

*Project-Team Orion**Intelligent Environments for Problem
Solving by Autonomous Systems**Sophia Antipolis*

THEME 3A

Activity
Report

2003

Table of contents

1. Team	1
2. Overall Objectives	1
2.1. Presentation	1
2.1.1. Research Themes	1
2.1.2. International and Industrial Cooperation	2
3. Scientific Foundations	2
3.1. Introduction	2
3.2. Program Supervision	2
3.3. Software Platform for KBS Design	4
3.4. Automatic Image Interpretation	6
4. Application Domains	7
4.1. Overview	7
4.2. Astronomic Imagery	8
4.3. Videosurveillance	8
4.4. Pollen Grain Recognition	9
4.5. Early Detection and Plant diseases Recognition	9
4.6. Synchronous Model Development	10
5. Software	10
5.1. Ocapi	10
5.2. Pegase	10
5.3. VSIP	10
6. New Results	11
6.1. Program Supervision	11
6.1.1. Program Supervision and Model Calibration	11
6.1.2. New Scheduling Engine	12
6.1.3. Distributed Program Supervision	13
6.2. Software Platform for KBS Development	13
6.2.1. Component Framework Verification	14
6.3. Automatic Interpretation of Image Sequences	14
6.3.1. Update of the Reference Image	15
6.3.2. People Detection	15
6.3.3. Human Posture Recognition	16
6.3.4. Classification of Lateral Forms	18
6.3.5. Temporal Scenario Recognition for Automatic Video Interpretation	19
6.3.6. Video Understanding Evaluation	21
6.3.7. Towards an Automatic Evaluation and Repair Framework for Video Interpretation	22
6.4. Complex Object Recognition	24
6.4.1. A Cognitive Vision Platform for Natural Complex Object Recognition	24
6.4.2. Ontology-Based Knowledge Acquisition and Visual Concept Learning	25
6.4.3. Image Formation of a 3D Translucent Microscopic Object (Application to Light Microscopy)	25
7. Contracts and Grants with Industry	26
7.1. Industrial Contracts	26
8. Other Grants and Activities	27
8.1. European projects	27
8.1.1. ECVision European Research Network	27
8.1.2. ADVISOR Project	27

9. Dissemination	27
9.1. Scientific Community	27
9.2. Teaching	28
9.3. Thesis	28
9.4. Participation to Conferences, Seminars, and Invitations	28
10. Bibliography	29

1. Team

Scientific leader

Monique Thonnat [DR Inria]

Administrative Assistant

Catherine Martin

Inria Researchers

François Brémond [CR Inria]

Sabine Moisan [CR Inria]

Annie Ressouche [CR Inria]

Research Engineers

Alberto Avanzi [Bull Engineer]

Binh Bui

Julien Canet [Fellow Engineer, arrived on 1er september 2003]

Frédéric Cupillard

Magali Mazière

Christophe Tornieri

Ph D Students

Boulay Bernard [INRIA Grant, Nice Sophia-Antipolis University, since october 2003]

Céline Hudelot [Regional Grant, Nice Sophia-Antipolis University]

Benoit Georis [Louvain Catholic University (UCL) Belgium]

Nicolas Maillot [INRIA Grant, Nice Sophia-Antipolis University]

Thinh Van Vu [INRIA Grant, Nice Sophia-Antipolis University]

Jean-Philippe Vidal [CEMAGREF Grant, Toulouse University]

Associates

Jean-Paul Rigault [Professor, Nice Sophia-Antipolis University]

Visitors

Tony Krakenes [Norwegian Defence Research Establishment, 6 months (January-July) 2003]

Undergraduates

Anne-Laure Barjon [ESSI UNSA, october 2002-april 2003]

Bernard Boulay [DEA,ESSI UNSA, april-september 2003]

Thi-Thanh-Tu Bui [DEPA,IFI Hanoi, april-september 2003]

Florent Fusier [DESS Bordeaux II University, april-december 2003]

Naoufel Khayati [ENSI Tunis, september-december 2003]

Vincent Rampal [ESSI UNSA, october 2002-april 2003]

2. Overall Objectives

2.1. Presentation

Orion is a multi-disciplinary team at the frontier of computer vision, knowledge-based systems(KBS), and software engineering.

The Orion team is interested in research on **intelligent reusable systems** and **cognitive vision**.

2.1.1. Research Themes

More precisely, our research themes deal with the design of intelligent systems based on knowledge representation, learning and reasoning techniques.

We study two levels of reuse: the reuse of programs and the reuse of tools for knowledge-based system design. We propose an original approach based on **program supervision** techniques which allow to plan

modules (or programs) and to control their execution. Our researches concern the representation of knowledge about the programs and their use as well as planning techniques. Moreover, relying on state of the art practices in software engineering and in object-oriented languages we propose a platform to facilitate the construction of knowledge-based systems.

In cognitive vision we study two kinds of **automatic image understanding**: video sequence understanding and complex object recognition. Our researches thus concern the representation of the knowledge about objects, of events and of scenarios to be recognised, as well as the reasoning processes useful for image understanding, like categorization for object recognition.

2.1.2. International and Industrial Cooperation

- Participation to an industrial project, CASSIOPEE, which aims at developing an automatic video surveillance platform in a bank context. This project involves a bank, a video system integrator, a telesurveillance operator (Bank Corporation, Eurotelis and Ciel), and INRIA.
- Participation to the European IST project ADVISOR Annotated Digital Video for Intelligent Surveillance and Optimised Retrieval with Racal Research (UK), Bull (France), The University of Reading (UK), King's College (UK) and Vigitec (Belgium).
- Contract with RATP (in video interpretation) for passenger detection and classification in real time.
- Participation to projects with ALSTOM and SNCF, both related to train visual surveillance.
- Cooperation with ENSI, GRIFT/ASI of Tunis (Tunisia) in the framework of Franco-Tunisian cooperations.
- Cooperation with the American ARDA network on video events to define standard language and ontology for video events.

3. Scientific Foundations

3.1. Introduction

In order to facilitate the design of KBS, we design engines, independent of specific applications, but yet *dedicated* to a particular task. The tasks that we study are program supervision and image understanding.

Developing dedicated tools allows us to provide systems that are adapted to express the necessary knowledge and that can be used in a wide range of application domains.

To design such engines, it is necessary to rely on models of both knowledge and reasoning mechanisms (problem solving methods) that are involved in the tasks we study.

3.2. Program Supervision

Key words: *program supervision, planning, program reuse.*

Participants: Sabine Moisan, Monique Thonnat.

Glossary

Program supervision aims at automating the reuse of complex software (for instance image processing program library), by offering original techniques to plan and control processing activities.

Knowledge-based systems are well adapted for the program supervision research domain. Indeed, these techniques achieve the twofold objective of program supervision: both to favour the capitalisation of knowledge about the use of complex programs and to operationalise this utilisation for users not specialised in the domain. We study the problem of modelling knowledge specific to program supervision, in order to define on the one hand, knowledge description languages and knowledge verification facilities for experts and, on the other hand, tools (e.g., inference engines) to operationalise program supervision knowledge into software systems dedicated to program supervision. To implement different program supervision systems, we have developed a generic and customisable framework: the LAMA platform [5], which is devoted both to knowledge base and inference engine design.

Program supervision aims at automating the reuse of complex software (for instance image processing library programs). To this end we propose original techniques to plan and control processing activities. Most of the work that can be found in the literature about program supervision is generally motivated by application domain needs (for instance, image processing, signal processing, or scientific computing). Our approach relies on KBS techniques. A knowledge-based program supervision system emulates the strategy of an expert in the use of the programs. It typically breaks up into:

- a library of executable programs in a particular application domain (e.g., medical image processing),
- a knowledge base for this particular domain, that encapsulates expertise on programs and processing; this primarily includes descriptions of the programs and of their arguments, but also expertise on how to perform automatically different actions, such as initialisation of program parameters, assessment of program execution results, etc.
- a general supervision engine, that use the knowledge stored in the knowledge base for effective selection, planning, execution and control of execution of the programs in different working environments.
- interfaces that are provided to users to express initial processing requests and to experts to browse and modify a knowledge base, as well as to trace an execution of a knowledge-based system.

Program supervision is a very general problem, and program supervision techniques may be applied to any domain where complex processing is necessary and where each sub-processing corresponds to a suitable chain of several basic programs. To tackle this generality, we provide both knowledge models and software tools. We want them to be both general (*i.e.* independent of any application and of any library of programs) and flexible, which means that the lack of certain type of knowledge has to be compensated by powerful control mechanisms, like sophisticated repair mechanisms.

Program Supervision Model

To better understand the general characteristics of the program supervision activity and to improve the (re)use of existing programs, the knowledge involved in this activity has to be modelled independently of any application. The knowledge model defines the structure of program descriptions and what issues play a role in the composition of a solution using the programs. It is thus a guideline for representing reusable programs. We have thus used knowledge modelling techniques to design an explicit description of program supervision knowledge to allow the necessary expertise to be captured and stored for supporting of a novice user or an autonomous system performing program supervision. We have modelled concepts and mechanisms of program supervision first for the OCAPAPI [41] engine, and then for our more recent engines. A preliminary work with KADS expertise model has been improved by using recent component reuse techniques (from Software Engineering area), planning techniques (from Artificial Intelligence area), existing program supervision systems, and our practical experience in various applications such as obstacle detection in road scenes, medical imaging, or galaxy identification.

Knowledge Base Description Language In the LAMA platform we have developed the YAKL language that allows experts to describe all the different types of knowledge involved in program supervision, independently of any application domain, of any program library, or of the implementation language of the knowledge-based system (in our case Lisp or C++).

The objective of YAKL is to provide a concrete means to capitalise in a both formal (system-readable) and readable (user-readable) form the necessary skills for the optimal use of programs, for user assistance, documentation, and knowledge management about programs. First, a readable syntax facilitates communication among people (*e.g.*, for documenting programs) and, second, a formal language facilitates the translation of abstract concepts into computer structures that can be managed by software tools.

YAKL is used both as a common storage format for knowledge bases and as a human readable format for writing and consulting knowledge bases. YAKL descriptions can be checked for consistency, and eventually translated into operational code. YAKL is an open extensible language which provides experts with a user-friendly syntax and a well defined semantics for the concepts in our model.

3.3. Software Platform for KBS Design

Key words: *frameworks, library, component reuse, software engineering.*

Participants: Sabine Moisan, Annie Ressouche, Jean-Paul Rigault.

Glossary

The LAMA software platform provides a unified environment to design not only knowledge bases, but also inference engines variants, and additional tools. It offers toolkits to build and to adapt all the software elements that compose a KBS.

The LAMA software platform allows to reuse all the software elements that are necessary to design knowledge-based systems (inference engines, interfaces, knowledge base description languages, etc.). It gathers several toolkits to build and to adapt all these software elements. The platform both allows to design program supervision and automatic image interpretation KBSs, and it facilitates the coupling between the two types of KBSs.

Designing dedicated tools for a particular task (such as program supervision) has two advantages: on the one hand to focus the knowledge models used by the tools on the particular needs of the task, and, on the other hand to provide unified formalisms, common to all knowledge bases belonging to the same task. We want to go one step further in order to facilitate also the *reuse of elements that compose a knowledge-based system* (inference engines, interfaces, knowledge base description languages, etc.). That is why we decided to design the generic and adaptable LAMA platform [5]. Such a platform allows us to tackle the problem of adapting a task like program supervision (as well as planning or classification) and tune it to specific domain requirements. LAMA provides both experts and designers task-oriented tools, i.e. tools that integrate a model of the task to perform, that would help reduce designer efforts and situate them at an appropriate level of abstraction. The platform thus provides a unified environment to design not only expert knowledge bases, but also variants of inference engines, and additional tools. It gathers several toolkits to build and to adapt all these software elements.

LAMA relies on recent techniques in software engineering. It is an object-oriented extensible and portable software platform that implements the program supervision model and provides “computational building blocks” (toolkits) to design dedicated tools. The toolkits are complementary but independent, so it is possible to modify, or even add or remove a tool without modifying the rest. Another objective of the platform is to couple KBSs performing different complementary tasks in a unified environment.

We have used LAMA to design different program supervision engines and variants of them. The platform has substantially simplified the creation of these engines, compared to the amount of work that had been necessary for the previous implementation of OCAPI.

