IN PARTNERSHIP WITH:
CNRS

Institut polytechnique de Grenoble

Université Joseph Fourier (Grenoble 1)

Activity Report 2011

Project-Team PRIMA

Perception, recognition and integration for observation of activity

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

RESEARCH CENTER
Grenoble - Rhône-Alpes

THEME
Vision, Perception and Multimedia Understanding

Table of contents

Project-Team PRIMA

Keywords: Computer Vision, User Interface, Perception, Ambient Computing, Ubiquitous Computing

1. Members

Research Scientists
Frédéric Devernay [CR1 INRIA]
Sabine Coquillart [DR2 INRIA, HdR]

Faculty Members
James Crowley [Team Leader, Professeur CE2 Grenoble INP, HdR]
Augustin Lux [Professeur Grenoble INP, HdR]
Patrick Reignier [Professeur Grenoble INP, HdR]
Dominique Vaufreydaz [MdC Université Pierre Mendès-France]
Alexandre Demeure [MdC Université Joseph Fourier]

Technical Staff
Amaury Negre [Engineer CNRS (Shared with E-Motion)]
Nicolas Gourrier [Engineer CDD INRIA]
Claudine Combe [Engineer CDD INRIA]
Thibault Serot [Engineer CDD INRIA]
Matthieu Volat [Engineer CDD INRIA]
Sylvain Duchêne [Engineer CDD INRIA]
Sergi Pujades-Rocamora [Engineer CDD INRIA]
Rémi Barraquand [Engineer CDD INRIA]
Lukas Rummelhard [Engineer CDD INRIA]

PhD Students
Han Yue [Doctoral Student (Bourse MENSR)]
John Alexander Ruiz Hernandez [Doctoral Student (Bourse INRIA)]
Antoine Meler [Doctoral Student (Bourse MENSR)]
Mathieu Guillame-Bert [Doctoral Student (Bourse MENSR)]
Marion Decrouez [Doctoral Student (Bourse CEA)]
Dimitri Masson [Doctoral Student (Bourse MENSR)]
Evanthia Mavridou [Doctoral Student (Bourse MESR)]
Julian Quiroga [Doctoral Student (Bourse Columbian)]
Varun Jain [Doctoral Student (Bourse INRIA)]

Administrative Assistant
Marion Ponsot [Secretary (SAR) INRIA]

Other
Wafa Benkaour [Masters Student]

2. Overall Objectives

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

Smart Environments, Computer Vision, Machine Perception, Human-Computer Interaction, Perceptual User Interfaces

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 human-computer 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, infrared other environmental 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 it's 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. Adding personal robots as peripheral extensions to a smart environment greatly extends the range of services that can be offered.

An environment may be considered as "interactive" when it is capable interacting to 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 [43]. 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 multi-modal 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, 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

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

2.2. Highlights

- In January 2010, the research laboratories of the Grenoble Universities were evaluated by the French AERES evaluation agency. Within this context, each of the 24 research groups of the Laboratoire Informatique de Grenoble (LIG) were individually evaluated with respect to 4 criteria: Scientific Quality, Visibility, Governance, and Scientific Project, as well as overall activity. The results were labeled with a grade (A+, A, B, C). The PRIMA research group received an overall score of A+, and was on of only two research groups in Grenoble to receive a score of A+ in all four areas: Scientific Quality, Visibility, Governance, and Scientific Project.
- Members of the PRIMA group have coordinated the submission of the proposal AMIQUAL4HOME to the second round of funding for Equipment for Scientific Excellence (EQUIPEX) of the programme Investissement d'Avenir. AMIQUAL4HOME has been ranked 45 of the 270 proposals submitted and is listed among the Laureat of the 2011 Equipex programme.
- PRIMA director James Crowley has been named as a senior member of the Institut Universitaire de France (IUF) with a research programme entitled "Perception for Social Interaction".

3. Scientific Foundations

3.1. Context Aware 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 [Johnson-Laird 83], [Johnson-Laird 98]. 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 [Crowley 03], [Coutaz 05].

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 context (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 (more details at the end of this section).
- *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 (in a group of meeting participants, who is interacting with whom and thus which interaction groups are formed)

Currently situation models are constructed by hand. Our current 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 curently exploring reinforcement learning approaches to solve this problem.

With a reinforcement learning approach, the system is rewarded and punished by user reactions to system behaviors. A simplified stereotypic interaction model assures a initial behavior. 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 camera-man 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. Robust architectures for multi-modal perception

3.2.1. Summary

Machine perception is notoriously unreliable. Even in controlled laboratory conditions, programs for speech recognition or computer vision generally require supervision by highly trained engineers. Practical real-world use of machine perception requires fundamental progress in the way perceptual components are designed and implemented. A theoretical foundation for robust design can dramatically reduce the cost of implementing new services, both by reducing the cost of building components, and more importantly, by reducing the obscure, unpredictable behaviour that unreliable components can create in highly complex systems. To meet this challenge, we propose to adapt recent progress in autonomic computing to the problem of producing reliable, robust perceptual components.

Autonomic computing has emerged as an effort inspired by biological systems to render computing systems robust [42]. Such systems monitor their environment and internal state in order to adapt to changes in resource availability and service requirements. Monitoring can have a variety of forms and raises a spectrum of problems. An important form of monitoring relies on a description of the system architecture in terms of software components and their interconnection. Such a model provides the basis for collecting and integrating information from components about current reliability, in order to detect and respond to failure or degradation in a component or changes in resource availability (auto-configuration). However, automatic configuration, itself, imposes constraints on the way components are designed, as well as requirements on the design of the overall system [38].

Robust software design begins with the design of components. The PRIMA project has developed an autonomic software architecture as a foundation for robust perceptual components. This architecture allows experimental design with components exhibiting, auto-regulation, self-description, self-Monitoring and self-configuration and self-repair. Maintenance of such autonomic properties can result in additional computing overhead within components, but can pay back important dividends in system reliability.

3.2.2. Detailed Description

Components based programming makes it possible to design systems that can be dynamically reconfigured during run-time. Reconfiguration can be achieved by having each component provide a description of its parameters, input data and output data using a standardized XML schema. Such XML descriptions can be recorded in a component registry and used to adapt interfaces either manually or automatically. Such XML descriptions are an example of the principle of self-description that characterizes Autonomic Systems [44]. Other such principles are defined at the component level and the systems integration level. At the component level, in addition to self-description one finds techniques for "auto-initialization", "self-regulation", self-monitoring and "performance reporting". At the systems level, one finds methods for "self-configuration", self-repair, and system supervision.

An important form of monitoring relies on a description of the system architecture in terms of software components and their interconnections. Such a model provides the basis for collecting and integrating information from components about current reliability, in order to detect and respond to failure or degradation in a component or changes in resource availability (auto-configuration). However, automatic configuration, itself, imposes constraints on the way components are designed, as well as requirements on the design of the overall system [38], [44]. The PRIMA group has taken a leading role in introducing autonomic system approachs to programming perceptual systems.

Robust software design begins with the design of components. The PRIMA project has developed an autonomic software architecture as a foundation for robust perceptual components. This architecture allows experimental design with components exhibiting:

Auto-criticism: Every computational result produced by a component is accompanied by an estimate of its reliability.

Auto-regulation: The component regulates its internal parameters so as to satisfy a quality requirement such as reliability, precision, rapidity, or throughput.

Self-description: The component can provide a symbolic description of its own functionality, state, and parameters.

Self-Monitoring: the component can provide a report on its internal state in the form of a set of quality metrics such as throughput and load.

Self-configuration: The component reconfigures its own modules so as to respond to changes in the operating environment or quality requirements [52].

Self-repair: The component can react to errors detected by self-monitoring by changing operating parameters, switching among execution modules, or editing internal data structures.

Maintenance of such autonomic properties can result in additional computing overhead within components, but can pay back important dividends in system reliability.

The PRIMA software architecture for supervised autonomic perceptual components [30], [31], is shown in figure 1. In this design, perceptual components use a supervisory controller to dynamically configure, schedule and execute a set of modules in a cyclic detection and tracking process.

The supervisory controller provides five fundamental functions: command interpretation, execution scheduling, event handling, parameter regulation, and reflexive description. The supervisor acts as a programmable interpreter, receiving snippets of code script that determine the composition and nature of the process execution cycle and the manner in which the process reacts to events. The supervisor acts as a scheduler, invoking execution of modules in a synchronous manner. The supervisor handles event dispatching to other processes, and reacts to events from other processes. The supervisor regulates module parameters based on the execution results. Auto-critical reports from modules permit the supervisor to dynamically adapt processing. Finally, the supervisor responds to external queries with a description of the current state and capabilities.

