recognition", Vilamoura, Algarve, Portugal, October 2014, pp. 46 - 54 [*DOI :* 10.1007/978-3-319-11758-4_6], https://hal.inria.fr/hal-01079658

[18] E. MAVRIDOU, M.-D. HOANG, J. L. CROWLEY, A. LUX. *Scale Normalized Radial Fourier Transform as a Robust Image Descriptor*, in "ICPR - 22nd International Conference on Pattern Recognition", Stockholm, Sweden, August 2014, https://hal.inria.fr/hal-01065463

[19] S. PUJADES, L. BOIRON, R. RONFARD, F. DEVERNAY. *Dynamic Stereoscopic Previz*, in "International Conference on 3D Imaging", Liege, Belgium, Professeur Jacques G. Verly, December 2014, https://hal.inria.fr/hal-01083848

[20] S. PUJADES, F. DEVERNAY, B. GOLDLUECKE. *Bayesian View Synthesis and Image-Based Rendering Principles*, in "CVPR - 27th IEEE Conference on Computer Vision and Pattern Recognition", Columbus, United States, June 2014, pp. 3906 - 3913 [*DOI :* 10.1109/CVPR.2014.499], https://hal.inria.fr/hal-00983315

[21] S. PUJADES, F. DEVERNAY. *Viewpoint Interpolation: Direct and Variational Methods*, in "ICIP - International Conference on Image Processing", Paris, France, October 2014, https://hal.inria.fr/hal-01058327

[22] J. QUIROGA, T. BROX, F. DEVERNAY, J. L. CROWLEY. *Dense Semi-Rigid Scene Flow Estimation from RGBD images*, in "ECCV 2014 - European Conference on Computer Vision", Zurich, Switzerland, September 2014 [*DOI :* 10.1007/978-3-319-10584-0_37], https://hal.inria.fr/hal-01021925

[23] L. RUMMELHARD, A. NÈGRE, M. PERROLLAZ, C. LAUGIER. *Probabilistic Grid-based Collision Risk Prediction for Driving Application*, in "ISER", Marrakech/Essaouira, Morocco, June 2014, https://hal.inria.fr/hal-01011808

[24] D. VAUFREYDAZ, A. NÈGRE. *MobileRGBD, An Open Benchmark Corpus for mobile RGB-D Related Algorithms*, in "13th International Conference on Control, Automation, Robotics and Vision", Singapour, Singapore, December 2014, https://hal.inria.fr/hal-01095667

National Conferences with Proceedings

[25] S. PUJADES, F. DEVERNAY. *Interpolation de points de vue : approches directe et variationnelle*, in "Reconnaissance de Formes et Intelligence Artificielle (RFIA) 2014", France, June 2014, https://hal.archives-ouvertes.fr/hal-00989031

Research Reports

[26] T. FRAICHARD. *Will the Driver Seat Ever Be Empty?*, March 2014, n^o RR-8493, https://hal.inria.fr/hal-00965176

[27] T. FRAICHARD, R. PAULIN, P. REIGNIER. *Human-Robot Motion: Taking Attention into Account*, March 2014, n^o RR-8487, https://hal.inria.fr/hal-00965157

Other Publications

[28] J. COUTAZ, E. FONTAINE, N. MANDRAN, A. DEMEURE. *About Composing Our Own Smart Home*, February 2014, EUD4Services (End-User Development for Services) workshop, in Conjunction with AVI 2010, Rome, May 2010, https://hal.inria.fr/hal-00953496

References in notes

[29] *Robosoft company*, 2013, http://www.robosoft.com/

[30] M. BARTLETT, G. LITTLEWORT, M. FRANK, C. LAINSCSEK, I. FASEL, J. MOVELLAN. *Fully Automatic Facial Action Recognition in Spontaneous Behavior*, in "Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on", april 2006, pp. 223 -230, http://dx.doi.org/10.1109/FGR.2006.55

[31] V. COLIN DE VERDIÈRE, J. CROWLEY. *Visual Recognition using Local Appearance*, in "ECCV98", Freiburg, June 1998, pp. 640–654

[32] J. COUTAZ, J. CROWLEY. *Plan « intelligence ambiante »: défis et opportunités*, Groupe de Travail « Intelligence Ambiante » du Groupe de Concertation Sectoriel (GCS3), October 2008

[33] J. COUTAZ, J. CROWLEY, S. DOBSON, D. GARLAN. *Context is Key*, in "Communications of the ACM", March 2005, vol. 48, n^o 3, Special issue on the Disappearing Computer

[34] J. CROWLEY, O. RIFF. *Fast Computation of Scale Normalised Gaussian Receptive Fields*, in "International Conference on Scalespace theories in Computer vision", Skye, UK, June 2003, pp. 584–598

[35] S. DAVIDOFF. *Routine as resource for the design of learning systems*, in "Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing - Adjunct", New York, NY, USA, Ubicomp '10 Adjunct, ACM, 2010, pp. 457–460, http://doi.acm.org/10.1145/1864431.1864486

[36] M. DECROUEZ, R. DUPONT, F. GASPARD, F. DEVERNAY, J. L. CROWLEY. *Détection et apprentissage automatique d'objets pour la modélisation de milieux intérieurs*, in "Actes du 18e congrès francophone sur la Reconnaissance des Formes et l'Intelligence Artificielle (RFIA)", Lyon, France, January 2012, vol. ISBN : 978-2-9539515-2-3, 8 p. , http://liris.cnrs.fr/rfia2012