The core of the platform (see figure 1) is a *framework* of re-usable components, called BLOCKS, it provides designers with a software framework (in the sense of software engineering). For instance the program supervision part of the framework offers reusable and adaptable components that implement generic data structures and methods for supporting a program supervision system. It also supplies a task ontology to construct a knowledge base. Additional toolkits are also provided in the platform: a toolkit to design knowledge

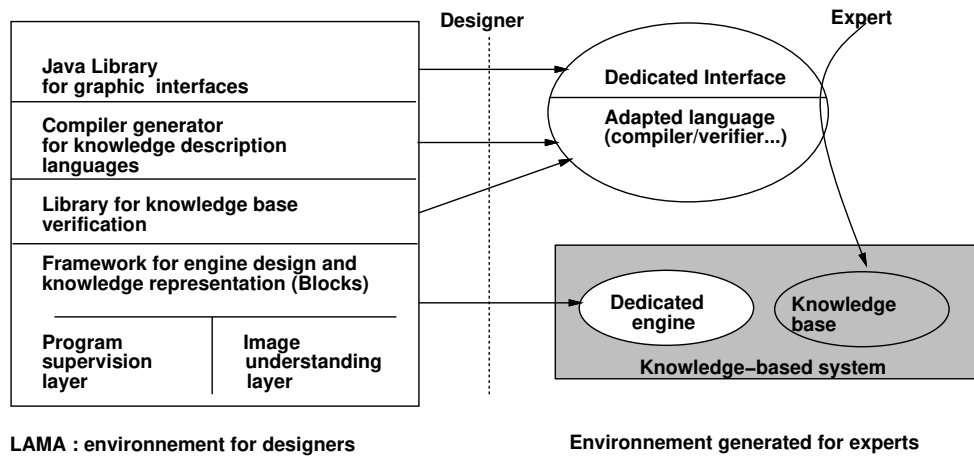


Figure 1. LAMA architecture and tools for engine design, knowledge base description, verification, and visualisation

base editors and parsers –to support the dedicated description languages, a knowledge verification toolkit –adapted to the engine in use–, a toolkit to develop graphical interfaces –both to visualize the contents of a knowledge base and to run the solving of a problem. The most important toolkits are briefly described below.

Framework for Engine Design

BLOCKS (Basic Library Of Components for Knowledge-based Systems) is a framework (in the software engineering sense), that offers reusable and adaptable components implementing generic data structures and methods for the design of knowledge-based systems engines. The objective of BLOCKS is to help designers create new engines and reuse or modify existing ones without extensive code rewriting.

The components of BLOCKS stand at a higher level of abstraction than programming language usual constructs. BLOCKS thus provides an innovative way to design engines. It allows engine designers to speed-up the development (or adaptation) of problem solving methods by sharing common tools and components. Adaptation is often necessary because of evolving domain requirements or constraints.

Using BLOCKS, designers can conveniently compose engines (in other words of problem solving methods) by means of basic reasoning components. They can also test, compare or modify different engines in a unified framework. Moreover, this platform allows the reuse of (parts of) existing engines.

This approach allows as well a unified vision of various engines and supplies convenient comparisons between them.

Engine Verification Toolkit

From a software engineering point of view, in order to ensure a safe reuse of BLOCKS components, we are working on a verification toolkit for engine behavior, which relies on *model-checking* techniques.

We propose a mathematical model and a formal language to describe the knowledge about engine behaviors. Associated tools may ensure correct and safe reuse of components, as well as automatic simulation and verification, code generation, and run-time checks.

Knowledge Base Verification Toolkit

Knowledge-based systems (KBS) require a safe building methodology to ensure a good quality. This quality control can be difficult to introduce into the development process due to its unstructured nature. The usual verification methods focus on syntactic verification based on formalisms that represent the knowledge (knowledge representation schemes, like rules or frames) .

Our aim is to provide tools to help experts during the construction of knowledge bases, in order to guarantee a certain degree of reliability in the final system. For this purpose we can rely not only on the knowledge representation schemes (frames and rules), but also on the underlying model of the task that is implemented in the KBS.

The toolkit for verification of knowledge bases is composed of a set of functions to perform knowledge verifications. These verifications are based on the properties of the modes of representation of the knowledge used in the KBSs (frames and rules), but it can be adapted to check the role which the various pieces of knowledge play in the task at hand. Our purpose is not only to verify the consistency and the completeness of the base, but also to verify the adequacy of the knowledge with regard to the way an engine is going to use it.

Graphic Interface Framework Interfaces are an important part of a knowledge-based system. The graphic interface framework is a Java library that follows the same idea as BLOCKS. It allows to customize interfaces for designing and editing knowledge bases, according to the engine.

3.4. Automatic Image Interpretation

Participants: Francois Brémond, Monique Thonnat.

Key words: *image interpretation, pattern recognition, scenario recognition, image sequences.*

Glossary

Automatic Image Interpretation consists in extracting the semantics of data depending on a predefined model. This is a specific part of the perception process : automatic interpretation of results coming from the image processing level.

Automatic image interpretation is a difficult problem which is the basis of many research activities in both computer vision and artificial intelligence. The difficulty of the interpretation task first comes from the type of the entities to be recognized. It is easier to recognize manufactured rigid objects than the behaviour of several natural objects in a dynamic context. The difficulty also depends on the type of interpretation to be performed. The problem can be a simple labeling of an entity detected in an image associated with a model or the detection and the consistency checking (e.g. spatial, temporal, structural) of an entity set.

The Orion team aims at the automatic interpretation of spatial and/or temporal images. Interpretation results can be object categorization but also event, situation or scenario recognition. The approach is composed of two steps. (1) An image processing step which aims at extracting entities of interest for interpretation. (2) The analysis of the extracted entities which can be selected for object categorization or for behaviour recognition.

The difficulty of the problem is that two types of knowledge are required. First, primitive and descriptor extraction relies on the execution of image processing programs and requires knowledge on vision algorithms. Second, the interpretation task requires expert knowledge of the application domain related to the entities to be recognized or analysed. Automatic image processing execution is related to program supervision and is a research activity by itself (cf. module 3.2).

The next points describe the proposed approaches for complex object recognition and for image sequence analysis.

Complex Object Recognition :

Complex object recognition refers to recognize non-geometric objects from abstract semantic models. In a first stage, image processing techniques are used to detect regions of interest so as to compute numerical descriptors. In a second stage, extracted descriptors are used by the interpretation system to classify the object in a predefined taxonomy of classes which define the semantic models. Three recursive tasks are involved in the classification process : data abstraction, class matching, recognition refinement. During the classification process, more information may have to be computed from the image. The operationalization of such systems requires an important work for the design of knowledge bases and for the implementation of image processing programs.

Interpretation of Image Sequences :

The interpretation of image sequences refers to giving a semantic explaining the human activities depicted from image streams provided by color, monocular and fixed cameras. The general base of the scene interpretation algorithm is based on the a priori knowledge (containing the scene model and a library of scenario models) and on the cooperation of the 4 following modules (cf figure 2): 1) mobile object detection and frame to frame tracking, 2) multi-cameras combination, 3) long term tracking, and 4) behaviour recognition. The first module is implemented as three instances, one foreach camera. It detects the mobile objects evolving in the scene and tracked them on 2 consecutive images. The second one combines the detected mobile objects from several cameras. This module is optional in the case of one camera. The third module tracks the mobile objects on a long term basis using model of the expected objects to be tracked. The last module consists, thanks to artificial intelligence techniques, in identifying the tracked objects and in recognizing their behaviour by matching them with predefined models of one or several scenarios.

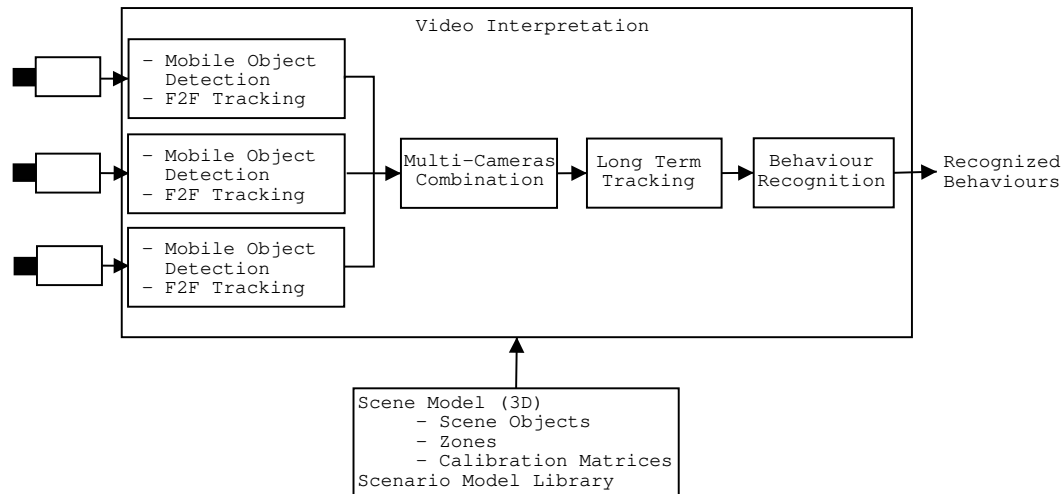


Figure 2. Algorithm of the interpretation of image sequences.

4. Application Domains

4.1. Overview

Key words: *astronomy, visualsurveillance, transportation, environnement, health, pollen, multimedia, bio computer science, hydraulics, safety properties.*

Applications achieved in the Orion team are useful to validate both our research directions and ours models. We are mainly involved in the following applicative domains : astronomy, health, environment, videosurveillance and transportation. Our approach supports two goals. A first scientific goal is to bring new technics in other domains. For instance, in astronomy for the automatic classification of galaxies. A second goal concerns industrial issues in order to develop operational systems as for instance intelligent visualsurveillance of underground stations. Besides imagery is the main focus of the team, we have also developed applications in numerical calculus domain such as program supervision for numerical simulation. For instance, in order to model physical processes related to river hydraulics people use simulation codes which are based on the discretisation of the simplified fluid mechanics equations (de Saint-Venant equations) that model streamflows. This task, called model calibration, is close to program supervision and it has a predominant role in good modelling practice in hydraulics and in water-related domains.

Moreover, our theoretical approach in the software engineering field may be applied in a more general context: as a consequence, the theory we developed to enforce safety properties of software tools we developed, can be applied to critical system verification.

4.2. Astronomic Imagery

The complete automation of galaxy description and classification with respect to their morphological type based on images is an historic application in our team [53][50]. We are expert in this domain both concerning the image processing of galaxies field and theoretical models of morphological classification. This application is a reference to validate our models and software related to interpretation for complex objects understanding and to program supervision [51][55].

4.3. Videosurveillance

In the domain of videosurveillance, the growing feeling of insecurity among the population led the private companies and also the public authorities to deploy security systems in order to protect their equipment or their commercial interests. For the safety of the public places, the video camera based surveillance techniques are more and more used, but the multiplication of the camera number leads to the saturation of transmission and analysis means (it is difficult to supervise simultaneously hundreds of screens). For example, 1000 cameras are now used for monitoring the subway network of Brussels. In the framework of our works on automatic video interpretation, we have studied since 1994 the conception of an automatic platform which can assist the videosurveillance operators.

The aim of this platform is to act as a filter, sorting the scenes which can be interesting for a human operator. This platform is based on the cooperation between an image processing component and an interpretation component using artificial intelligent techniques. Thanks to this cooperation, this platform has to automatically recognize different scenarios of interest in order to alert the operators. These works have been realised with academic and industrial partners, like European projects Esprit Passwords, AVS-PV and AVS-RTPW and more recently, European project ADVISOR and industrial projects RATP, Cassiopee, ALSTOM and SNCF. A first set of very simple applications for the indoor night surveillance of supermarket (AUCHAN) showed the feasibility of this approach. A second range of applications has been investigated and it corresponds to the parking monitoring where the rather large viewing angle makes it possible to see many different objects (car, pedestrian, trolley) in a changing environment (illumination, parked cars, trees shaken by the wind, etc.). This set of applications allowed us to test various methods of tracking, trajectory analysis and recognition of typical cases (occultation, creation and separation of groups, etc).