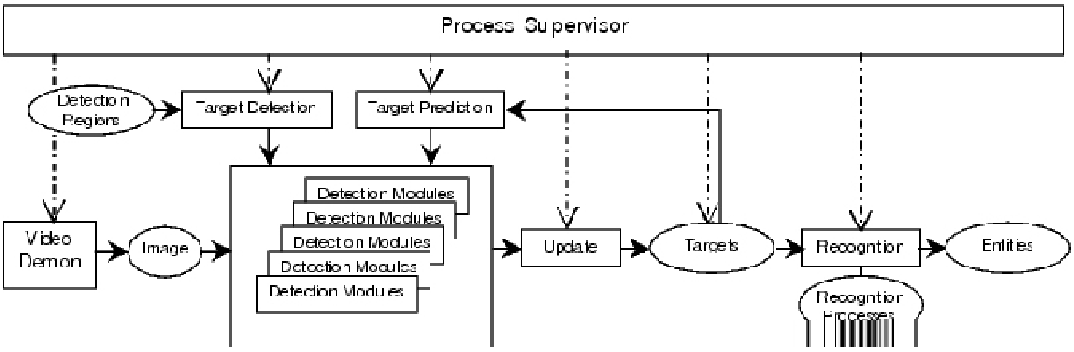


Figure 1. Architecture for an autonomic perceptual component

Real-time visual processing for the perceptual component is provided by tracking. Tracking conserves information about over time, thus provides object constancy. Object constancy assures that a label applied to a blob at time T1 can be used at time T2. Tracking enables the system focus attention, applying the appropriate detection processes only to the region of an image where a target is likely to be detected. Also the information about position and speed provided by tracking can be very important for describing situations.

Tracking is classically composed of four phases: Predict, observe, detect, and update. The prediction phase updates the previously estimated attributes for a set of entities to a value predicted for a specified time. The observation phase applies the prediction to the current data to update the state of each target. The detect phase detects new targets. The update phase updates the list of targets to account for new and lost targets. The ability to execute different image processing procedures to process target information with an individual ROI is useful to simultaneously observe a variety of entities.

Homeostasis, or "autonomic regulation of internal state" is a fundamental property for robust operation in an uncontrolled environment. A process is auto-regulated when processing is monitored and controlled so as to maintain a certain quality of service. For example, processing time and precision are two important state variables for a tracking process. These two may be traded off against each other. The component supervisor maintains homeostasis by adapting module parameters using the auto-critical reports from modules

An auto-descriptive controller can provide a symbolic description of its capabilities and state. The description of the capabilities includes both the basic command set of the controller and a set of services that the controller may provide to a more abstract supervisor. Such descriptions are useful for both manual and automatic assembly of components.

In the context of recent National projects (RNTL ContAct) and European Projects (FAME, CAVIAR, CHIL), the PRIMA perceptual component has been demonstrated with the construction of perceptual components for

1. Tracking individuals and groups in large areas to provide services,
2. Monitoring a parking lot to assist in navigation for an autonomous vehicle.
3. Observing participants in an meeting environment to automatically orient cameras.
4. Observing faces of meeting participants to estimate gaze direction and interest.
5. Observing hands of meeting participants to detect 2-D and 3D gestures.
6. Observing and monitoring the activities of groups to understand social interaction

3.3. Service Oriented Architectures for Intelligent Environments

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 (activity modeling, personal assistant). OMiSCID has facilitated cooperation by experts from within the PRIMA project as well as in projects with external partners.

Experiments with semantic service description and spontaneous service composition are conducted around the OMiSCID middleware. In these experiments, attention is paid to usability. A dedicated language has been designed to allow developers to describe the functionalities that their services provide. This language aims at simplifying existing semantic web services technologies to make them usable by a normal developer (i.e. that is not specialized in the semantic web). This language is named the User-oriented Functionality Composition Language (UFCL).

UFCL allows developers to specify three types of knowledge about services:

- The knowledge that a service exposes a functionality like a “Timer” functionality for a service emitting message at a regular frequency.
- The knowledge that a kind of functionality can be converted to another one. For example, a “Metronome” functionality issued from a music centered application can be seen as a “Timer” functionality.
- The knowledge that a particular service is a factory and can instantiate other services on demand. A TimerFactory can for example start a new service with a “Timer” functionality with any desired frequency. Factories greatly helps in the deployment of service based applications. UFCL factories can also express the fact that they can compose existing functionalities to provide another one.

To bring the UFCL descriptions provided by the developers to life, a runtime has been designed to enable reasoning about what functionalities are available, what functionalities can be transformed to another one and what functionalities could be obtained by asking factories. The service looking for a particular functionality has just to express its need in term of functionalities and properties (e.g. a “Timer” with a frequency of 2Hz) and the runtime automates everything else: gathering of UFCL descriptions exposed by all running services, compilation of these descriptions to some rules in a rule-based system, reasoning and creation of a plan to obtained the desired functionality, and potentially invoking service factories to start the missing services.

3.4. User needs analysis method for smart home

3.4.1. Summary

Ubiquitous computing promises unprecedented user services from the flexible and robust combination of software services with the physical world. Our hypothesis is that end users are willing to shape their own interactive spaces by coupling smart artifacts, building imaginative new functionalities that were not anticipated by system designers. Our work is concerned with the fundamental meaning (and human needs) of building confederation of interoperating smart artifacts. The Social sciences offer tools and methods for exploring human needs and behavior. However, the novelty of our problem requires to solicit participants imagination while at the same time controlling the experimentation and respecting the privacy of their intimate home. We set up the DisQo method [29] to explore how far people are ready to envision the interconnection of everyday devices to improve their lives. Results show that services suggested by our family members fall into four categories: Service substitution, Service improvement, Service chaining and Service "starter".

3.4.2. Detailed Description

Drawing on Davidoff's et al. method and conclusions (i.e. "families want more control of their lives" [33]), we focused on "busy" families. The participants have been solicited through bulletin board advertisements, email, as well as from personal relationships.

We are interested in determining how far people are ready to envision the interconnection of everyday devices to improve their lives. For so doing, we have used a combination of interview (good for clarification), playful cultural probe (appropriate for respecting privacy and for improving subjects involvement). The presence of the experimental team (ourselves, from 1 to 3 persons) was limited to 1h30 per family home. Fieldwork was structured as a four-step process: photographing, interview, game, and debriefing.

Step 1: Photographing. Two volunteer family members were asked to take pictures of 10 objects at the rate of 2 objects per room. For each of the 5 rooms of their choice, they were asked to take a picture of one object that they considered to be necessary in their everyday life or that would help them in organizing their lives, as well as a picture of one object that they considered to be superfluous but valuable (typically, a painting). The volunteers (in general, the parents) were not supposed to be in the same room at the same time so that they would not know which pictures the other member had taken. Meanwhile, the experimental team would wait sitting at a place indicated by the parents (typically, the living room where they usually meet with friends and visitors).

Step 2: Interview. We then conducted an interview with all the family members, using the pictures as input material. Questions were directed at understanding the reasons for their choices, the value attached to the objects or the services provided in daily use. Special attention was given to the (many) remote controller(s) typically found in the household environment. We progressively oriented our questions towards novel uses of smart artifacts. In particular, we asked which objects of the house (including those on the pictures) they would qualify as "programmable" (e.g., TV's, washing machines, alarm clocks), "communicating" (e.g., computers, mobile phones), or emotional (i.e. carrying intimate value). This was used as a means to elicit routines and exceptional needs as well as to prepare the game developed in Step 3.

Step 3: Association game. The association game drew on people creativity using the pictures as play cards. Pictures were sorted randomly and presented two at a time (then, three at a time) on the tablet PC. Family members were asked to imagine which service(s) and value(s) these two (or three) objects coupled together would provide them with. Random coupling was designed to solicit imagination in unexpected ways as solutions creativity grows with the semantic distance between elements [48].

Step 4: Debriefing and informal discussion. The last stage was dedicated to debriefing, including opened friendly discussions.

Overall, we have collected comments and objective data for 349 couplings for a total duration of 25 hours of our presence in the 17 family homes. We found a number of facts that are quite consistent with the results reported in prior literature:

1. “Wake-up” time, “on-the-way-to-home” and “arrivinghome” times are key moments to people. To save time and improve efficiency, activities are organized into wellpolished procedures. As a result, exceptions to these routine tasks are sources of stress. Support for avoiding or for solving exceptions is one class of services expected from a smart home.
2. With regard to programming, attitudes range from "I do not want to be assisted" to “It will work 99% of the time, but it will be hell for the other 1%”. Motivation for programming is systematically grounded on a clear straight forward observable benefit

Our data from the association game shows two important results:

1. Family members are prone to envision new services when coupling involves one “communicating” object, or one "programmable" object, at least.
2. The “communicating” capability has more impact than “programmability” on the capacity of family members to imagine new services.

The services suggested by our family members fall into four categories. We illustrate them with the most typical examples drawn from our fieldwork :

1. Service substitution. People have for instance observed that, for the same (sport) events, commentaries on radio broadcasts are richer than those provided by TV.
2. Service improvement. Some household appliances such as washing machines and storage areas, do not provide any convenient way to control and monitor their current internal state. Appliances than are not sufficiently equipped could be improved by coupling them with additional input and output facilities such as those of the TV set.
3. Service chaining. Service chaining is intended to improve comfort, wellbeing as well as resources for the routine, but hectic, activities. For example, ?picking up the towel after the shower would trigger the coffee machine so that coffee would be ready just in time, at the right temperature, along with the radio turned on in the kitchen broadcasting the news using the appropriate sound level?.
4. Service "starter". We have observed that some appliances serve as triggers for services that are expected to be precomposed to support routine activities. Not surprisingly, people also want to have an explicit and reliable control over the home (cf. the worry that 1time, the house would turn into hell).

Based on this results, we currently explores interactive systems that enable end users to programm their smart environment.

3.5. Robust view-invariant Computer Vision

3.5.1. 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.