[37] A. K. DEY, T. SOHN, S. STRENG, J. KODAMA. *iCAP: interactive prototyping of context-aware applications*, in "Proceedings of the 4th international conference on Pervasive Computing", Berlin, Heidelberg, PERVASIVE'06, Springer-Verlag, 2006, pp. 254–271, http://dx.doi.org/10.1007/11748625_16

[38] P. EKMAN, W. V. FRIESEN. *Facial Action Coding System: A Technique for the Measurement of Facial Movement*, Consulting Psychologists Press, 1977

[39] W. FREEMAN, E. ADELSON. *The Design and Use of Steerable Filters*, in "Pattern Analysis and Machine Intelligence", September 1991, vol. 13, n^o 9, pp. 891–906

[40] N. GOURIER, D. HALL, J. CROWLEY. *Facial feature detection robust to pose, illumination, and identity*, in "International Conference on Systems, Man and Cybernetics, Special track on Automatic Facial Expression Analysis", October 2004, pp. 617-622

[41] D. HALL, J. CROWLEY. *Détection du visage par caractéristiques génériques calculées à partir des images de luminance*, in "Reconnaissance des formes et intelligence artificelle", Toulouse, France, 2004, to appear

[42] E. HALL. *A system for the notation of proxemic behaviour*, in "American Anthropologist, New Series", Oct. 1963, vol. 65, n^o 5, pp. 1003-1026

[43] D. HALL, V. COLIN DE VERDIÈRE, J. CROWLEY. *Object Recognition using Coloured Receptive Fields*, in "European Conference on Computer Vision", Dublin, Ireland, June 2000, pp. I 164–177

[44] P. JOHNSON-LAIRD. *How we reason*, Oxford University Press, 2006

[45] P. JOHNSON-LAIRD. *Mental Models: Toward a Cognitive Science of Language, Inference and Consciousness*, Harvard University Press, 1983

[46] P. JOHNSON-LAIRD. *Computer and the Mind: An Introduction to Cognitive Science*, Harvard University Press, 1998

[47] A. KENDON. *Spacing and orientation in co-present interaction*, in "Development of Multimodal Interfaces: Active Listening and Synchrony", A. ESPOSITO, N. CAMPBELL, C. VOGEL, A. HUSSAIN, A. NIJHOLT (editors), Lecture Notes in Computer Science, Springer, 2010, vol. 5967 [*DOI :* 10.1007/978-3-642-12397-9]

[48] J. J. KOENDERINK, A. J. VAN DOORN. *Representation of Local Geometry in the Visual System*, in "Biological Cybernetics", 1987, n^o 55, pp. 367-375

[49] S. KOO, D.-S. KWON, FILTERING EXTENDED KALMAN. *Recognizing Human Intentional Actions from the Relative Movements between Human and Robot*, in "Nonlinear Dynamics", 2009, pp. 939–944

[50] T. LINDEBERG. *Feature Detection with Automatic Scale Selection*, in "International Journal of Computer Vision", 1998, vol. 30, n^o 2, pp. 79–116

[51] F. LINDNER, C. ESCHENBACH. *Towards a Formalization of Social Spaces for Socially Aware Robots*, in "Spatial Information Theory", M. EGENHOFER, N. GIUDICE, R. MORATZ, M. WORBOYS (editors), Lecture Notes in Computer Science, Springer, 2011, vol. 6899 [*DOI :* 10.1007/978-3-642-23196-4]

[52] D. LOWE. *Object Recognition from Local Scale-Invariant Features*, in "ICCV", Corfu, Greece, September 1999, pp. 1150–1157

[53] J. MAISONNASSE, N. GOURIER, O. BRDICZKA, P. REIGNIER. *Attentional Model for Perceiving Social Context in Intelligent Environments*, in "Artificial Intelligence Applications and Innovations", I. MAGLOGIANNIS, K. KARPOUZIS, M. BRAMER (editors), IFIP Int. Federation for Information Processing, Springer, 2006, vol. 204 [*DOI :* 10.1007/0-387-34224-9_20]

[54] A. NEGRE, C. BRAILLON, J. CROWLEY. *Visual navigation from variation of intrinsic scale*, 2006

[55] E. A. SCHEGLOFF. *Body Torque*, in "Social Research", 1998, vol. 65, n^o 3, pp. 535–596 [*DOI :* 10.1017/S0140525X04400053], http://www.questia.com/googleScholar.qst?docId=95791590

[56] B. SCHIELE, J. CROWLEY. *Recognition without Correspondence using Multidimensional Receptive Field Histograms*, in "International Journal of Computer Vision", January 2000, vol. 36, n^o 1, pp. 31–50

[57] K. N. TRUONG, E. M. HUANG, G. D. ABOWD. *CAMP: A Magnetic Poetry Interface for End-User Programming of Capture Applications for the Home*, in "Proceedings of UBICOMP 2004: The 6th International Conference on Ubiquitous Computing", Nottingham, England, September 2004, pp. 143–160

[58] A. VINCIARELLI, M. PANTIC, H. BOURLARD. *Social signal processing: Survey of an emerging domain*, in "Image and Vision Computing", 2009, vol. 27, n^o 12, pp. 1743 - 1759, Special issue : Visual and multimodal analysis of human spontaneous behaviour [*DOI :* 10.1016/J.IMAVIS.2008.11.007], http://www.sciencedirect.com/science/article/pii/S0262885608002485

[59] D. VOGEL, R. BALAKRISHNAN. *Interactive public ambient displays: transitioning from implicit to explicit, public to personal, interaction with multiple users*, in "Proceedings of the 17th annual ACM symposium on User interface software and technology", ACM, 2004, pp. 137–146