Since 1997, we have studied and developed videosurveillance techniques in the transport domain which requires the analysis and the recognition of groups of persons observed from lateral and low position viewing angle in subway stations (subways of Nuremberg, Brussels, Charleroi and Barcelona). More recently, we work in cooperation with Bull company in the Dyade Telescope action, on the conception of a video surveillance intelligent platform which is independent of a particular application. The principal constraints are the use of fixed cameras and the possibility to specify the scenarios to be recognised, which depend on the particular application, based on scenario models which are independent from the recognition system. The collaboration with Bull has been continued through the European project ADVISOR until March, 2003. Also, we experimented in the framework of a national cooperation, the application of our video interpretation techniques to the problem of the media based-communication. In this case, the scene interpretation is a way to decide which information has to be transmitted by a multimedia interface.

In parallel of the videosurveillance of subway stations, since 2000, new projects based on the video understanding platform have started for new applications, like bank agency monitoring and train car surveillance. The

new challenge in bank agency monitoring is to handle a cluttered environment and in train car surveillance is to take into account the motion of the cameras.

4.4. Pollen Grain Recognition

A part of Orion activities related to healthcare and environment are dedicated to automatic pollen grain recognition. We aim at providing tools with the palynologist so that they can process large amounts of data in a short time. For that purpose we use complex object recognition techniques which rely on image processing, knowledge based systems and pattern recognition.

The aim is to quantify the correlation between the environmental stress (so-called envi-contamination factor that is a combination of the concentration of allergens, the concentration of atmospheric pollutants including ozone and black dusts), and some indicators of the population health (medical data, hospitalisation statistics, school and work absenteeism, medicine consumption). The task of the palynologist technician is to recognise the pollen particles present on a microscope slide, to give every pollen a name (family, genus, specie, group) and to finally produce a pollen spectrum for the given day. Not only because of the time required to obtain the pollen measurements from the sensor samples but also because possible human errors of counting and identifying the pollen grains can occur, it is of major interest to develop a system capable to recognise the pollen grains and to count them per types, this means to make possible an automatic evaluation of the atmospheric pollen concentration.

In this context, two main directions are studied : global counting of the number of pollen grains found on a slice and individual recognition of each pollen grain found on a slice. The second approach gives the accurate quantity of each type of pollen grain. Automatic global counting has been studied by using image processing techniques ([45] and [54]). Nevertheless, individual recognition remains of great interest. That is why the Orion team has been studying this problem since 1996 [48].

Due to the complexity of the different types of pollen grains, palynologist knowledge is taken into account [23]. That is why a cooperative approach between image processing techniques and artificial intelligence techniques is proposed. Image processing techniques allow to extract and analyse relevant regions in images. Knowledge based systems use taxonomic knowledge to perform interpretation. They are also used for the supervision of image processing operators. It can be noted that similar approaches have been used in the past for galaxy, zooplankton and foraminifer recognition [52][49][46][47].

The European project ASTHMA started in 1998 and finished in 2001. One of the goals of this project was to provide near real time accurate information on aeroallergens and air quality to the sensitive users. During ASTHMA, the Orion team was in charge of the conception and the study of a platform dedicated to the recognition of 3D pollen grain images (cf. module 6.4).

4.5. Early Detection and Plant diseases Recognition

In the Environment domain, Orion is interested in the automation of the early detection of plant diseases. The goal is to detect, to identify and to accurately quantify the first symptoms of diseases or pest initial presence. As plant health monitoring is still carried out by humans, the plant diagnosis is limited by the human visual capabilities whereas most of the first symptoms are microscopic. Due to the visual nature of the plant monitoring task, computer vision techniques seem to be well adapted. We make use of complex object recognition methods including image processing, pattern recognition, scene analysis, knowledge based systems. Our work takes place in a large-scale and multidisciplinary research program (IPC: Integrated Crop Production) ultimately aimed at reducing pesticidal application. We focus on the early detection of powdery mildew on greenhouse rose trees. Powdery mildew has been identified by the Chamber of Agriculture as a major issue in ornamental crop production. As the proposed methods are generic, the expected results concern all the horticultural network.

Objects of interest can be fungi or insects. Fungi appear as thin networks more or less developed and insects have various shapes and appearances. We have to deal with two main problems: the detection of the objects and their semantic interpretation for an accurate diagnosis. In our case, due to the various and complex

structures of the vegetal support and to the complexity of the objects themselves, a purely bottom up analysis is insufficient and explicit biological knowledge must be used. Moreover, to make the system generic, the system has to process images in an intelligent way, i.e. to be able to adapt itself to different image processing requests and image contexts (different sensors, different acquisition conditions). We proposed a generic cognitive vision platform based on the cooperation of three knowledge based systems.

This work takes part in a two year research agreement between the Orion team and INRA (Institut National de Recherche Agronomique) started in November 2002. This research agreement continues the COLOR (COoperation LOcale de Recherche) HORTICOL started in september 2000 (see also 6.4.1).

4.6. Synchronous Model Development

In collaboration with V. Roy (CMA Ecole des Mines de Paris) and J-Y Tigli (CNRS-UNSA) we develop a specific synchronous/asynchronous architecture dedicated to critical system specification. The motivation of this work is to introduce safety property checking of critical features in this domain and our approach is based on the synchronous model we develop to enforce safety properties of the engines built with the BLOCKS library. We begin to apply this approach to wearable computer applications designing (computer system for all around usage in any environment by a mobile user). Applications in this domain fit the SAS architecture and this last meets the requirement of such applications([25])

On another hand, in collaboration with V. Roy and D. Gaffé (Sports, CNRS and UNSA), we consider the problem of compilation of synchronous languages in a modular way. In order to be efficient in application specification and verification, we must be able to deal with large systems. Hence, we introduce a new synchronous model with modularity facilities and sound mathematical semantics in terms of process algebra model.

5. Software

5.1. Ocapi

Until 1996 the Orion team has developed and distributed the OCAPI version 2.0 program supervision engine. The users belong to industrial domains (NOESIS, Geoimage, CEA/CESTA) or academic ones (Observatoire de Nice, Observatoire de Paris à Meudon, University of Maryland).

5.2. Pegase

Since september 1996, the Orion team distributes a new program supervision engine PEGASE, based on the LAMA platform. The Lisp version has been used at Maryland University and at Genset (Paris). The C++ version (PEGASE+) is now available.

5.3. VSIP

VSIP (figure 3) is a Video Surveillance Interpretation Platform written in C and C++. The goal is to build a generic environment applicable as a first step to video surveillance of banks and subways. Besides the image acquisition hardware, the platform is built from three software components: image processing for people detection, human tracking, and interpretation of behaviours relative to the people evolving in the scene. The platform takes as input video streams from several cameras, a geometric description of the unoccupied scene and a set of behaviors of interest specified by experts of the application domain. For each detected event, the algorithms emit automatically an alert and store an annotation in accordance with the set of predefined behavior models. The system has been validated in March 2003 through ten days of multi-camera live recording of a metro station in Barcelona. The next validation will be performed in a bank agency near Paris in December, 2003.

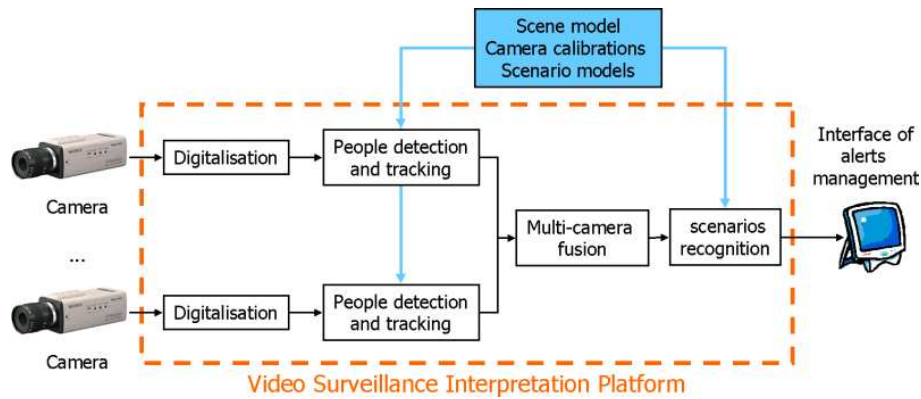


Figure 3. Components of the Video Surveillance Interpretation Platform (VSIP).

6. New Results

6.1. Program Supervision

Participants: Anne-Laure Barjon, Sabine Moisan, Vincent Rampal, Annie Ressouche, Jean-Paul Rigault, Jean-Philippe Vidal.

This year we mainly focus on tasks related to supervision. Adaptations and improvements of our research in this topics are always required. The motivation of this research are applications in the river hydraulics domain and astronomy.

6.1.1. Program Supervision and Model Calibration

Participants: Julien Canet, Sabine Moisan, Jean-Philippe Vidal.

In the framework of a co-directed PhD thesis (with Cemagref Lyon and INPT Toulouse), we study the relationships between program supervision and model calibration tasks. Our application domain is river hydraulics. When a numerical model is built up (e.g., for a river reach and its corresponding hydraulic phenomena, such as flood propagation), the model must be as representative as possible of physical reality. To this end, some numerical and empirical parameters must be adjusted to make numerical results match observed data. This activity – called model calibration – can be considered as a task in the artificial intelligence sense. Model calibration is an essential step in physical process modelling. We propose an approach to model calibration support that combines heuristics and optimisation methods: knowledge-based supervision techniques have been adapted to complement standard numerical modelling ones in order to help end-users of simulation codes.

After a first attempt to model knowledge involved in calibration task with the CommonKADS formalism, we have developed an extended model of objects and sub-tasks using UML class and activity diagrams. This formalism – which proved to be adequate to our problem – allowed us to write a knowledge base expressed with YAKL language to be used with PEGASE+ program supervision engine. We identified three knowledge levels throughout the building of this knowledge base:

- domain-independent knowledge about model calibration: this first level is bound to be as generic as possible to be reused for other application domains. Indeed, this knowledge should be applied when a numerical model of a physical system requires field measurements for calibration of its parameters.
- monodimensional river hydraulics knowledge: the core of the knowledge base is composed of descriptive knowledge – graphical objects involved in expert calibration – and reasoning knowledge represented by expert rules. This second level has been defined in order to be independent of the simulation code used.

- knowledge about a specific hydraulic simulation code. In order to get an operational tool, we used a simulation code called MAGE – and the associated pre and post-processors – developed at CEMAGREF and we have formalised knowledge about its use with a variant of the program supervision approach.

We have currently achieved an important phase which was the specification of artificial intelligence language and tools dedicated to the calibration task, with a focus on its application to hydraulics. The first level knowledge model led us to design specifications for a new language derived from YAKL. This new language is meant to be specific for representing the generic knowledge involved in model calibration. The specifications of a new calibration engine has been completed and its implementation is under way within the LAMA platform. This approach has been presented in the seventh International Conference on Knowledge-Based Intelligent Information and Engineering Systems (KES'2003) [26].

Enhancement of the current knowledge base will be performed thanks to confrontation with hydraulic experts on more and more complex model calibration cases. Validation of the resulting system will be carried out by conducting experiments. Calibration processes and results of the resulting system and expert calibration will be compared on real-life cases.

New Symbolic Curves Module

The Symbolic Curves Module in the LAMA platform was adapted to hydraulic model calibration. This module computes the symbolic description of a sampled curve representing cartesian functions. Such curves are often experiments or observations results stored as a list of points. Our module generates a symbolic description for these curves given a symbolic dictionary that defines the descriptors to use and the digital values they represent. The symbolic description of a curve consists in filtering the curve to obtain a simplified curve and then describing the different parts (such as segments, peaks...) of this simplified curve with symbolic descriptors. For instance, a symbolic description of curve may look like: a medium highly increasing segment, then a sharp advanced peak, then a short highly decreasing segment, and a long flat segment.

The main changes in the new release of the module are:

- New modular architecture (not a single program but a collection of reusable objects). This will allow an easier integration into BLOCKS and the reuse of the objects in other projects.
- STL template objects replaced by BLOCKS ones (to produce "lightweight" executables).
- New symbolic objects/operators to fit the needs for the hydraulic application. For instance, we can now describe the symbolic position of a point compared to a curve. Slope breaks are also described within a symbolic curve description.
- Each symbolic description and comparison can now use its own dictionary.

6.1.2. New Scheduling Engine

Participants: Anne-Laure Barjon, Sabine Moisan, Vincent Rampal.