3.5.2. Detailed Description

The visual appearance of a neighbourhood can be described by a local Taylor series [45]. 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 [46] 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 [55]. 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 [46], and normalized in orientation using direction of the first derivatives [55].

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" [47]. 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" [37] of Gaussian derivatives. The problem is to determine the appropriate orientation. In earlier work by PRIMA members Colin de Verdiere [28], Schiele [55] and Hall [41], proposed normalising the local jet independently at each pixel to the direction of the first derivatives calculated at the intrinsic scale. This has provided promising results for many view invariant image recognition tasks as described in the next section.

Color is a powerful discriminator for object recognition. Color images are commonly acquired in the Cartesian color space, RGB. The RGB color space has certain advantages for image acquisition, but is not the most appropriate space for recognizing objects or describing their shape. An alternative is to compute a Cartesian representation for chrominance, using differences of R, G and B. Such differences yield color opponent receptive fields resembling those found in biological visual systems.

Our work in this area uses a family of steerable color opponent filters developed by Daniela Hall [41]. These filters transform an (R,G,B), into a cartesian representation for luminance and chrominance (L,C1,C2). Chromatic Gaussian receptive fields are computed by applying the Gaussian derivatives independently to each of the three components, (L, C1, C2). The components C1 and C2 encodes the chromatic information in a Cartesian representation, while L is the luminance direction. Chromatic Gaussian receptive fields are computed by applying the Gaussian derivatives independently to each of the three components, (L, C1, C2). Permutations of RGB lead to different opponent color spaces. The choice of the most appropriate space depends on the chromatic composition of the scene. An example of a second order steerable chromatic basis is the set of color opponent filters shown in figure 2.

Key results in this area include

1. Fast, video rate, calculation of scale and orientation for image description with normalized chromatic receptive fields [32].
2. Real time indexing and recognition using a novel indexing tree to represent multi-dimensional receptive field histograms [53].
3. Robust visual features for face tracking [40], [39].

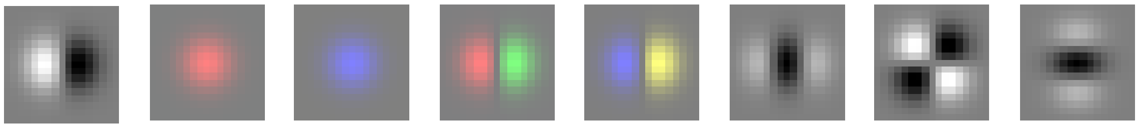


Figure 2. Chromatic Gaussian Receptive Fields ($G_x^L, G^{C_1}, G^{C_2}, G_x^{C_1}, G_x^{C_2}, G_{xx}^L, G_{xy}^L, G_{yy}^L$).

4. Affine invariant detection and tracking using natural interest lines [57].
5. Direct computation of time to collision over the entire visual field using rate of change of intrinsic scale [49].

We have achieved video rate calculation of scale and orientation normalised Gaussian receptive fields using an O(N) pyramid algorithm [32]. 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 at video rates. This method has been used in the MinImage project to provide real time detection, tracking, and identification of faces. It has also been used to provide techniques for estimating age and gender of people from their faces

3.6. New forms of man-machine interaction based on perception

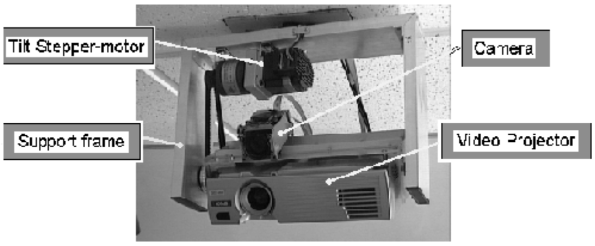
Surfaces are pervasive and play a predominant role in human perception of the environment. Augmenting surfaces with projected information provides an easy-to-use interaction modality that can easily be adopted for a variety of tasks. Projection is an ecological (non-intrusive) way of augmenting the environment. Ordinary objects such as walls, shelves, and cups may become physical supports for virtual functionalities [51]. The original functionality of the objects does not change, only its appearance. An example of object enhancement is presented in [27], where users can interact with both physical and virtual ink on a projection-augmented whiteboard.

Combinations of a camera and a video projector on a steerable assembly [26] are increasingly used in augmented environment systems [50] [54] as an inexpensive means of making projected images interactive. Steerable projectors [26] [51] provide an attractive solution overcoming the limited flexibility in creating interaction spaces of standard rigid video-projectors (e.g. by moving sub windows within the cone of projection in a small projection area [58]).

The PRIMA group has constructed a new form of interaction device based on a Steerable Camera-Projector (SCP) assembly. This device allows experiments with multiple interactive surfaces in both meeting and office environments. The SCP pair, shown in figure 3, is a device with two mechanical degrees of freedom, pan and tilt, mounted in such a way that the projected beam overlaps with the camera view. This creates a powerful actuator-sensor pair enabling observation of user actions within the camera field of view. This approach has been validated by a number of research projects as the DigitalDesk [59], the Magic Table [27] or the Tele-Graffiti application [56].

In October 2008, Doctoral students from the PRIMA group have created the start up company HI-Labs. HI-LABS sells interactive technologies and content management systems for interactive publicity and information kiosks in public places. In 2010 HILABS has begun delivery of its UBICITY interactive store window and by end of 2010 had installed over 100 systems.

4. Application Domains



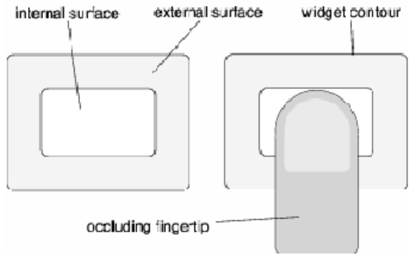


Figure 3. Steerable camera-projector pair (left) and surfaces defined to detect touch-like gestures over a widget (right)

4.1. The Smart Spaces Research Plateforme

Participants: Patrick Reignier, Dominique Vaufreydaz, Remi Barraquand, Augustin Lux, James Crowley, Lukas Rummelhard, Amaury Negre.

Project PRIMA has recently moved to a new Smart Spaces Research Plateforme in order to develop and test components and services for context aware human centered services. The Smart Spaces Research Plateforme is a 50 Square Meter space equipped with a microphone array, wireless lapel microphones, wide angle surveillance cameras, panoramic cameras, steerable cameras, scanning range sensors and two camera-projector video-interaction devices, and a KNX smart electrical system. The microphone array is used as an acoustic sensor to detect, locate and classify acoustic signals for recognizing human activities. The wide-angle and panoramic cameras provide fields of view that cover the entire room, and allows detection and tracking of individuals. Steerable cameras are used to acquire video of activities from any viewing direction.

Context aware human centered services may categorized as tools, advisors, or media. Tool services are designed to perform a specific task or function as robustly as possible. If any adaptation is involved, it should serve to adapt the function to a changing environment. The user interface, and any interaction with users should be perfectly predictable. The degree to which the operation of a tool should be transparent, visible or hidden from the user is an open research question. Advisor services observe the users actions and environment in order to propose information on possible courses of actions. Advisors should be completely obedient and non-disruptive. They should not take initiatives or actions that cannot be overridden or controlled by the user. Media services provide interpersonal communications, entertainment or sensorial extension.

Examples of human centered tool services include:

- An activity log recording system that records the events and activities of an individual's daily activities.
- A service that integrates control of heating, air-conditioning, lighting, windows, window-shades, exterior awnings, etc to provide an optimum comfort level defined in terms of temperature, humidity, CO2, acoustic noise and ambient light level.
- A service that manages the available stock of supplies in a home and orders supplies over the Internet to assure that the appropriate level of supplies are always available.
- A service to measure the walking rate, step size and posture of an elderly person to estimate health and predict the likelihood of a fall.

Some examples of advisor services include:

- A service that provides shopping advice about where and when to shop.

- A service that can propose possible menus based on the available food stuffs in the kitchen.
- A service that observes the activities of humans and appliances within the home and can suggest ways to reduce the cost of heating, electricity or communications.
- A service that observes lifestyle and can offer advice about improving health.

Some examples of media services include

- A service that maintains a sense of informal non-disruptive presence with distant family members.
- A robot device that communicates affection.
- A device that renders the surface temperature of wall, floors and windows to show energy consumption and loss within a house.
- Services that enable seamless tele-presence for communication with others

4.2. Ambient Services for Assisted Living

Participants: Remi Barraquand, Frédéric Devernay, Amaury Negre, James Crowley.

The continued progress in extending life-span, coupled with declining birth rates have resulted in a growing number of elderly people with varying disabilities who are unable to conduct a normal life at home, thereby becoming more and more isolated from society. Governmental agencies including hospitals, healthcare institutions and social care institutions are increasingly overburdened with care of this growing population. Left unchecked, economic and man-power requirements for care of the elderly could well trigger a societal and economic crisis. There is an urgent societal need for technologies and services that allow elderly people to live autonomously in their own environments for longer periods. Smart environments provide a promising new enabling technology for such services.

Adapting smart environments to enhance the autonomy and quality of life for elderly require:

- Robust, plug-and-play sensor technologies monitor the activities and health of elderly in their own home environments.
- Easy to use communications services that allow people to maintain a sense of presence to avoid isolation without disrupting privacy or distracting attention from normal daily activities.
- Architectural frameworks that allow ad hoc composition of services from distributed heterogeneous components scattered throughout the environment.
- Distributed system architectures which allow the cooperation of independent emergency services to work together to provide emergency care,
- Technologies interpret activity to warn of loss of mobility or cognitive function.
- Engineering approaches for the customization/personalization/adaptation of living assistance systems at installation and run time,
- Social, privacy, ethical and legal safeguards for privacy and control of personal data.

4.3. 3-D video processing

Participants: Frédéric Devernay, Matthieu Volat, Sylvain Duchêne, Sergi Pujades-Rocamora.

Stereoscopic cinema has seen a surge of activity in recent years, and for the first time all of the major Hollywood studios released 3-D movies in 2009. This is happening alongside the adoption of 3-D technology for sports broadcasting, and the arrival of 3-D TVs for the home. Two previous attempts to introduce 3-D cinema in the 1950s and the 1980s failed because the contemporary technology was immature and resulted in viewer discomfort. But current technologies — such as accurately-adjustable 3-D camera rigs with onboard computers to automatically inform a camera operator of inappropriate stereoscopic shots, digital processing for post-shooting rectification of the 3-D imagery, digital projectors for accurate positioning of the two stereo projections on the cinema screen, and polarized silver screens to reduce cross-talk between the viewers left- and right-eyes — mean that the viewer experience is at a much higher level of quality than in the past. Even so, creation of stereoscopic cinema is an open, active research area, and there are many challenges from acquisition to post-production to automatic adaptation for different-sized display [35], [36].

Until recently, in order to view stereoscopic 3-D video, the user had to wear special glasses. Recent advances in 3-D displays provide true 3-D viewing experience without glasses. These screens use either a micro-lenticular network or a parallax barrier placed in front of a standard LCD, plasma, or LED display, so that different viewpoints provide different images. If the characteristics of the network and the screen are carefully chosen, the user will perceive two different images from the viewpoints of the left and right eyes. Such glasses-free 3-D screens usually display between 8 and a few dozen different viewpoints.

When the 3-D scene which has to be displayed is computer-generated, it is usually not a problem to generate a few dozen viewpoints. But when a real scene has to displayed, one would have to shoot it through the same number of synchronized cameras as there are viewpoints in order to display it properly. This makes 3-D shooting of real scenes for glasses-free 3-D displays mostly unpractical. For this reason, we are developping high-quality view-interpolation techniques, so that the many different viewpoints can be generated from only a few camera positions [14].

Our research focuses on algorithms derived from Computer Vision and Computer Graphics, applied to live-action stereoscopic 3-D content production or post-production, including [34]:

- Live monitoring of stereoscopic video: geometric image misalignment, depth budget (i.e. limits on horizontal disparity), left-right color balance, left-right depth-of-field consistency [16].
- Live correction of stereoscopic video: correct the above defects in real-time when it is possible, with the help of GPU-based architectures.
- Adaptation of the stereoscopic content to the display size and distance, to avoid divergence or geometric deformations [14].
- Novel camera setups and algorithms for unconstrained stereoscopic shooting (especially when using long focal length).
- Novel camera setups and algorithms for glasses-free 3D displays.
- Stereoscopic inpainting.
- Stereoscopic match-moving.
- Compositing stereoscopic video and matte painting without green screen.
- Relighting of stereoscopic video, especially when videos are composited.

4.4. Simultaneous localization and mapping (SLAM)

Participants: James Crowley, Frédéric Devernay, Marion Decrouez.

Live processing of a video sequence taken from a single camera enables to model an a priori unknown 3D scene. Metrical SLAM (Simultaneous Localization and Mapping) algorithms track the camera pose while reconstructing a sparse map of the visual features of the 3D environment. Such approaches provide the geometrical foundation for many augmented reality applications in which informations and virtual objects are superimposed on live images captured by a camera. Improving such systems will enable in the future precise industrial applications such as guided-maintenance or guided-assembly in wide installations.

A problem with current methods is the assumption that the environment is static. Indoor environments such as supermarket ailes and factory floors may contain numerous objects that are likely to be moved, disrupting a localization and mapping system. We explore methods for automatic detection and modeling of such objects. We define the scene as a static structure that may contain moving objects and objects are defined as a set of visual features that share a common motion compared to the static structure [19]. Using several explorations of a camera in the same scene, we detect and model moved objects while reconstructing the environment. Experiments highligĥt the performance of the method in a real case of localization in an unknown indoor environment.

4.5. User localization in large-scale Smart Spaces

Participants: Dominique Vaufreydaz, Yan Hue, Lukas Rummelhard, Amaury Negre.

Ad-hoc assemblies of mobile devices embedding sensing, display, computing, communications, and interaction provide an enabling technology for smart environments. In the PRIMA project we have adopted a component oriented programming approach to compose smart services for such environments. Common services for smart spaces include

- Services to manage energy in building, including regulating temperature, illumination, and acoustic noise,
- Ambient assisted living services to extend the autonomy of elderly and infirm,
- Logistics management for daily living,
- Communication services and tools for collaborative work,
- Services for commercial environments,
- Orientation and information services for public spaces, and
- Services for education and training.

We are pursuing development of components based on the concept of "large-scale" smart space that is an intelligent environment which will be deployed on a large surface containing several buildings (as a university campus for example). We also define the "augmented man" concept as a human wearing one or many mobile intelligent wireless devices (telephone, Smartphone, pda, notebook). Using all these devices, one can use many different applications (read emails, browse the Internet, file exchange, etc.). By combining the concepts of large-scale perceptive environments and mobile computing, we can create intelligent spaces, it becomes possible to propose services adapted to individuals and their activities. We are currently focussing on two aspects of this problem: the user profile and the user location within a smart space.

A fundamental requirement for such services is the ability to perceive the current state of the environment. Depending on the nature of the service, environment state can require sensing and modeling the physical properties of the environment, the location, identity and activity of individuals within the environment, as well as the set of available computing devices and software components that compose the environment. All of these make up possible elements for context modeling.

Observing and tracking people in smart environments remains a challenging fundamental problem. Whether it is at the scale of a campus, of a building or more simply of a room, we can combine several additional localization levels (and several technologies) to allow a more accurate and reliable user perception system. Within the PRIMA project, we are currently experimenting with a multi-level localization system allowing variable granularity according to the available equipment and the precision required for the targeted service.

5. Software

5.1. OMiSCID Middleware for Distributed Multi-Modal Perception

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

OMiSCID is new 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* [25]). 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 informations and documentations about OMiSCID has been set up to improve the visibility and promote the use of this middleware.

The OMiSCID graphical user interface (GUI) is an extensible graphical application that facilitates analysis and debugging of service oriented applications. The core functionality of this GUI is to list running services, their communication channels and their variables. This GUI is highly extensible and many modules (i.e. plugins) have been created by different members of the team: figure 4 shows an example of some of these modules. OMiSCID GUI is based on the Netbeans platform and thus inherits from its dynamic installation and update of modules.

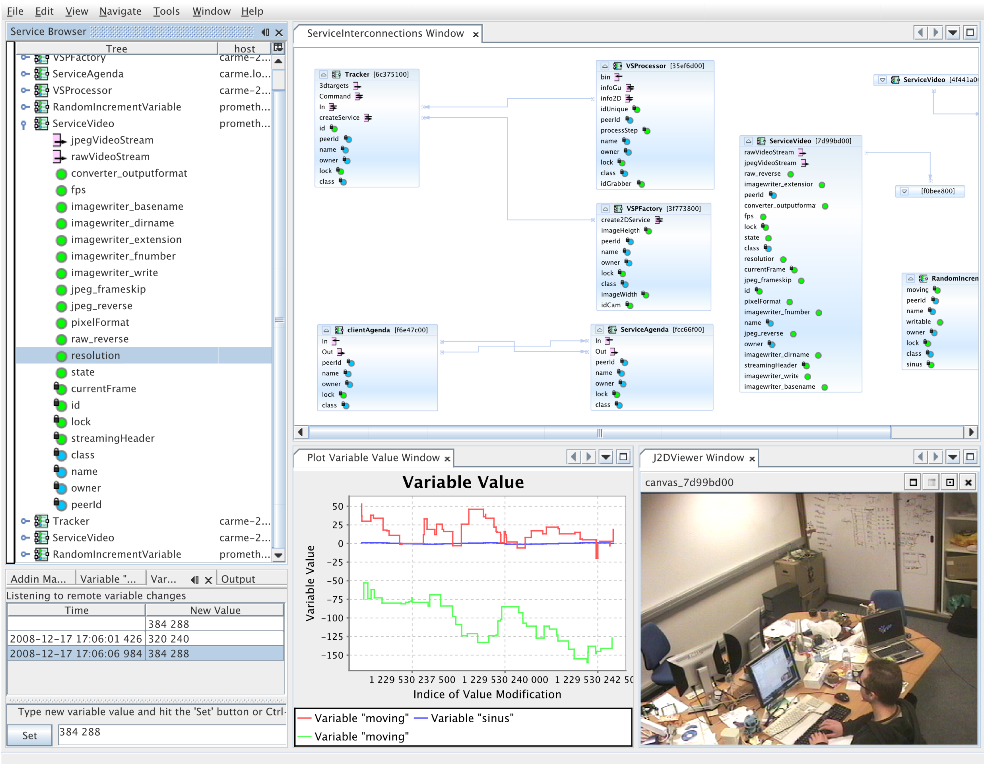


Figure 4. OMiSCID GUI showing a list of running services and some modules for service interconnections, variable plotting, live video stream display and variable control

5.2. 3D Bayesian Tracker

Participants: James Crowley [correspondant], Amaury Negre, Lukas Rummelhard.