In a collaboration with the Centre d'Étude Spatial et du Rayonnement (*CESR*) in Toulouse, on automated telescopes, we have proposed this year a prototype implementation of an automatic scheduler, based on the specifications developed last year. A scheduler is currently in use (since 2000) on the autonomous TAROT telescope, but it may have difficulties when handling periodic or constrained requests with a large time interval and it is unable to give to the telescope users and operators the visibility over the schedule. Moreover, if an unexpected event happens, the current schedule is aborted and a new one is computed for the remaining of the night to address the alert observations. The normal process resumes the day after, including the un-observed scheduled requests, which is not optimal.

The new versatile scheduler for automated telescope observations and operations aims at optimising telescope use, while taking alerts (e.g., Gamma-Ray Bursts), weather conditions, and mechanical failures into account. We propose a two-step approach. First, a daily module develops *plan schemes* during the day that offer several possible scenarii for a night and provide alternatives to handle problems. Secondly, a nightly

module uses a reactive technique –driven by events from different sensors– to select at any moment the “best” block of observations to launch from the current plan scheme. In addition to a classical scheduling problem under resource constraints, we also want to provide dynamic reconfiguration facilities. The proposed approach is general enough to be applied to any other type of telescope, provided that reactivity is important. For the daily module, we have implemented the necessary structures to manage the various constraints (astronomical constraints: e.g., orientation, observing windows, target visibility, etc. or resource constraints: e.g., availability of filters, time quotas, or sky zone occultations). This module also applies a “fairness” and scientific priority policy to ensure a fair distribution of observing time among users (i.e. scientific campaigns/programs). For the nightly module, among reactive techniques, we have chosen to use the SyncCharts [40] graphical modeling language. SyncCharts allow to express a complex reactive behavior by simply drawing automata that (logically) communicate by broadcasting signals and favor extensibility. This module should be associated with a visualization interface to graphically display the plan scheme, the current execution line, and the reorganizations due to alerts or problems [32].

6.1.3. Distributed Program Supervision

Participants: Naoufel Khayati, Sabine Moisan.

In the framework of a “STIC cooperation” with Tunisia (ENSI dTunis) on distributed program supervision for medical imagery, a new PhD will start next year. To prepare the PhD study, the student spent a training period to study first, how to interface the PEGASE+ supervision engine and Matlab programs and, second, to improve the program supervision server prototype, that has been developed during previous periods (T. Ben Salah 2000, A. Omrane 2001).

The first step was a communication interface between C++ and Matlab which is used for launching Matlab only once and handling its environment via a C++ program. This interface allows PEGASE+ to control codes written in Matlab (which is the case for the Tunisian medical imaging programs).

The second step is to improve the program supervision server (which allows via Internet to manage, in a distributed way, resources and their users while guaranteeing an acceptable level of safety). The objective is to update the 3D- Reconstruction and Indexation knowledge bases with images and significant programs of the Tunisian GRIFT/ENSI research unit.

The objective of the forthcoming thesis will be to study various distribution methods of program supervision knowledge-based systems (KBS) for medical imagery. Given distributed data, programs, and knowledge, the aim of this thesis will be to propose convenient and efficient models of distributed program supervision, to execute distant physician queries. The first part will concern the development of a knowledge base concerning representative medical imagery programs of research teams in Tunisia and France. This step has been started this year. Second, we will study a prototype of a local KBS to allow the execution of queries on data which are local to the physician sites. Then, we will propose an architecture for a distributed KBS which must allow the remote execution of the same types of queries (e.g., in the form of Web services). Finally, based on the results of the first steps, we will specify the various possible distribution methods according to the needs of the physicians, the working environments, the size of data, etc.

6.2. Software Platform for KBS Development

Participants: Sabine Moisan, Annie Ressouche, Jean-Paul Rigault.

We wish to improve the LAMA platform with tools devoted to the verification of engine behavior. The basic idea to get reusable and efficient development of knowledge based system (KBS) is to adopt a component based approach to support this development process. Hence, a generic framework BLOCKS allowing KBS construction by subtyping of its components has been defined and we have studied how to ensure a safe subtyping with respect to the component properties. The least we expect is that the derived classes respect the properties that the base classes implement and guarantee. But, it turns out that this concern (called substitutability principle) is popular in component framework approach and is not only a BLOCKS usage concern. To enforce a safe subtyping, we want to apply formal methods of verification. But, we do not want

to use neither testing methods since they are not complete nor theorem prover techniques since there are not totally automatic. Thus, we consider model-checking techniques: they are exhaustive, automatic and well-suited to our problem.

This year, we have completed the formal approach started last year. In this approach, the substitutability principle can be ensured by applying model-checking techniques.

6.2.1. Component Framework Verification

Participants: Sabine Moisan, Annie Ressouche, Jean-Paul Rigault.

This year, we have first defined a *synchronous* mathematical model and a restriction operation which characterizes the notion of substitutability. Of course, the model fits component behavior representation, but to be practicable in the description of these behaviors, we defined a dedicated language and a semantics that bridges the gap between the language and the model. The semantics is structurally defined. The restriction operation leads to a preorder relation in the mathematical model and this preorder relation is compositional with respect to the language operators. On another hand, this preorder relation also preserves safety properties and thus ensures us that these last are preserved through subtyping. We also show that we can apply modular model checking techniques in our model and we define some practicable *design rules* at the language level whose application guarantees safe subtyping. Our formalism and its properties is detailed in a technical report ([30]) and has been presented in a workshop dedicated to component-based systems specification ([24]).

Besides, our approach allows to practically ensure a safe usage of component framework, its drawback is that the restriction operation is too strong and lead to reject too many classes as not safe derivations of basic ones. A future work will be to study how to relax the restriction operation in order to accept some derivations which are rejected by the current implementation.

We have applied this formal approach to build a tool for correct BLOCKS manipulation. The description language we defined is implemented and this year we focus on model-checking utilities in the tool. To this aim, we have integrated the NUSMV model-checker. NUSMV verifies safety properties in a very efficient way, by using a symbolic approach to represent models and by using powerful verification methods based on “SAT” solvers (very efficient tools based on propositional logic formula satisfaction). This integration works in two ways: our behavior description programs are translated as NUSMV models and the counter examples given back by NUSMV in case of violated properties are interpreted in the language. The future work related to this practical issue is to implement the substitutability analyzer based both on the design rules and on the restriction operation we defined.

6.3. Automatic Interpretation of Image Sequences

Participants: Alberto Avanzi, François Brémond, Bernard Boulay, Binh Bui, Thi-Thanh-Tu Bui, Frédéric Cupillard, Benoit Georis, Florent Fusier, Magali Mazière, Monique Thonnat, Christophe Tornieri, Thinh Van Vu.

The goal of this activity is to automatise the understanding of the activities happening in a scene. Sensors are mainly one or several fixed and monocular video cameras in indoor or outdoor scenes ; the observed mobile objects are mainly humans and vehicules. Our objective is the modeling of the interpretation process of image sequences and the validation of this model through the development of a generic interpretation platform. These techniques are applied in the framework of six projects : the transfer action Telescope2, the European project ADVISOR and the four following industrial projects: RATP, CASSIOPEE, ALSTOM and SNCF.

The problem in which we are interested is the interpretation of the behaviour of people acting in a scene; i.e. to find a meaning to their evolution in the scene. This scene is observed by one or several fixed video cameras. To realise the interpretation, we need to solve two sub-problems. The first one is to provide for each frame measures about the scene content and the system in charge of this problem is called “perceptual” module. The second one is to understand this content. So, we try to recognise predefined scenarios based on visual invariants. The system in charge of the second problem is the module of scenario recognition. Our approach to image sequence interpretation is based on the a priori modeling of the observed environment.

This year, we have extended our works on the modelisation of the reference image and on people detection. We have proposed new approaches for the recognition of human postures and the real time recognition of temporal scenarios. We have also put an emphasis on the evaluation of the video understanding platform by end users and we have started building a new framework for the automatic evaluation and repair of the video understanding process.

6.3.1. Update of the Reference Image

Participants: Christophe Tornieri, François Brémond, Monique Thonnat.

Motion detection is one of the main steps of video analysis. This step is done by subtracting the current image from a reference image corresponding to the background image. In order to be robust, the platform must update at each image the reference image to take into account the illumination changes and also changes of the environment such as a moving door. Classical updating methods gradually integrate a portion of the current image into the reference image. These methods are well adapted for slow illumination changes. However they do not handle sudden illumination changes nor environment changes. We propose a new algorithm that consists in the modeling of the integration process to discriminate parts of the image that correspond to illumination and environment changes (opening a cupboard) from those that correspond to individuals.

The main idea of our algorithm is to compute moving regions corresponding to individuals evolving in the scene and to compute stationary regions corresponding to noise and to integrate the stationary regions into the reference image without integrating the individuals. A moving region is a tracked blob from two consecutive images. It often corresponds to an individual evolving in the scene and should not be integrated in the reference image. A stationary region is a part of the current image that does not appear in the reference image and that is always detected at the same location in the current images. A stationary region usually corresponds to a noise such as a sudden illumination change that lasts in the current images. It is represented by a rectangular zone in the image to which we associate a template. The template gives the precise shape of the pixels constituting the stationary zone. With this template, we are able to evaluate motion in the zone and determine if the motion is due to noise or to human movements.

We have tested our algorithm on different video sequences: six videos of subway station, one video of an outdoor railway (1 hour), four videos of a bank office (up to two hours) and live cameras in a office. Even if there are a lot of people, our algorithm manages to include only illumination and environment changes. This algorithm has been presented in IDSS 2003 symposium([17]) and has been integrated in the ORION visual surveillance platform.

6.3.2. People Detection

Participants: François Brémond, Florent Fusier, Monique Thonnat.

People detection is an important step in video interpretation. We proposed an algorithm to locate people in videos for video understanding applications. The input of the algorithm is the list of the moving regions computed from the difference of the current image and the background image. A moving region can correspond either to an isolated individual, a part of an individual (e.g. the legs) or a group of people when people are overlapping each other. The proposed algorithm counts the number of persons inside each mobile region and their localization (2D and 3D) in the scene using only the difference image. This algorithm is based on two complementary methods :

- a method based on person head and feet detection.
- a method based on a human-shape ellipsoid mask.

The goal of the algorithm is to refine a rough classification step which consists in associating a type (a part of a person, a person, a group, a crowd) for each mobile region based on its size. The first method consists in detecting the head and if it is possible the feet of each person in the mobile region. Two models of head are defined : one based on the global shape of a head, detected by analyzing the projections of the potential head, as shown on Fig. 4. The other model is based on the Omega shape composed of the head, neck and

shoulders. The head shape model is composed by four projection models depending on the viewpoint of the camera : closed-field, mid-field (including face and side), far-field of the camera. When heads and feet in a mobile region have been detected, we verify if their characteristics (like the real size, the localization of each part of body,...) match the model of a person.

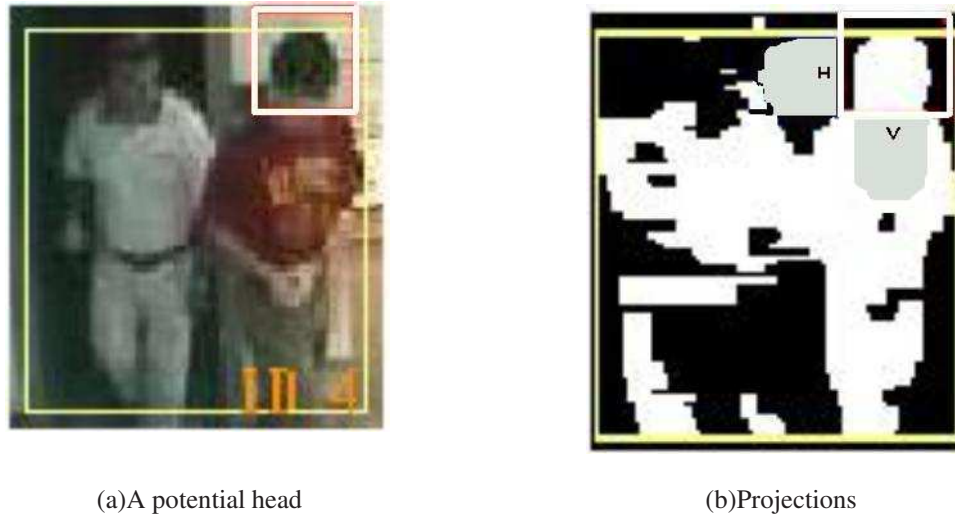


Figure 4. A potential head (in the highlighted box of left figure) and its projections (right figure): its horizontal projection(H) and its vertical projection(V).