The 2DBT and 3DBT tracking systems are autonomic perceptual components originally created for the IST CAVIAR project and the IST CHIL projects. Both systems are autonomous perceptual components managed by a autonomic supervisor. The Autonomic supervisor provides self monitoring, self repair, self configuration, auto-regulation of parameters and self-description.

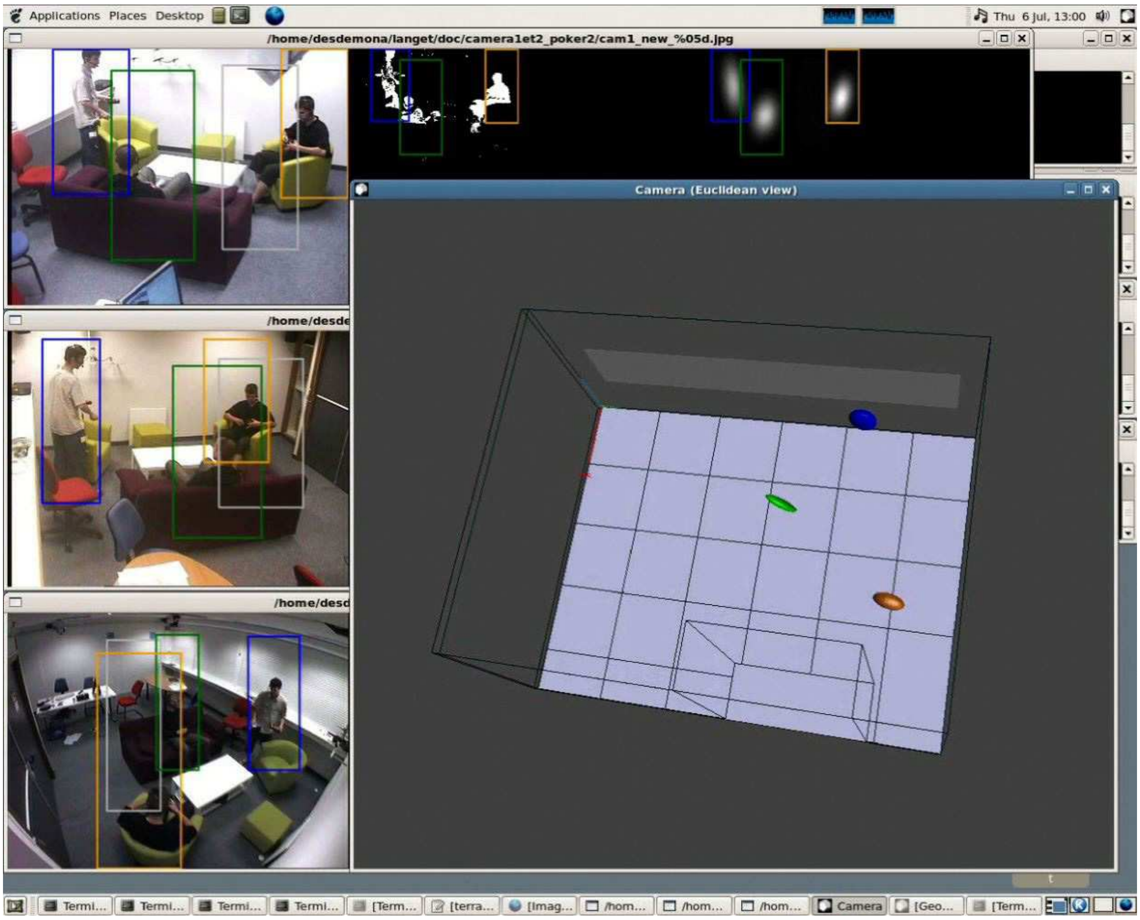


Figure 5. The 3D Bayesian tracker integrates observations from multiple sensors

The INRIA 3D Bayesian body tracker is used to detect, locate and track multiple 3D entities in real time. It is 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.

This perceptual component can be configured to monitor and track the activity within a smart space. The tracker receives its observations from 2D detection process that can use any available pixel level detection algorithm. The tracker currently integrates information from adaptive background subtraction, motion detection, skin color detection, and local appearance using scale normalised Gaussian derivatives. A common scenario is to use the motion to detect and initialise tracking, adaptive background subtraction to track 3D bodies, and skin color to track hands and faces. Cameras may be connected dynamically.

This work is currently supported by ICT Labs thematic actions on Smart Spaces and Smart Energy systems.

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.

5.3. Stereo Viewfinder

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

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

5.4. Tracking Focus of Attention for Large Screen Interaction

Participants: Claudine Combe, John Alexandre Ruiz Hernandez, Varun Jain, James Crowley [correspondant].

Large multi-touch screens may potentially provide a revolution in the way people can interact with information in public spaces. Technologies now exist to allow inexpensive interactive displays to be installed in shopping areas, subways and urban areas. Thesis displays can provide location aware access to information including maps and navigation guidance, information about local businesses and and commercial activities. While location information is an important component of a users context, information about the age and gender of a user, as well as information about the number of users present can greatly enhance the value of such interaction for both the user and for local commerce and other activities.

The objective of this task is to leverage recent technological advances in real time face detection developed for cell phones and mobile computing to provide a low-cost real time visual sensor for observing users of large multi-touch interactive displays installed in public spaces. The initial requirements for this system were expressed by the recent INRIA start-up HiLabs, created in 2008. By the end of 2010, HiLabs had installed over 100 interactive displays in public spaces, mostly in the form of interactive shop windows for travel agents, real-estate agents and banks. HiLabs customers indicated a potential important gain in market if such displays could be made aware of the number, gender and age of users.

The software developed for this activity builds on face detections software that has recently been developed by INRIA for the French OSEO project MinImage. MinImage was a five year, multi-million euro project to develop next generation technologies for integrated digital imaging devices to be used in cellphones, mobile and lap-top computing devices, and digital cameras, that has begun in February of 2007. The project scope included research on new forms of retinas, integrated optics, image formation and embedded image processing. INRIA was responsible for embedded algorithms for real time applications of computer vision.

Within MinImage, INRIA developped embedded image analysis algorithms using image descriptors that are invariant to position, orientation and scale and robust to changes in viewing angle and illumination intensity. INRIA proposed use of a simple hardware circuit to compute a scale invariant Gaussian pyramid as images acquired by the retina. Sums and differences of image samples from the pyramid provide invariant image descriptors that can be used for a wide variety of computer vision applications including detection, tracking and recognition of visual landmarks, physical objects, commercial logos, human bodies and human faces. Detection and tracking of human faces was selected as benchmark test case. This work has been continued with support from EIT ICTlabs, to provide context information for interaction with large multi-touch interactive displays installed in public spaces.

Multitouch interactive displays are increasingly used in outdoor and public spaces. This objective of this task is to provide a visual observation system that can detect and count users of a multitouch display and to estimate information such as the gender, and age category of each user. us rendering the system sensitive to environmental context.

SuiviDeCiblesCouleur locates individuals in a scene for video communications. FaceStabilsationSystem renormalises the position and scale of images to provide a stabilised video stream. SuiviDeCiblesCouleur 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 revised APP is under preperation for new versions of this software for face detection, face tracking, gender and age estimation, and orientation estimation.

6. New Results

6.1. A Lightweight Augmented Virtuality System for Providing a Faithful and Spatially Manipulable Visual Hand Representation

Participants: Sabine Coquillart, Olivier Martin, Andreas Pusch.

We introduced the technical foundations of a system designed to embed a lightweight, faithful and spatially manipulable representation of the user's hand into an otherwise virtual world - Augmented Virtuality (AV). A highly intuitive control during pointing-like near space interaction can be provided to the user, as well as a very flexible means to experimenters, in a variety of contexts. Our approach essentially relies on stereoscopic video see-through Augmented Reality (AR) technology and a generic, extendible framework for managing 3-D visual hand displacements. Research from human-computer interaction, perception and motor control has contributed to the elaboration of our proposal which combines a) acting in co-location, b) avoiding occlusion violations by assuring a correct scene depth ordering and c) providing a convincing visual feedback of the user's hand. This system has already successfully been used in one case and further promising applications are studied [17], [18].

6.2. Effects of Hand Feedback Fidelity on Near Space Pointing Performance and User Acceptance

Participants: Sabine Coquillart, Olivier Martin, Andreas Pusch.

We conducted an experiment to test the effects of different hand representations on near space pointing performance and user preference. Subjects were presented with varying levels of hand realism, including real hand video, a high and a low level 3D hand model and an ordinary 3D pointer arrow. Behavioural data revealed that an abstract hand substitute like a 3D pointer arrow leads to significantly larger position estimation errors in terms of lateral target overshooting when touching virtual surfaces with only visual hand movement constraints. Further, questionnaire results show that a higher fidelity hand is preferred over lower fidelity representations for different aspects of the task [18].