The second method is based on an ellipsoid human shape mask, as shown on Fig. 5. The principle is to fit the ellipsoid mask with a potential person to check either the moving pixels inside the ellipsoid mask matched the model of a person (density of moving pixels, 3D size of this mask, height/width ratio). This method is used in two different ways associated to two situations : (1) a person is detected by the head/feet detection method, (2) when no people have been detected by the head/feet method in a large part of a mobile region. When a person has been detected by the head/feet method, the aim of the second method is to verify the characteristics of the detected person to avoid false detections. The second case (when nobody has been detected) occurs often when the mobile region has a strong inclination, due to the camera geometric deformation. In this case, the mask is computed on the whole mobile region. To apply this mask, we compute the inclination of the mobile region using camera information to compensate the geometric deformation at this position in the scene. Then the mask characteristics are checked as previously described.

The algorithm has been evaluated on more than thousand mobile objects, the results are encouraging with approximately 82 % of true positive and 18 % of false positive. The rate of false positive is mainly due to a bad quality of detection and people overlapping each other. In order to solve the problem related to geometric deformation of camera, we are applying the compensation step of the inclination of mobile objects to the head detection and projections phases. This work is described in the report [37].

6.3.3. Human Posture Recognition

Participants: François Brémond, Bernard Boulay, Monique Thonnat.

Posture recognition is one step in the global process of analyzing human behavior. Behavior analysis is an important field dealing with many applications such as video surveillance or domotics. Usually, human behavior is recognized through the study of trajectories and positions of persons and using a priori knowledge about the scene (localization of doors, walls, areas of interest,...). This method is well adapted to a scene



Figure 5. Ellipsoid masks of detected persons on the real image (highlighted ellipses)

with large field of view observing the full trajectories of people. But, when we consider a cluttered scene where there is no continuous observation of people displacement, we often do not have enough information to accurately determine behavior. Recognizing posture then is a necessary step to recognize behavior more accurately.

To determine the posture of a person in a video, we use the mobile object detected by the detection module of VSIP platform (Video Surveillance Intelligent Platform).



Figure 6. The T-shape posture and its blob

First, we determine which postures we want to recognize. Up to now, there are seven postures which must be classified in three categories: the standing postures (standing with arms near the body, standing with arm to left or to right and T-shape posture (cf. figure 6)), the seated postures (seated on the floor or on a chair) and the bending posture. Then, we propose two 2D appearance-based methods and an approach which combines 2D appearance-based methods and a 3D human model, to recognize the postures. The first one uses horizontal and vertical (H.&V.) projections of mobile object on the reference axis of the 2D binary image. A learning phase

is made to determine the average (H.&V.) projections for each posture. Then the current (H.&V.) projections are compared to the average (H.&V.) projections by using a sum of squared differences. The second method decomposes the human silhouette into blocks and compute the density of pixel in movement for each block to obtain a vector. The learning phase is made by using a PCA to compute an average vector for each posture. Then the current vector is compared to the average vectors by using the Mahalanobis distance. And finally, we use 3D models of each posture (cf. figure 7) to make the previous methods independent of the camera view point. The orientation and the 3D position of the mobile object are computed and applied to the 3D models. Then the current mobile object is compared with the mobile objects of the 3D models by using the previous 2D method.

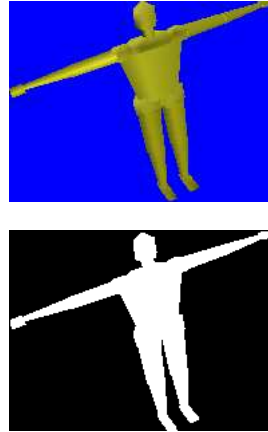


Figure 7. T-shape posture of 3D model and its blob

We obtain good results by using the 2D appearance-based methods (76% of correct recognition for the projections method and 80% for the block density method, tests realized on more than 1000 mobile objects). Combining 2D methods and 3D model gives encouraging results. But we need to automate the process to validate our approach. This work was published in proceedings joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance (VS-PETS 2003) [16][33].

6.3.4. Classification of Lateral Forms

Participants: François Brémond, Binh Bui, Monique Thonnat.

In a system of automatic video interpretation, one of the essential stages consists of classifying the mobile objects in the scene because they can correspond to several types of different objects (a person, a push-chair, etc). To solve this problem, we proposed a new method using a Bayesian network for the classification of mobile object shapes observed from the side (lateral form). For that, we improve the site equipped last year with a camera observing the scene from the top and with five lateral cameras observing from the side optical fibers placed on the other side.

Models of Lateral Forms

Initially, we have to build the models of the lateral forms for different mobile objects. These models are built from the characteristics of the objects detected by the cameras (for example density of the fibers hidden by the object). We decompose the lateral form of a mobile object into several zones (three or nine zones). The form is decomposed into three zones if it is partially detected (the mobile object has just entered or leaved the site). The form is decomposed into nine zones if it is completely detected (the mobile object is completely in the site). The size of each zone is proportionally defined from the size of the mobile object. For each zone, we compute the number of free and hidden fiber parts. The models of lateral forms are built from the combination

of these zones. To refine the models of lateral forms, in addition to these zones, we also added the length, the width and the height of the mobile object.

Automatic Learning of Lateral Forms

For each model (class) of lateral forms, we use ten typical sequences (representative of the class), corresponding to hundreds of frames. For each frame, we compute and save the values of the density of free/hidden fibers for each zone. We count the number of mobile objects having the same density value for a given zone z and thus we obtain the frequency for a given mobile object class to have the density value d of free/hidden fibers for the zone z .

This year, to automate the learning stage, we have developed a software for the user to easily enter/save the characteristics of the mobile objects in the scene. Once the user has chosen a video sequence, the software shows each frame of the sequence for the user to delimit (draw a 2D bounding box) the mobile object seen in the frame and to choose its class c . Then, for each mobile object, the software:

- computes automatically its 3D bounding box from its 2D bounding box thanks to the camera calibration.
- decomposes the 3D bounding box into three or nine zones according to the position of the mobile object in the scene (partially or completely seen).
- for each zone, computes the value of density of free/hidden fibers and then updates the model for class c .

For each frame of the sequence, the number of objects and their class are also saved in several files. These files also enable in a post-treatment to evaluate the classification module. For the evaluation, we compare the output of classification module with the real data called “ground truth”, in other words, with the content of these files.

This software is written in C/gtk++.

6.3.5. Temporal Scenario Recognition for Automatic Video Interpretation

Participants: François Brémond, Thi-Thanh-Tu Bui, Monique Thonnat, Thinh Vu Van.

Our goal is to study the problem of *Temporal Scenario Recognition for Automatic Video Interpretation*. In particular, we want to design an algorithm recognizing in real-time temporal scenarios pre-defined by experts and taking as input individuals tracked by a vision module and a priori knowledge of the observed environment. We have proposed a novel approach taking advantages of two state of the art approaches: one proposed by N. Rota in his thesis defended in ORION research team in 2001 and the other proposed by M. Ghallab & C. Dousson at LAAS in 1994. To do this, we have to solve five problems: Firstly, the representation of a scenario by state of the art algorithm is not intuitive. In particular, the scenarios often correspond to instantaneous events detected uniquely at a given instant. We added in our formalism the possibility to define scenarios in both time interval and time point. For example, Fig. 8 shows a representation of a scenario "bank attack" in the context of a bank agency. This work had been published in the proceedings of KES2002 conference [13]. Besides representation issues, reducing the complexity of the temporal scenario recognition algorithm is a challenge. The recognition can be viewed as a Temporal Constraint Satisfaction Problem (TCSP). So, the algorithm to solve this problem belongs generally to the NP-complete class. Thus, the other tasks of our work concern the reduction of the processing time for the recognition algorithm.

Secondly, to get processing time reduction, we optimized the search of a given scenario instance in the set of recognized scenario instances by indexing this set by a graph (called *graph of solutions*) that is constituted of scenario models and actor lists of already recognized scenarios. This indexing based on scenario models allows to reduce notably the processing time of the algorithm. This work has been published in the proceedings of the workshop "Modeling and Solving Problems with Constraints" held during ECAI2002 conference [12][34].

Thirdly, we also proposed to reduce the processing time of the algorithm by factorizing composed scenario models into simpler scenario models. The similar algorithms of the state of the art re-verify the temporal constraints with previously recognized scenario instances until the given scenario is recognized.

```

Scenario(bank_attack,
Characters(com : Person, rob : Person)
Sub-Scenarios (
  (com_at_pos : inside_zone (com, "back_counter"))
  (rob_enters : changes_zone(rob, "entrance_zone", "front_counter"))
  (com_at_safe : inside_zone (com "safe_zone"))
  (rob_at_safe : inside_zone (rob, "safe_zone")) )
Forbidden-Scenarios (
  (any_in_bank : inside_zone (anyP, "bank"))      )
Constraints (
  (com_at_pos before com_at_safe) // (1)
  (rob_enters before rob_at_safe) // (2)
  (rob_enters during com_at_pos) // (3)
  (anyP ≠ com) // (4)
  (anyP ≠ rob) // (5)
  (any_in_bank during com_at_safe) // (6)

```

Figure 8. Representation of the scenario "bank attack": this scenario is composed of four sub-scenarios. This description contains six constraints. The constraints (1)-(3) express a sequence of sub-scenarios composing the scenario. Constraints (4) - (6) in connection with the forbidden scenario express that there is no other person (*anyP* is a person different from *com* and *rob*) in the bank during the attack

This verification implies the search for sub-scenario instances in the list of already recognized scenarios, thus it can lead to a combinatorial explosion problem. To solve this problem, we propose to decompose scenarios in an initial phase of pre-compilation in order for each compiled scenario to contain only two sub-scenarios. To decompose a scenario model, we have to check the consistency of the constraints defined within the scenario and order its sub-scenarios by their ending time using a graphical method. This decomposition enables the recognition algorithm to look for only one sub-scenario instance at each instant, thus it implies only one search in a linear time in function of the number of recognized scenario instances. This work has been published in the proceedings of ICVS2003 conference [28].

Fourthly, state of the art algorithms try all combinations of actors to recognize composed scenarios. As a consequence the recognition algorithm can become exponential with respect to the actor number. To solve this problem, we propose to reconstruct the list of actors of a composed scenario instance from the actors of its sub-scenarios instead of trying all combinations of these actors. By using this new method the time to recognize a compiled composed scenario is closed to a linear algorithm with respect to the actor number. This work has been published in the proceeding of IJCAI2003 conference [27].

Fifthly, the scenario models to be recognized can be defined manually by expert or generated automatically in the compilation of composed scenario models. Thus, the redundancy of information (i.e. scenario models) in a scenario knowledge base is often important. To eliminate the redundancy, we proposed that for a group of scenario models having the same semantic, we keep only one of them. A main problem to be solved concerns the comparison of two scenario models. Two scenario models are identical if they have the same list of actors, the same list of sub-scenarios and their set of constraints have a same semantic. As the compiled composed scenario models contain only two sub-scenarios, we also eliminate the redundancy inside composed scenario models. Finally, to validate the proposed algorithm, we integrated the recognition algorithm in the platform of Automatic Video Interpretation VSIP (Video Surveillance Intelligent Platform) and realized four different types of tests: (1) on recorded videos taken in a bank agency and in two metro stations (one in Belgium and one in Spain) to verify if the algorithm can correctly recognize the predefined scenario models, (2) on live videos acquired on-line from cameras installed in an office, in a metro station and in a bank branch to verify

if the algorithm can work robustly on a long time mode, (3) on recorded videos taken in a bank agency and on simulated data to study how the complexity of the algorithm depends on the scenario models (i.e. number of sub-scenarios and actors) and (4) on simulated data to study how the complexity of the algorithm depends on the complexity of the scene (i.e. number of persons in the scene). The experimental results in term of processing time show that the proposed recognition algorithm is quasi-linear in function of the number of sub-scenarios as shown on Fig.9 and in function of the number of actors as shown on Fig.10.