7. Contracts and Grants with Industry

7.1. European and National Projects

7.1.1. FUI 3Dlive

Participants: Frédéric Devernay, Matthieu Volat, Sylvain Duchêne, Vijay Ch. A. V..

3Dlive (http://3dlive-project.com) is a collaborative project, supported by French Ministry of Industry, and involving 3 industry and research clusters: Images & Réseaux (Brittany and Pays-de-la-Loire regions), Imaginove (Rhône-Alpes region), Cap Digital (Paris region).

There are eight partners:

- R&D/industry:
  - Orange Labs (project leader),
  - Technicolor (3D R&D),
  - Thomson Video Networks (encoders),
  - Thales Angenieux (optics).
- Small companies:
  - AMP (TV shooting),
  - Binocle (specific 3D HW & SW manufacturer).
- University labs:
  - INRIA/PRIMA,
  - Institut Telecom.

The objectives of this project are to create expertise in France for the live filming and transmission of 3D stereo contents, and to help French industry and universities to be major global 3D actors.

The role of PRIMA within this project is to develop new algorithms for real-time processing of stereoscopic video streams. This includes:

- stereoscopic video rectification and geometric adjustments.
- view interpolation, and extraction of stereoscopic metadata for the adaptation of the stereoscopic content to the projection screen.

These algorithms rely on view- and scale- invariant feature extraction, feature matching, dense stereoscopic reconstruction, and computer graphics techniques (matting, and accelerated processing and rendering using the GPU).

3Dlive won the *Loading the Future* trophy from the Images & Réseaux cluster in 2011.

7.1.2. OSEO Project MinImage: Embedded Integrated Vision Systems

Start Date: 1 march 2008

Duration: 60 months

The consortium consists of:

- STMicroelectronics
- Saint-Gobain Recherche
- CEA-LETI and LIST
- Varioptic
- INRIA Grenoble Rhone-Alpes Research Centre
- DxO

The goal of the MinImage project is to develop integrated micro-cameras for portable telephones. This is a 141 Million Euro development program provided with 70 Million Euros of Aide by OSEO/AII. The program includes major development efforts in micro-electronics, optics, image processing, and image analysis.

Within the MinImage program, PRIMA has created a fast integer-coefficient O(N) algorithm for computing scale and orientation normalized Gaussian derivatives that is suitable for implementation as a dedicated image processing component within an CMOS integrated vision system. The PRIMA feature extraction engine is currently under evaluation for use in the next generation integrated vision systems for mobile devices sold by ST Microelectronics.

Within MinImage, we have achieved video rate calculation an image pyramid with exactly scale invariant impulse responses using an integer coefficient O(N) algorithm suitable for embedded computer vision. Our software implementation software provides a practical method for obtaining invariant image features from very large retinas for detection, tracking and recognition at video rates. This method is at the core of the real time embedded image description system for mobile applications being developed by ST Microelectronics and the CEA.

John-Alexandre Ruiz-Hernandez has recently demonstrated that the steerable scale invariant Gaussian derivative features outperform the popular "Integral Images" method for face detection using a cascade of linear classifiers popularized by Viola and Jones. We are currently extending these results other applications such as gender recognition, character recognition and place recognition. Key results in this area include

1. Fast, video rate, calculation of scale and orientation for image description with normalized chromatic receptive fields.
2. Real time indexing and recognition using a novel indexing tree to represent multi-dimensional receptive field histograms.
3. Robust visual features for face tracking, bodies, and other objects.

8. Partnerships and Cooperations

8.1. European Initiatives

8.1.1. Major European Organizations with which you have followed Collaborations

EIT KIC ICTlabs

ICTLabs is the KIC for ICT (http://eit.ictlabs.eu/ict-labs/) ICTlabs is set up as a network of 5 "co-location" centers in Helsinki, Stockholm, Berlin, Paris, Eindhoven. The Paris node is run by INRIA with partners Alcatel Lucent, Orange, University Paris Sud and Institut Telecom. PRIMA actively participates in the thematic actions: Smart Spaces, Smart Energy Systems and Health and Well Being.

8.2. International Initiatives

8.2.1. INRIA International Partners

Since the PERSPOS project (BQR Grenoble INP 2008-2009), the MICA center (UMI 2954 CNRS) and PRIMA has a long time collaboration. We wish to develop the concept of "large-scale" perceptive space that is an intelligent environment which will be deployed on a large surface containing several buildings (as a university campus for example). The user is wearing one or many mobile intelligent wireless devices (telephone, Smartphone, PDA, notebook). Using all these devices, one can use many different applications (read emails, browse the Internet, file exchange, etc.). By combining the concepts of large-scale perceptive environments and mobile computing, we can create intelligent spaces to propose services adapted to individuals and their activities. Our collaboration is focussing the user location within such a smart space. Tracking people in smart environments remains a challenging fundamental problem. Whether it is at the scale of a campus, of a building or more simply of a room, we can dynamicaly combine several localization levels (and several technologies) to allow a more accurate and reliable user localization system. This collaboration was concrete with the Ph.D. thesis from Han Yue (started in November 2008). This thesis was co-supervided between Grenoble INP and Hanoi Polytechnical Institute.

8.2.2. Visits of International Scientists

8.2.2.1. Internships

Ch. A. V. Vijay

Subject: Visual fatigue assessment on stereoscopic movies based on image processing: will this 3-D movie give you a headache?
Institution: IIT Bombay (India)

9. Dissemination

9.1. Animation of the scientific community

James L. Crowley Has served as area-chair for the 2011 IEEE International Conference on Robotics and Automation (ICRA2011) in Shanghai China, August 2011.

Sabine Coquillart has served as Program Committee co-chair for the 2011 Joint Virtual Reality Conference of EGVE - EuroVR, Nottingham, Sep. 2011.

Dominique Vaufreydaz has served as co-organizer of *The distinguished lecture series of LIG* (see http://liglab.fr/spip.php?article884) and as organizer of LIG seminars.

Sabine Coquillart has served as

- Panels co-chair for IEEE Virtual Reality 2010, Waltham, Massachusetts, USA, March 2010.
- member of the EUROGRAPHICS Executive Committee reelected for a three year term (2011-2013).
- member of the EUROGRAPHICS Working Group and Workshop board.
- member of the EuroVR Executive Committee.
- member of the Editorial Board of the Journal of Virtual Reality and Broadcasting.
- co-Guest Editor of an issue of Presence: Teleoperators & Virtual Environments Journal to appear in 2011.

Patrick Reignier is a board member of the Association Francaise pour l'Intelligence Artificielle (AFIA).

Frédéric Devernay is a member of the GDR 720 ISIS (Information, Signal, Images et ViSion) executive commitee, responsible for the "Geometry, 3D and motion" theme, and was a member of the organizing committee for the Journees ORASIS 2012 (French workshop for your researchers in computer vision).

9.1.1. Participation on Conference Program Committees

James L. Crowley has served as a member of the program committee for the following conferences and Workshops

- HCI 2011 - IEEE International Workshop on Human-Computer Interaction:?Real-time vision aspects of natural user interfaces, November 7, 2011, Barcelona, Spain
- ICMI-2011 - 13th International Conference on Multimodal Interaction, ICMI 2011, Alicante, Spain, 14-18 November 2011
- ICVS2011 - International Conference on Vision Systems, Sophia Antipolis, 20-22 Sept 2011.
- CVPR2011 - IEEE Computer Vision and Pattern Recognition (CVPR) 2011 Colorado Springs June 20-25, 2011
- FG2011 - the Ninth IEEE International Conference on Automatic Face and Gesture Recognition (FG 2011), Santa Barbara, March 21-25, 2011
- VS2011 - Eleventh International Workshop on Visual Surveillance 2011, Barcelona, Spain, November 13th 2011,

Sabine Coquillart has served as a member of the program committee for the following conferences.

- CGI 2011 - Computer Graphics International - Ottawa, Canada, June 2011.
- CGVCVIP 2011 - IADIS Computer Graphics, Visualization, Computer Vision and Image Processing - Rome, Italy, July 2011.
- GRAPHICON'2011 - International Conference on Computer Graphics and Vision - Moscou, Russie, Sept. 2011.
- GRAPP'2011 - International Conference on Computer Graphics Theory and Applications - Portugal, March 2011.
- WSCG 2011 - International Conferences in Central Europe on Computer Graphics, Visualization and Computer Vision - Prague, Czech Republic, Jan-Feb 2011.
- ICIG 2011 - International Conference on Image and Graphics - Chine, Aug. 2011.
- ISMAR 2011 - IEEE International Symposium on Mixed and Augmented Reality - Basel, Suisse, Oct. 2011.
- WINVR 2011 - ASME World Conference on Innovative Virtual Reality - Milan, Italy, June 2011.
- VRIC 2011 - Haptics for Telepresence, Teleoperation and Collaborative Environments - Laval, France, April, 2011.
- ISVC 2011 - International Symposium on Visual Computing - Las Vegas, USA, Sept. 2011.
- SCCG 2011 - Spring Conference on Computer Graphics - Slovak Republic, April 2011.
- SVR 2011 - Symposium on Virtual and Augmented Reality - Brasil , May 2011.
- ICAT 2011 - International Conference on Artificial Reality and Telexistence - Osaka, Japon, Dec. 2011.
- Siggraph Asia 2011 IPC of Technical Sketches and Posters - Hong-Kong, Dec. 2011.

Sabine Coquillart has served as a member of the best paper award committee for the 2011 Joint Virtual Reality Conference of EGVE

Alexandre Demeure is co-organizer of the EICS 2011 workshop: Enhancing interaction with supplementary Supportive User Interfaces.

Dominique Vaufreydaz has served as a member of the program committee for the following conferences/journals :

- UBICOMM 2011, Fifth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, Lisbon, Portugal, November 2011.
- MTEL2011 at ISM2011, Reviewer and member of the program committee of the Sixth Workshop on Multimedia Technologies for E-Learning in conjunction with the IEEE International Symposium on Multimedia 2011, Dana Point (California), USA, December 2011.
- Special journal extension of the Fifth IEEE International Workshop on Multimedia Technologies for E-Learning.

Frédéric Devernay has served as a member of the program committee for the following conferences and Workshops:

- CVMP 2011 - International Conference on Visual Media Production, London, UK.
- GRETSI 2011 - Traitement du Signal et des Images, Bordeaux, France.
- ICVS2011 - International Conference on Vision Systems, Sophia Antipolis, 20-22 Sept 2011.

Dominique Vaufreydaz was an expert for the *Jeunes Chercheuses et Jeunes Chercheurs* Program from Agence Nationale pour la Recherche (ANR).

Patrick Reignier was an expert for the *Jeunes Chercheuses et Jeunes Chercheurs* Program from Agence Nationale pour la Recherche (ANR).

9.1.2. Participation In International Programs

James Crowley is an expert for the European Commission Future and Emerging Technologies program.

Sabine Coquillart has served on the following advisory Panels

- steering committee for the ICAT conference - International Conference on Artificial Reality and Telexistence.
- Chairman of the steering committee for the EGVE Working Group - EUROGRAPHICS Working group on Virtual Environments.
- Evaluator for proposals submitted for European Commission funding under the framework of FET Open.
- Reviewed proposals for the Austrian Science Fund (FWF).

9.1.2.1. Invited Presentations by James Crowley

James Crowley has given the following invited presentations in 2011

- Embedded View-invariant Visual Detection and Recognition, Grasp Lab seminar, University of Pennsylvania, Philadelphia, 12 December 2011.
- Smart Environments as an Inside-Out Robot, National Seminar for Young Robotics Researchers, JNRR, La Rochelle, 19 Octobre 2011
- Context Aware Systems and Services, Presentation to CNRS INS2I Scientific Council, Paris, 29 september 2011,
- Context Aware Services for Smart Habitats, Summer School on Ambient Intelligence, Lilles, France, 5 July 2011.
- Perception of Human Activity for Smart Environments, Schneider Electric Co. Lecture Series, Grenoble 11 April 2011
- Intelligence dans les Murs : L'invasion progressive de l'Informatique Omniprésent, Grenoble Rotary Club, 28 march 2011
- Context Aware Human Centered Services, invited Presentation to IFIP Working Group 2.7, Grenoble, 27 January 2011.

9.1.2.2. Invited Presentations by Sabine Coquillart

Sabine Coquillart has given the following invited presentations in 2011

- "3D User Interfaces: from Virtual Reality to Pseudo-VR", XEROX Research Center, Meylan, April 2011.
- Panelist at the "Building a Career in Virtual Reality" panel organized by Mark Billinghurst for the 2011 IEEE Virtual Reality Conference, Singapore, March 2011.

9.2. Teaching

Augustin Lux and James Crowley were co-responsable for the Master of Science in Informatics at Grenoble (MoSIG).

Patrick Reignier is responsable for the Licence Pro Développeur Web pour l'Entreprise.

James Crowley:

> Master : Intelligent Systems, M1 year of the Master of Science in Informatics at Grenoble. This course is common with ENSIMAG 2nd year.
>
> Master : Computer Visio, M2 year of the Master of Science in Informatics at Grenoble. This course is common with ENSIMAG 2nd year.

Sabine Coquillart:

Master : Virtual Reality and 3D User Interfaces, M2 GVR (both course and TPs), INPG.

Master : Interactive Systems in Virtual Reality, M2 ARIES, Chalon sur Saône.

Master : Ing$\acute{\eta}$ierie du Virtuel & de l'Innovation, M2 Laval.

Frédéric Devernay:

Licence : Informatics (28h), L1, INPG, France

Master : Algorithmics (48h), M2, ENSIMAG, France.

Patrick Reignier:

Master : Web Programming, Java EE, Programming Langages

Licence : Algorithmics, Web programming

Dominique Vaufreydaz is responsible of several courses from the Economic Faculty of Grenoble:

Licence : Traitement de texte et tableur pour l'économie, 48h eq TD ,L1, Grenoble II, France.

Licence : Informatique appliquée à l'économie et à la gestion, enseignement à distance, Licence, Grenoble II, France.

Licence : Pratique avancée du Tableur, 72 h eq TD, L3, Grenoble II, France.

Licence Professionnelle : Enquêtes et traitement d'enquêtes avec le logiciel Sphinx, 11 h eq TD, Licence pro Métiers de l'Emploi et de la Formation, 11 h eq TD, Grenoble II, France.

Licence Professionnelle : Administration Windows, 39.5 h eq TD, Licence pro Métiers de l'Emploi et de la Formation, Grenoble II, France.

Master : Enquêtes et traitement d'enquêtes avec le logiciel Sphinx, 11 h eq TD, M2 STRATEGIES ECONOMIQUES DU SPORT ET DU TOURISME, Grenoble II, France.

Master : Pratique avancée du Tableur, 22 h eq TD, M1 Économie internationale et stratégies d'acteurs, Grenoble II, France.

Master : Mise à niveau Informatique pour l'économie, 22 h eq TD, M2 DIAGNOSTIC ECONOMIQUE D'ENTREPRISE, Grenoble II, France.

PhD & HdR:

PhD : John Alexandre Ruiz-Hernandez, "Analyse faciale avec dérivées gaussiennes", Univeristy of Grenoble, 23 september 2011, Directed by James Crowley and Augustin Lux [11]

PhD in progress : Han Yue, Localisation d'un utilisateur dans un espace perceptif à grande échelle par analyse multimodale hétérogène, Oct 2008, Directed by Éric Castelli (Laboratoire MICA, Hanoi, Vietnam) and co-directed by Dominique Vaufreydaz

PhD in progress : Remi BARRAQUAND, Apprentissage de Comportement pour l'Interaction Sociale, Oct 2008, James Crowley and Patrick Reignier

PhD in progress : Antoine MELER, Mise En Corresponspondance Rapide, Oct 2008, James Crowley and Augustin Lux

PhD in progress : Mathieu GUILLAME-BERT, Induction des Patternes de Activités, Oct 2009, James Crowley

PhD in progress : Marion Decrouez, Localisation visuelle mobile embarquée sur plateformes de calcul massivement parallèle, date du Oct 2009, Frederic Devernay

PhD in progress : Varun Jain, Perception Visuelle des Emotions, Oct 2011, James Crowley

PhD in progress : Julian Quiroga, Perception de l'Activité Humaine, Oct 2011, Frederic Devernay

PhD in progress : Evanthia Mavridou, Une Nouvelle invariante Visuelle pour la Reconnaissance, Oct 2011, James Crowley and Augustin Lux

10. Bibliography

Major publications by the team in recent years

[1] S. BORKOWSKI, J. LETESSIER, F. BÉRARD, J. CROWLEY. *User centric design of a vision system for interactive applications*, in "2006 IEEE International Conference on Computer Vision Systems, ICVS'06", New York, January 2006, p. 9-16.

[2] S. BORKOWSKI, J. LETESSIER, J. CROWLEY. *Spatial Control of Interactive Surfaces in an Augmented Environment*, in "EHCI '04, Engineering Human Computer Interaction and Interactive Systems", New York, July 2004, p. 228-244.

[3] 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://www-prima.inrialpes.fr/prima/pub/Publications/2009/BCR09/.

[4] 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, p. 47–57, http://www-prima.inrialpes.fr/prima/pub/Publications/2009/BMRC09/.

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

[6] 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.

[7] 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, p. 54-64.

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

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

[10] D. HALL, F. PÉLISSON, O. RIFF, J. CROWLEY. *Brand Identification Using Gaussian Derivative Histograms*, in "Machine Vision and Applications", 2004, vol. 16, n^o 1, p. 41-46.

Publications of the year

Doctoral Dissertations and Habilitation Theses

[11] J. A. RUIZ HERNANDEZ. *Analyse faciale avec dérivées gaussiennes*, Université de Grenoble, September 2011, http://tel.archives-ouvertes.fr/tel-00646718/en/.

Articles in International Peer-Reviewed Journal

[12] A. PUSCH, O. MARTIN, S. COQUILLART. *A Lightweight Augmented Virtuality System for Providing a Faithful and Spatially Manipulable Visual Hand Representation*, in "Studies in Health Technologies and Informatics, Journal of CyberTherapy and Rehabilitation", 2011, vol. 167, p. 170-175.

International Conferences with Proceedings

[13] R. BARRAQUAND, P. REIGNIER, N. MANDRAN. *The Sorceress of Oz*, in "Workshop for Pervasive Intelligibility, part of the Pervasive Conference", San Francisco, USA, July 2011.