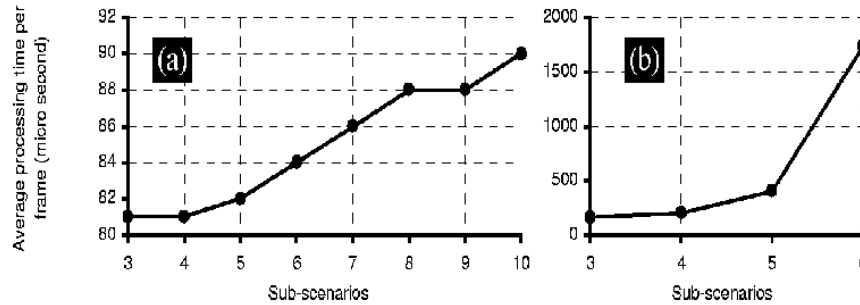


Figure 9.

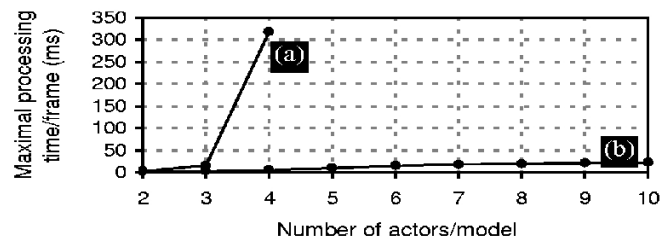


Figure 10.

6.3.6. Video Understanding Evaluation

Participants: François Brémond, Alberto Avanzi, Frédéric Cupillard, Magali Mazière, Monique Thonnat, Christophe Tornieri.

We have put this year an emphasis on the building of experimental sites and on the evaluation of the video understanding platform by end users.

Experimental Site for Bank Monitoring:

In December 2002, the Orion VSIP platform adapted for the CASSIOPEE project has been installed on a PC located in a bank agency. The platform processes color images coming from a surveillance video camera looking at the agency, at a framerate of 10 images per second. The goal of the platform is to detect a predefined scenario called *bank attack*. This scenario involves 2 actors : an employee and a robber. The robber enters the bank agency, goes towards the employee, standing behind the counter and threatens the employee to open the safe door.

A performance evaluation has been done by playing several times with actors the same bank attack scenario. Other “normal” scenarios (scenarios which are not supposed to arise an alarm) have been played too, always

by actors, to evaluate the robustness of the platform against false alarms. The results showed 80 % of true positive and 0 % of false positive.

Ten Days of Live Demonstration in Barcelona Metro Station:

In March 2003, at the end of the European Project ADVISOR, evaluation, validation and demonstration of the prototype have been done at the TMB headquarter in Barcelona to various guests, including the European Commission, project Reviewers and representatives of Bruxelles and Barcelona Metro. Together with this demonstration, an evaluation has been done by security operators of the metro of Barcelona and Brussels in charge of the videosurveillance during a week at the Sagrada Familia Metro Station in Barcelona.

The evaluation, validation and demonstration were conducted using both live and recorded videos. For the validation task, the system was tested in live conditions using four input channels in parallel, the four channels being composed of three recorded sequences and one live input stream from the main hall of the Sagrada Familia Metro station. The three recorded sequences enabled to test the system with rare scenarios of interest, not always available during the demonstration. The three recorded data sequences were constructed using thirty-two shorter prerecorded sequences, showing five different predefined scenarios, four of them (*fighting*, *blocking*, *jumping over the barrier* and *vandalism*) played by actors and one (*overcrowding*) coming from original videos. The live camera allowed to evaluate the system against scenarios which often happen (e.g. *overcrowding*) and which can occur during the demonstration. It also allowed to evaluate the system against false alarms.

In total, out of 21 fighting incidents in all the Demonstrator sequences, 20 alarms were correctly generated, giving a very good detection rate of 95%. These twenty correctly identified alarms had an average report accuracy of 68% (by accuracy we mean the temporal overlap between intervals corresponding to the detected alarm and the ground truth). Out of nine blocking incidents, seven alarms were generated, giving a detection rate of 78%. These seven alarms were found to be 60% accurate on average. Out of 42 instances of jumping over the barrier, including repeated incidents, the behaviour was detected 37 times, giving a success rate of 88%. The two sequences of vandalism were always detected with an overall accuracy of 71%, over six instances of vandalism. Finally, the two overcrowding alarms in camera C11 were consistently detected, with an overall accuracy of 80% over 7 separate instances of the alarms. The overcrowding alarms were also consistently detected in the live camera C10, with some 28 separate events being detected.

In conclusion, the ADVISOR demonstration has been evaluated very positively by end-users and European Committee. The algorithms responded very successfully to the input data, with high detection rates, less than 5% of false alarms and with all the reports being above approximately 70% accurate. A documentary movie of 11 minutes has been realised to present the ADVISOR system and to show the live demonstration in Barcelona. This work has been reported in [43][42] [17]

Permanent Live Demonstration at INRIA:

Since September 2003, the experimental videosurveillance platform VSIP is running at Inria, using two videocameras installed in an office. The platform receives two live color data stream (sequences of images at 5-10 frames per second, coming from the two CCTV cameras), and performs person and object detection, person tracking (on each camera), data fusion (with objects and persons detected by each camera), event and scenario recognition. The platform has been used to test the recognition of the scenario *bank attack* and constitutes a permanent test bed to validate all new video understanding algorithms.

6.3.7. Towards an Automatic Evaluation and Repair Framework for Video Interpretation

Participants: François Brémond, Benoît Georis, Monique Thonnat.

In the past few years, many interpretation systems have been developed but none of them have been successfully applied to real world applications. One major weakness of these systems is the tracking process. Tracking is still a central issue in scene interpretation, as the loss of a tracked object prevents the analysis of its behaviour. Tracking has been extensively studied for many years. Various techniques have been explored, both model-based and model-free. Nevertheless, the tracking problem remains unsolved since there are

many sources of ambiguities like shadows, illumination changes, over-segmentation and mis-detection. These difficulties need to be handled in order to make the correct matching decision.

During this year, we have proposed a new framework for automatic evaluation and repair of video interpretation systems, which is currently applied on the short term tracking algorithm. The proposed framework (figure 11) is composed of four main parts:

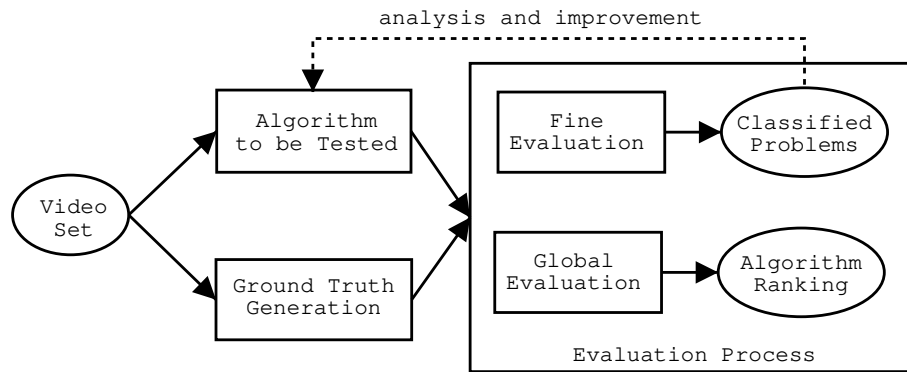


Figure 11. Proposed framework for algorithm automatic evaluation

- Algorithm to be Tested:** Currently, we have applied this framework on the short term tracking algorithm which is composed of two steps: the mobile object detection procedure and the frame to frame tracking procedure. Up to now, other tracking algorithms can be evaluated with this framework for the global evaluation process only.
- Representative Video Set:** A careful selection of test video sequences is mandatory in order to perform a relevant evaluation. Indeed, tracking algorithms are developed following precise hypotheses which describe the type of video sequences algorithms can process. Thus, tracking evaluation strongly depends on the input sequence type. Moreover, all tracking algorithms succeed when video sequences are simple and fail when they are difficult. We have first listed all difficulties we can encounter in video understanding, such as camera and scene motion, slow or fast illumination changes, automatic gain control of cameras, sensor type influence,... Then we have chosen to focus on 3 difficulties: the average number of people in the scene, the detection quality and the number of crossings between persons. Finally, according to these 3 criteria, we have currently selected 6 indoor scene video sequences from three different applications: a bank agency, a metro platform and an office.
- Ground Truth Generation:** Tracking algorithms are designed in function of specific results to achieve related to the target application. Some algorithms want to correct detection results and process a whole person (e.g., recover his/her feet) even if this person is partially observable due to an occlusion (e.g., his/her feet are behind a desk). Other algorithms do not need to recompute accurately the detection of the person. Therefore, the issue of defining ground truth consists in being impartial for all tracking algorithms. So, we have chosen to draw a full bounding box for each mobile object as soon as we see an evidence of its presence in the image (e.g., a hand, a shadow, ...). Then, we have defined 8 ground truth attributes: the width, the height and the position of the mobile object, computed both in 2D and in 3D, the type (PERSON, VEHICLE, ...) and an identifier. To this end, we have used a software interface called ViPER (cf : <http://lamp.cfar.umd.edu/media/research/viper/>).

- **Evaluation Algorithms and Result Criteria.** The evaluation is done at two levels. First, a global evaluation process is applied. During this step, tracking algorithms are seen as black boxes. The goal is to rank all algorithm types using global criteria such as the number of missed tracks or the number of identifiers per mobile object. Second, a fine evaluation process, which is algorithm dependent, analyses in detail each sub step of the algorithm and produces a precise classification and diagnosis of tracking errors. This is an iterative procedure. A classification of tracking errors is defined and the associated errors are identified automatically. Once this identification is done, the goal is to diagnose which parameter or which part of the algorithm generated an error class. The last step consists in repairing manually the algorithm. Finally, the new tracking algorithm is re-evaluated.

The second evaluation level and the associate repair methodology to improve the tracking is the main contribution of this work. We are currently working towards an automatic repair process. For this purpose, we are investigating the extraction of all sensitive parameters from each algorithm sub step and to relate/to tune them according to the knowledge of the 3D scene and the application.

This work has been described in the Master Thesis report [38] and has been presented at the 3rd International IASTED Conference on Visualization, Imaging and Image Processing in September 2003 [19].

6.4. Complex Object Recognition

This year, research activities on complex object recognition have been continued. A cognitive vision platform composed of three knowledge based systems is currently under development. This work is conducted in cooperation with INRA Sophia Antipolis. The platform will be used for the detection of plant diseases. Ontology based knowledge acquisition for object description has also been studied. In particular, we have built a learning architecture which can be used for visual concept recognition. A complete image formation model for microscopic 3D translucent objects has been developed. N. Dey has defended his PhD thesis on this subject.

Key words: *cognitive vision, classification, categorization, learning, image formation, environment, plant diseases, pollen.*

6.4.1. A Cognitive Vision Platform for Natural Complex Object Recognition

Participants: Céline Hudelot, Nicolas Maillot, Sabine Moisan, Monique Thonnat.

Our goal is to provide a generic and re-usable cognitive vision platform for the automatic recognition of natural complex objects in their natural environment. The task of natural object recognition is an hard task which can be divided into more tractable sub-problems:

1. The semantic data interpretation problem
2. The mapping between high level representations of physical objects and image numerical data
3. The image processing problem, i.e. segmentation and numerical description

To separate the different types of knowledge and the different reasoning strategies involved in the object recognition process, we propose a distributed architecture based on three highly specialized Knowledge Based System (KBS). Each KBS is specialized for the corresponding sub-problem of object recognition. The beginning of this year has been dedicated to the conception of the architecture ([21]) of the cognitive vision platform. Then the formalization of the different types of knowledge adapted to the different types of expertise and the specification of the 3 engines with adapted reasoning involved in the platform have been done ([20]).

The proposed platform is currently under implementation with the development platform LAMA. In cooperation with INRA (URIH de Sophia Antipolis), the early detection of plant diseases, in particular rose diseases is used to evaluate and validate our platform ([15]).

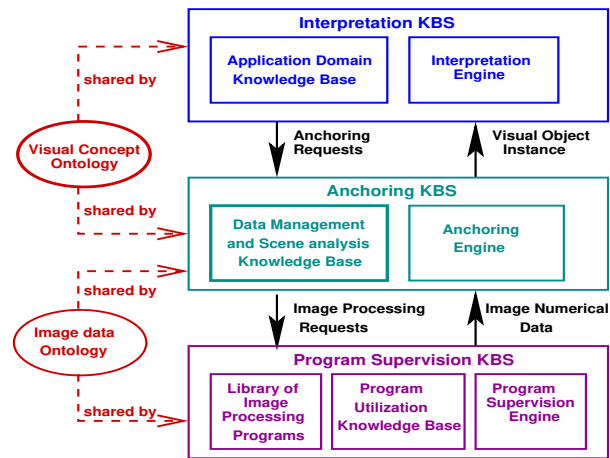


Figure 12. Global architecture of our cognitive vision platform

6.4.2. Ontology-Based Knowledge Acquisition and Visual Concept Learning

Participants: Céline Hudelot, Nicolas Maillot, Sabine Moisan, Monique Thonnat.

Experts often use a well-defined vocabulary to describe complex objects (e.g. in palynology, in astrophysics, etc.). Our goal is to capture this knowledge to use it into our cognitive vision platform for automatic recognition of natural complex objects. We propose an ontology-based acquisition process to guide knowledge acquisition. A visual concept ontology has been designed for that purpose. This ontology is structured in several parts : spatio-temporal concepts, texture concepts, color concepts, relations between concepts (e.g. spatio-temporal relations), and context concepts (e.g. point of view, acquisition device).

The knowledge base resulting from the acquisition process is used by the cognitive vision platform. A knowledge acquisition tool called OntoVis has been developed and used for the description of pollen grain images. This tool allows domain knowledge acquisition (i.e. domain objects and their subparts) and visual description guided by the ontology. This tool also provides an efficient module to manage image samples. This year has been dedicated to the diffusion of these results ([23] and [22]). We have also designed and implemented a learning architecture to learn and recognize the visual concept used during visual description. Image processing operators are applied on image samples to extract numerical descriptors used during the learning phase.

The learning phase uses the tree structure of the visual concept ontology to learn the visual concepts at different levels of granularity. For that purpose, a binary classifier is attached to each visual concept C_i and is specialized in its recognition. The output of a classifier gives the probability $P(C_i|\mathbf{x})$; it represents the probability of an unknown pattern \mathbf{x} to be a representative sample of C_i . One of the interests of this approach is that classification consistency can be checked at numerical level (ambiguity and distance rejection) and at the knowledge level (i.e. a reasoner can use the ontology to check the validity of the classification results). This architecture is currently applied to texture recognition.

6.4.3. Image Formation of a 3D Translucent Microscopic Object (Application to Light Microscopy)

Participants: Nicolas Dey, Monique Thonnat.

A complete image formation model for microscopic 3D translucent objects has been developed. In our model, we define a translucent object as a discrete repartition of refractive indexes and absorption coefficients. To simulate the trajectory of light, we propose a physical model using ray tracing (rays are traced from the light source to the observer). This physical model is used to calculate the lit object space. To simulate the image

generation by the optical system, we use some wave optic principles. We model the 3D transfer function of the microscope, which depends on the amount of defocusing. We first choose the focused plane, and then we calculate a simulated image using this transfer function with each plane of the lit object space that is more or less defocused. If we change the focused plane, we obtain a simulated image sequence.

A PhD thesis [44] on this subject was defended by N.Dey on November 26th 2002. Results have been published in [18].

7. Contracts and Grants with Industry

7.1. Industrial Contracts

Participants: Alberto Avanzi, François Brémond, Binh Bui, Frédéric Cupillard, Magali Mazière, Monique Thonnat, Christophe Tornieri.

In 2003, ORION team has been involved in 5 industrial projects : ADVISOR project on subway visual surveillance which finished in march 2003, RATP project on subway user classification, CASSIOPEE project on bank agency visual surveillance, ALSTOM project and SNCF project on train visual surveillance.

- **ADVISOR** : This project lasted 3 years and finished in march 2003 with 2.4MF financial resources for ORION. The aim of this project was to automatically analyse (recognize scenarios of interest) and annotate subway video sequences [35][36][31][39].
- **RATP** : This project has a duration of 3 years and will provide 1.6MF financial resources to ORION. The aim of this project is to classify in real-time different types of subway users.
- **CASSIOPEE** : This project has a duration of 3 years and will provide 450KEuros financial resources to ORION. The aim of this project is to develop and test a visualsurveillance platform allowing automatic detection of predefined scenarios in a bank agency environment.
- **ALSTOM** : This project began in april 2003 and finished in july 2003. It lasted 4 months (funding for Orion). The aim of this project was to realize a feasibility study on an industrial application allowing to detect human behaviors in trains using techniques developed in the ORION team.
- **SNCF** : This project has begun in september 2003 for a duration of 18 months (funding for Orion). The aim of this project is to automatically detect human behaviors in trains. During this project ORION team will develop techniques to recognize human behaviors and scenarios of interest.
- **TELESCOPE 2 & TELESCOPE 3** : The TELESCOPE 2 project ended in march 2003 after a two years duration and will be followed by the TELESCOPE 3 project. The aim of both projects is to complement an initial project(ended in 2001) in which a toolkit in the domain of cognitive video interpretation for videosurveillance applications(VIS) has been achieved . The purpose of these two projects is to improve this toolkit in order to facilitate its usage, to ensure more robustness and to extend its functionalities.
- **Videa** : This project began in november 2003 and has a 2 years duration. The aim of this project is to transfer a part of the videosurveillance technology of the ORION team into industrial products. During this project, the ORION team will develop and transfer to a videosurveillance company two applications enabling the recognition of specific human behaviours.

8. Other Grants and Activities

8.1. European projects

ORION team has been involved this year in two european projects on image interpretation : ADVISOR project and ECVision European Research Network (IST-type).

8.1.1. *ECVision European Research Network*

The ECVision European Research Network has begun in march 2002 for 3 years. This research network was formed to promote and merge activities of 50 european laboratories working in cognitive vision (see <http://www.ECVision.info/home/Home.htm>).

8.1.2. *ADVISOR Project*

ADVISOR project began in january 2000 and finished in march 2003. The aim of this project was to develop an intelligent system which would automatically detect events of interest and inform subway operators of such events. Furthermore, the system had to annotate and archive video sequences in order to be able to later retrieve interesting video sequences by a post-processing stage. During this project, ORION team has developed innovative activities such as fusion from multiple cameras, real-time processing of video sequences, annotation of video sequences and real-time recognition of scenarios of interest. Project members were Thales (U.K.), Bull (France), Vigitec (Belgium), Kingston University (U.K.), Reading University (U.K.) and INRIA (France).

9. Dissemination

9.1. Scientific Community

- M. Thonnat is an area leader of the ECVision European Excellence Network in cognitive vision domain since march 2002 for 3 years (50 teams and 12 countries).
- M. Thonnat is an expert for the RNTL program
- M. Thonnat is a reviewer for AIJ (Artificial Intelligence Journal), PATREC, IJVC(International Journal on Computer Vision), CVIU (Journal of Computer Vision and Image Understanding), IJPRAI and RIA review (Revue d'Intelligence Artificielle).
- M. Thonnat is a reviewer for the conferences: RFIA, TAIMA, PETS and CVPR 2004 for High Level Vision and for Visualsurveillance (Area Chair).
- F. Brémond and M. Thonnat are members of the US standardisation group ARDA in order to define an ontology dedicated to video event recognition.
- M. Thonnat is a member of the Joint Executive Committee to organize cooperations between the NSC (Taiwan) and french research teams. Franco-Taiwan conferences related to Multimedia and Web Technologies.
- F. Brémond is a reviewer for the RFIA conference.
- S. Moisan is a member of the program committee of the IC'2003 conference, about knowledge software
- S. Moisan is a member of the 27^e department of specialist committee at UNSA (Nice Sophia Antipolis University)

9.2. Teaching

- Orion is a hosting team for the DEA of computer Science of UNSA
- Teaching at DESS of Computer Science at Essi (UNSA), Object-oriented Analysis and Conception lecture (25h S. Moisan).
- Teaching at DEA of Astronomy , image and gravitation (UNSA) classification lectures (9h M. Thonnat and 3h F. Brémond)
- Teaching at ISIA (Institut d'Informatique et d' Automatique, Ecole des Mines de Paris) grammar analysis lecture and TP (16h A. Ressouche)
- Contribution to a MIG (Module d'Intégration Générale) Seminar on Verification Methods and managing of student projects (A. Ressouche)
- Contribution to the DEUG Math-info of UNSA (Nice University): Computer Science and Programming lectures and practice (Céline Hudelot)
- Java language lectures (50h) and Unix system programming in C lectures (25h) at Nice IUT GTR (Génie des Telecommunications et Réseaux) (N. Maillot).

9.3. Thesis

M. Thonnat has recently presented her habilitation thesis (habilitation à diriger des recherches) untitled: "Towards Cognitive Vision Knowledge and Reasoning for Image Analysis and Interpretation" [14].

The following Phd theses are in progress in the Orion team:

- Celine Hudelot : Interprétation automatique d'images in situ de végétaux pour la détection et le suivi de pathologies, Nice Sophia Antipolis University.
- Nicolas Maillot : Système cognitif d'interprétation d'images pour la reconnaissance d'images d'objets 3D, Nice Sophia Antipolis University.
- Tinh van Vu : Visualisation de comportements humains pour l'interprétation de séquences video, Nice Sophia Antipolis University.
- Jean-Philippe Vidal : Equifinalité dans les modèles numériques en hydraulique à surface libre : méthodologie de calage de paramètres, Polytechnique National Institute of Toulouse.
- Benoit Georis : Knowledge-based reconfigurable tracker, Louvain Catholic University.
- Bernard Boulay : Reconnaissance de postures pour l'interprétation d'activités humaines, Nice Sophia Antipolis University.

9.4. Participation to Conferences, Seminars, and Invitations

Member of the Orion team have presented papers in the following conferences:

- WACV 2002 (Workshop on Application in Computer Vision) (Orlando, USA)
- 4th Sino Franco Workshop on Multimedia and Web Technologies Workshop (Tapei, Tawain)
- IDSS (Intelligent Distributed Surveillance Systems Workshop) (London, UK)
- ICVS 2003 (3rd International Conference on Computer Vision Systems)(Graz, Austria)
- SAVCBS'2003 (Specification and Verification of Component-Based Systems) (Helsinki, Finland)
- KES 2003 (7th International Conference Knowledge-Based Intelligent Information and Engineering Systems) (Oxford, UK)
- IJCAI 2003 (Acapulco, Mexico)
- VIIP conference (Benalmadena Costa del Sol, Spain)

- ARDAV Video Challenge Workshops (San Diego USA, Monterey USA)
- ICCV et VS PETS 2003 (Nice France).
- ICTAI 2003 (International Conference on Tools with Artificial Intelligence) (Sacramento, USA).
- Cognitive Vision Workshop (reseau d'excellence ECVision)(Dagstuhl, Germany)
- ISWC 2003(Sanibel Island Florida, USA)
- 2003 IEEE (International Conference on Systems, Man and Cybernetics, SMC'03) (Whashington, USA)

10. Bibliography

Major publications by the team in recent years

- [1] F. BRÉMOND, M. THONNAT. *Issues of representing context illustrated by video-surveillance applications*. in « International Journal of Human-Computer Studies, Special Issue on Context », volume 48, 1998, pages 375-391.
- [2] N. CHLEQ, F. BREMOND, M. THONNAT. *Advanced Video-based Surveillance Systems*. edition C.S. Regazzoni, G. Fabri, and G. Vernazza, Kluwer A.P. , Hangham, MA, USA, November, 1998, chapter Image Understanding for Prevention of Vandalism in Metro Stations, pages 108-118.
- [3] V. CLÉMENT, M. THONNAT. *A Knowledge-Based Approach to Integration of Image Procedures Processing*. in « CVGIP: Image Understanding », number 2, volume 57, March, 1993, pages 166–184.
- [4] S. LIU, P. SAINT-MARC, M. THONNAT, M. BERTHOD. *Feasibility study of automatic identification of planktonic foraminifera by computer vision*. in « Journal of foramineferal research », number 2, volume 26, April, 1996, pages 113–123.
- [5] S. MOISAN. *Une plate-forme pour une programmation par composants de systèmes à base de connaissances*. Habilitation à diriger les recherches, université de Nice, avril, 1998.
- [6] S. MOISAN, A. RESSOUCHE, J.-P. RIGAULT. *Blocks, a Component Framework with Checking Facilities for Knowledge-Based Systems*. in « Informatica, Special Issue on Component Based Software Development », number 4, volume 25, November, 2001, pages 501-507.
- [7] M. THONNAT, V. CLÉMENT, J. OSSOLA. *Automatic Galaxy classification*. in « Astrophysical Letters and Communication », number 1-6, volume 31, 1995, pages 65-72.
- [8] M. THONNAT, M. GANDELIN. *Un système expert pour la description et le classement automatique de zooplanctons à partir d'images monoculaires*. in « Traitement du signal, spécial I.A », number 5, volume 9, November, 1992, pages 373–387.
- [9] M. THONNAT, S. MOISAN. *What can Program Supervision do for Software Re-use?*. in « IEE Proceedings - Software Special Issue on Knowledge Modelling for software components reuse », number 5, volume 147, 2000.