[14] F. DEVERNAY, S. DUCHÊNE, A. RAMOS-PEON. *Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis*, in "Stereoscopic Displays and Applications XXII", San Francisco, California, United States, A. J. WOODS, N. S. HOLLIMAN, N. A. DODGSON (editors), SPIE, January 2011, vol. 7863, 786302 [*DOI :* 10.1117/12.872883], http://hal.inria.fr/inria-00565131/en.

[15] D. MASSON, A. DEMEURE, G. CALVARY. *Example Galleries Generated by Interactive Genetic Algorithms*, in "Second International Conference on Creativity and Innovation in Design (DESIRE'11)", 2011, http://dx.doi.org/10.1145/2079216.2079225.

[16] S. PUJADES, F. DEVERNAY. *Stereoscopic Focus Mismatch Monitoring*, in "CVMP - European Conference on Visual Media Production", London, Royaume-Uni, 2011, http://hal.inria.fr/hal-00651843/en/.

[17] A. PUSCH, O. MARTIN, S. COQUILLART. *A Lightweight AV System for Providing a Faithful and Spatially Manipulable Visual Hand Representation*, in "Proceedings of CyberTherapy 16", Gatineau, Canada, June 2011.

[18] A. PUSCH, O. MARTIN, S. COQUILLART. *Effects of Hand Feedback Fidelity on Near Space Pointing Performance and User Acceptance*, in "Proceedings of ISVRI 2011", Singapore, March 2011.

National Conferences with Proceeding

[19] M. DECROUEZ, R. DUPONT, F. GASPARD, F. DEVERNAY, J. CROWLEY. *Modélisation explicite des objets et de l'environnement en combinant les approches topologique et métrique pour la localisation*, in "ORASIS - Congrès des jeunes chercheurs en vision par ordinateur", Praz-sur-Arly, France, INRIA Grenoble Rhône-Alpes, 2011, http://hal.inria.fr/hal-00651823/en/.

[20] M. DECROUEZ, R. DUPONT, F. GASPARD, F. DEVERNAY, J. CROWLEY. *Détection et apprentissage automatique d'objets pour la modélisation de milieux*, in "RFIA - Congrès francophone sur la Reconnaissance des Formes et l'Intelligence Artificielle", Lyon, France, 2012, http://hal.inria.fr/hal-00651830/en/.

[21] A. DEMEURE, D. MASSON, G. CALVARY. *Graphs of models for exploring design spaces in the engineering of Human Computer Interaction*, in "Proceeding of the 2nd SEMAIS workshop of the IUI 2011 conference", Springer HCI, 2011, http://iihm.imag.fr/dmasson/publications/2011/Semais11-Demeure-GraphOfModels.pdf.

[22] D. MASSON, A. DEMEURE, G. CALVARY. *Inspirer, explorer: Magellan, un environnement interactif évolutionnaire pour la génération d'exemples*, in "23rd French Speaking Conference on Human-Computer Interaction", New York, NY, USA, IHM '11, ACM, 2011, p. 22:1–22:7, http://doi.acm.org/10.1145/2044354.2044382.

Scientific Books (or Scientific Book chapters)

[23] S. ZAIDENBERG, P. REIGNIER. *Reinforcement Learning of User Preferences for a Ubiquitous Personal Assistant*, in "Advances in Reinforcement Learning", Intech, January 2011, n^{o} ISBN 978-953-307-369-9, 470.

Research Reports

[24] N. BALACHEFF, B. BUCCIO, P. CHAPUIS, J. COUTIN, J. COUTAZ, J. CROWLEY, Y. DENNEULIN, L. DU BOUSQUET, A. DUDA, R. ECHAHED, M.-C. FAUVET, C. GARBAY, E. GAUSSIER, I. GUILLET, C. LAUGIER, Y. LEDRU, A. LEGRAND, N. MANDRAN, H. MARTIN, J.-F. MEHAUT, T. MORTURIER, B. PLATEAU, E. PONS, J. PREVOST, F. PROST, V. QUINT, P. REIGNIER, F. ROUSSEAU, M.-C. ROUSSET, E. RUTTEN, M. VACHER. *The First Four Years (2007-2010) and Beyond, Volume 1: Research Program and Activity Report*, LIG, Grenoble, France, 2011, n^{o} RR-LIG-015, http://rr.liglab.fr/research_report/RR-LIG-015.pdf.

References in notes

[25] *DNS-SD : DNS Service Discovery*, 2000, http://www.dns-sd.org.

[26] S. BORKOWSKI, O. RIFF, J. CROWLEY. *Projecting Rectified Images in an Augmented Environment*, in "PROCAMS'03 Workshop", 2003.

[27] F. BÉRARD. *The magic table: computer-vision based augmentation of a whiteboard for creative meetings*, in "Workshop on Projector-Camera Systems", 2003.

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

[29] J. COUTAZ, E. FONTAINE, N. MANDRAN, A. DEMEURE. *DisQo : A user needs analysis method for smart home*, in "Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries (NordiCHI '10)", ACM, 2010, p. 615-618 [*DOI :* 10.1145/1868914.1868987], http://iihm.imag.fr/publication/CFM+10b/.

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

[31] J. CROWLEY, P. REIGNIER. *An Architecture for Context Aware Observation of Human Activity*, in "Workshop on Computer Vision System Control Architectures (VSCA 2003)", 2003.

[32] 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, p. 584–598.

[33] S. DAVIDOFF, M. K. LEE, C. YIU, J. ZIMMERMAN, A. K. DEY. *Principles of Smart Home Control*, in "Ubicomp", P. DOURISH, A. FRIDAY (editors), Lecture Notes in Computer Science, Springer, 2006, vol. 4206, p. 19-34, http://dblp.uni-trier.de/db/conf/huc/ubicomp2006.html.

[34] F. DEVERNAY, P. BEARDSLEY. *Stereoscopic Cinema*, in "Image and Geometry Processing for 3-D Cinematography", R. RONFARD, G. TAUBIN (editors), Geometry and Computing, Springer Berlin Heidelberg, 2010, vol. 5, p. 11-51 [*DOI :* 10.1007/978-3-642-12392-4], http://hal.inria.fr/inria-00509217/en.

[35] F. DEVERNAY, S. DUCHÊNE. *New View Synthesis for Stereo Cinema by Hybrid Disparity Remapping*, in "International Conference on Image Processing (ICIP)", Hong-Kong Hong Kong, Sep 2010, http://hal.inria.fr/inria-00531369/en.

[36] F. DEVERNAY, A. RAMOS-PEON. *Novel View Synthesis for Stereoscopic Cinema: Detecting and Removing Artifacts*, in "3DVP 2010: ACM Workshop on 3D Video Processing", Italie Firenze, Oct 2010, http://hal.inria.fr/inria-00531364/en.

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

[38] D. GARLAN, S. CHENG, A. HUANG, B. SCHMERL, P. STEENKISTE. *Rainbow: Architecture-based, self-adaptation with reusable infrastructure*, in "IEEE Computer", October 2004, p. 46-54.

[39] 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, p. 617-622.

[40] 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.

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

[42] P. HORN. *Autonomic Computing; IBM's Perspective on the state of information technology*, October 2001, http://researchweb.watson.ibm.com/autonomic.

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

[44] J. KEPHART, D. CHESS. *The vision of autonomic computing*, in "IEEE Computer", 2003, vol. 36(1), p. 41–50.

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

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

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

[48] S. A. MEDNICK. *The associative basis of the creative process*, in "Psychological Review", 1962, p. 220–232.

[49] A. NEGRE, C. BRAILLON, J. CROWLEY. *Visual navigation from variation of intrinsic scale*, in "under review", 2006.

[50] G. PINGALI, C. S. PINHANEZ, A. LEVAS, R. KJELDSEN. *Steerable Interfaces for pervasive computing spaces*, in "IEEE PerCom", March 2003.

[51] C. S. PINHANEZ. *The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces*, in "Proceedings of Ubiquitous Computing 2001 Conference", September 2001.

[52] V. POLADIAN. *Dynamic Configuration of Resource-aware services*, in "Int. Conf. on Software Engineering", September 2001.

[53] F. PÉLISSON, D. HALL, O. RIFF, J. CROWLEY. *Brand identification using Gaussian derivative histograms*, in "International Conference on Vision Systems", Graz, Austria, April 2003, p. 492–501.

[54] R. RASKAR, G. WELCH, M. CUTTS, A. LAKE, L. STESIN, H. FUCHS. *The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays*, in "Proceedings of the ACM SIGGRAPH'98 Conference", 1998.

[55] 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, p. 31–50.

[56] N. TAKAO, J. SHI, S. BAKER. *Tele Graffiti: a camera projector based remote sketching system with hand based user interface and automatic session summarization*, in "Int. Journal of Computer Vision", 2003, vol. 53, p. 115-133.

[57] T. T. H. TRAN, A. LUX. *A method for ridge extraction*, in "Asian Conference on Computer Vision", Jeju, Korea, 2004, p. 960–966.

[58] F. VERNIER, N. LESH, C. SHEN. *Visualization Techniques for Circular Tabletop Interfaces*, in "Advanced Visual Interfaces", 2002.

[59] P. WELLNER. *The DigitalDesk Calculator: Tactile Manipulation on a Desk Top Display*, in "ACM Symposium on User Interface Software and Technology (UIST '91)", November 1991, p. 27–33.