- [10] M. THONNAT, S. MOISAN. *Knowledge-based systems for program supervision..* in « First international workshop on Knowledge-Based systems for the (re)Use of Programs libraries KBUP'95 », INRIA, pages 4–8, Sophia Antipolis, France, March, 1995.
- [11] M. THONNAT, N. ROTA. *Image Understanding for Visual Surveillance Applications.* in « Third International Workshop on Cooperative Distributed Vision », Kyoto, Japan, November, 1999.
- [12] V.-T. VU, F. BRÉMOND, M. THONNAT. *Temporal Constraints for Video Interpretation.* in « 15th European Conference on Artificial Intelligence », Lyon, France, 2002.
- [13] V.-T. VU, F. BRÉMOND, M. THONNAT. *Video surveillance: human behaviour representation and on-line recognition.* in « The Sixth International Conference on Knowledge-Based Intelligent Information & Engineering Systems (KES'02) », Crema, Italy, 2002.

Doctoral dissertations and “Habilitation” theses

- [14] M. THONNAT. *Vers une vision cognitive: mise en oeuvre de connaissances et de raisonnements pour l'analyse et l'interprétation d'images..* Habilitation à diriger les recherches, université de Nice, octobre, 2003.

Publications in Conferences and Workshops

- [15] P. BOISSARD, C. HUDELLOT, M. THONNAT, G. PEREZ, P. PYRRHA, F. BERTAUX. *Automated Approach to Monitor the Sanitary Status and to Detect Early Biological Attacks on Plants in Greenhouse.* in « International Symposium on Greenhouse Tomato, Avignon », CTIFL, IEEE, editor, 2003.
- [16] B. BOULAY, F. BREMOND, M. THONNAT. *Human Posture Recognition in Video Sequence.* in « Proceedings Joint IEEE International Workshop on VS-PETS, Visual Surveillance and Performance Evaluation of Tracking and Surveillance », pages 23-29, OCT 11-12, 2003.
- [17] F. CUPILLARD, F. BRÉMOND, M. THONNAT. *Behaviour Recognition for Individuals, Groups of people and Crowd.* in « IEE Proceedings of the Intelligent Distributed Surveillance Systems Symposium - IDSS2003 », London, UK, February, 2003.
- [18] N. DEY, A. BOUCHER, M. THONNAT. *Un modèle pour la formation de l'image d'un objet microscopique 3D translucide.* in « Actes du Troisième colloque francophone Traitements & Analyse d'Informations, Méthodes & Applications TAIMA'2003 », volume 1, pages 219-224, September, 2003.
- [19] B. GEORIS, F. BREMOND, M. THONNAT, B. MACQ. *Use of an Evaluation and Diagnosis Method to Improve Tracking Performances.* in « Proceedings of VIIP'03- 3rd IASTED International Conference on Visualization, Imaging and Image Processing », Benalmadera, SPAIN, SEP 8-10, 2003.
- [20] C. HUDELLOT, M. THONNAT. *A cognitive Vision Platform for Automatic Recognition of Natural Complex Objects.* in « International Conference on Tools with Artificial Intelligence, ICTAI, Sacramento », IEEE, editor, 2003.
- [21] C. HUDELLOT, M. THONNAT. *An Architecture for Knowledge Based Image Interpretation.* in « Workshop on Computer Vision System Control Architectures, VSCA, Graz », 2003.

- [22] N. MAILLOT, M. THONNAT, A. BOUCHER. *A Visual Concept Ontology for Automatic Image Understanding*. in « Poster Session of the 2nd International Semantic Web Conference, ISWC », 2003.
- [23] N. MAILLOT, M. THONNAT, A. BOUCHER. *Towards Ontology Based Cognitive Vision*. in « Computer Vision Systems, Third International Conference, ICVS », series Lecture Notes in Computer Science, volume 2626, Springer, J. L. CROWLEY, J. H. PIATER, M. VINCZE, L. PALETTA, editors, 2003.
- [24] S. MOISAN, A. RESSOUCHE, J.-P. RIGAUT. *Behavioral Substitutability in Component Frameworks: a Formal Approach*. in « Specification and Verification of Component-Based Systems (SAVCBS'2003) Workshop at ESEC/FSE 2003 », pages 22–28, Helsinki, Finland, September, 2003.
- [25] A. RESSOUCHE, V. ROY, J. TIGLI, D. CHEUNG. *SAS architecture: Verification Oriented Formal Modeling of Concrete Critical Systems*. in « 2003 IEEE International Conference on Systems, Man and Cybernetics, SMC'03 », Washington, USA, October, 2003.
- [26] J.-P. VIDAL, J.-B. MOISAN. *Knowledge-Based Hydraulic Model Calibration*. in « Knowledge-Based Intelligent Information and Engineering Systems (KES) », series LNAI, number 2773, volume I, Springer, L. J. V. PALLADE, editor, pages 676–683, Oxford, UK, September, 2003.
- [27] V.-T. VU, F. BRÉMOND, M. THONNAT. *Automatic Video Interpretation: A Novel Algorithm for Temporal Scenario Recognition*. in « The Eighteenth International Joint Conference on Artificial Intelligence (IJCAI'03) », Acapulco, Mexico, 2003.
- [28] V.-T. VU, F. BRÉMOND, M. THONNAT. *Automatic Video Interpretation: A Recognition Algorithm for Temporal Scenarios Based on Pre-compiled Scenario Models*. in « The 3rd International Conference on Vision System (ICVS'03) », Graz, Austria, 2003.

Internal Reports

- [29] S. MOISAN. *Program Supervision: YAKL and PEGASE+ Reference and User Manual*. Technical Report, number 5066, INRIA, December, 2003, <http://www.inria.fr/rrrt/rr-5066.html>.
- [30] S. MOISAN, A. RESSOUCHE, J.-P. RIGAUT. *A Behavior Model of Component Framework*. Research Report, number 5065, INRIA, December, 2003, <http://www.inria.fr/rrrt/rr-5065.html>.

Miscellaneous

- [31] A. AVANZI. *Watermarking Functionality for the ADVISOR System*. in « Advisor deliverables 040 », May, 2003.
- [32] A.-L. BARJON, V. RAMPAL. *Séquenceur d'observations pour télescope autonome*. Technical report, ESSI, 2003, Rapport de stage dernière année d'école d'ingénieur.
- [33] B. BOULAY. *Reconnaissance de postures caractéristiques d'un comportement humain*. in « Rapport de Stage de DEA Image-Vision - Université de Nice Sophia-Antipolis », November, 2003.

- [34] T.-T.-T. BUI. *Optimisation de la bibliothèque des modèles de scénarios temporels pour l'interprétation automatique de séquences vidéos*. in « DEPA thesis », 2003.
- [35] F. CUPILLARD, C. TORNIERI, F. BREMOND, M. THONNAT, A. AVANZI. *Calibration. Modelling, Motion and Tracking for Multiple Cameras*. in « Advisor deliverables 041: R2.2 », March, 2003.
- [36] F. CUPILLARD, C. TORNIERI, F. BREMOND, M. THONNAT, P. REMAGNINO, A. AVANZI. *Advanced Behaviour Recognition*. in « Advisor deliverables 042: R3.2 », March, 2003.
- [37] F. FUSIER, F. BREMOND, M. THONNAT. *Segmentation de personnes évoluant dans une scène*. in « Rapport de Stage de DESS Sciences Cognitives Appliquées - Université Bordeaux 2 », September, 2003.
- [38] B. GEORIS. *Towards an Evaluation and Repair Framework for Video Interpretation*. 2003, Master Thesis, Université Catholique de Louvain.
- [39] M. RENARD. *ADVISOR: Final Evaluation Report*. in « Advisor deliverables 039:R8.4 », March, 2003.

Bibliography in notes

- [40] C. ANDRÉ. *Representation and Analysis of Reactive Behaviors: A Synchronous Approach*. in « CESA'96 Proceedings », IEEE-SMC, Lille, France, 1996.
- [41] V. CLÉMENT, M. THONNAT. *A Knowledge-Based Approach to Integration of Image Procedures Processing*. in « CVGIP: Image Understanding », number 2, volume 57, March, 1993, pages 166–184.
- [42] F. CUPILLARD, F. BRÉMOND, M. THONNAT. *Group Behavior Recognition With Multiple Cameras*. in « IEEE Proceedings of the Workshop on Applications of Computer Vision - ACV2002 », Orlando, USA, December, 2002.
- [43] F. CUPILLARD, F. BRÉMOND, M. THONNAT. *Tracking Group of People for Video Surveillance*. edition Paolo Remagnino, Graeme A. Jones, Nikos Paragios and Carlo S. Regazzoni, series Video-Based Surveillance Systems, volume The Kluwer International Series in Computer Vision and Distributed Processing, Kluwer Academic Publishers, 2002, pages 89-100.
- [44] N. DEY. *Etude de la Formation de l'Image d'un Objet Microscopique 3D Translucide - Application à la microscopie optique*. Ph. D. Thesis, Université de Nice - Sophia Antipolis, November, 2002.
- [45] S. KAWASHIMA, Y. TAKAHASHI, S. AIKAWA, T. NAGOYA. *An attempt of applying the image processing for the automatic estimation of sampled airborne pollen*. in « Japanese Journal of Allergology », number 9, volume 44, 1995, pages 1150–1158.
- [46] S. LIU, P. SAINT-MARC, M. THONNAT, M. BERTHOD. *Feasibility study of automatic identification of planktonic foraminifera by computer vision*. in « Journal of foraminiferal research », number 2, volume 26, April, 1996, pages 113–123.
- [47] S. LIU, M. THONNAT, M. BERTHOD. *Automatic Classification of Planktonic Foraminifera by a Knowledge-*

based System. in « The Tenth Conference on Artificial Intelligence for Applications », IEEE Computer Society Press, pages 358–364, San Antonio, Texas, March, 1994.

- [48] M. MAZIÈRE. *Étude de faisabilité pour la reconnaissance automatique de grains de pollen*. 1997, Rapport de DEA, université de Nice-Sophia Antipolis.
- [49] J. OSSOLA. *Coopération de systèmes à base de connaissances pour l'analyse et la reconnaissance d'objets naturels complexes : application au classement de galaxies ou de zooplanctons*. Ph. D. Thesis, université de Nice-Sophia Antipolis, May, 1996.
- [50] M. THONNAT, A. BIJAOUÏ. *Knowledge-based galaxy classification systems*. A. HECK, F. MURTAGH, editors, in « Knowledge-based systems in astronomy », series Lecture Notes in Physics, volume 329, Springer Verlag, 1989.
- [51] M. THONNAT, V. CLÉMENT, J. OSSOLA. *Automatic Galaxy classification*. in « Astrophysical Letters and Communication », number 1-6, volume 31, 1995, pages 65-72.
- [52] M. THONNAT, M. GANDELIN. *Un système expert pour la description et le classement automatique de zooplanctons à partir d'images monoculaires*. in « Traitement du signal, spécial I.A », number 5, volume 9, November, 1992, pages 373–387.
- [53] M. THONNAT. *The World of Galaxies*. edition Corwin, H. Jr. and Bottinelli, L., Springer Verlag, 1989, chapter Toward an automatic classification of galaxies, pages 53-74.
- [54] R. TOMCZAK, P. BONTON, C. AUROYER, D. CAILLAUD, C. ROUQUET. *Traitement d'images et reconnaissance des formes appliquées à la mesure des taux de pollens dans l'air*. in « Journées Automatique, Agriculture et Agro-alimentaire », Clermont-Ferrand, France, 1997.
- [55] R. VINCENT, M. THONNAT, J. OSSOLA. *Program Supervision for Automatic Galaxy Classification*. in « Proc. of the International Conference on Imaging Science, Systems, and Technology CISST'97 », June, 1997.