Title: Computational and Cognitive Vision Systems: A Training European Network
Acronym: VISIONTRAIN

Human Resources and Mobility Activity
Marie Curie Actions
Research Training Networks

PART B

Call identifier: FP6-2002-Mobility-1

November 19, 2003

1	Institut National de Recherche en Informatique et Automatique (France), coordinator	INRIA
2	International School for Advanced Studies (Italy)	SISSA
3	University of Malmö (Sweden)	Malmö
4	IT University of Copenhagen (Denmark)	IT-UC
5	The Technion (Israel)	Technion
6	University of Oxford (Great Britain)	Oxford
7	University of Ljubljana (Slovenia)	UOL
8	Ecole Polytechnique Fédérale de Lausanne (Switzerland)	EPFL
9	Czech Technical University (Czech Republic)	CTU
10	Universiteit Utrecht (The Nederlands)	U. Utrecht
11	University of Mannheim (Germany)	U. Mannheim

Contents

List of Tables

1 Scientific quality of the project

1.1 Research topic

Vision is a complex process which transforms optical stimuli into high-level representations of the real world. The scientific objective and research topic of this Marie Curie Research Training Network is to carry out, collaboratively, a number of theoretical and methodological studies which will eventually enable computers to "see" and to "understand" what they see. Our research approach is to explore simultaneously computer and biological vision in order to have a better knowledge of the mechanisms of human visual perception and thus have a significant impact on algorithm design, performance evaluation, and interfacing of computer vision systems with humans.

Systems for handling and understanding visual information are expected to have as a great impact on society over the next decades, as what computers and telecommunication have on today's society. In particular they will relieve humans from the burden of communicating with increasingly complex systems, be they technical or deriving from everyday needs. They will make many new applications possible.

Computer vision systems have already been built. They can operate in very restrictive domains or in carefully controlled environments – artificially constrained worlds – where models can be constructed. Attempts to build more general-purpose computer vision systems are numerous. Currently such systems perform well because they make use of human-assisted steps.

> *The ambitious overall objective of this project is to make progress towards the understanding of the computational, cognitive, and biological bases of visual processes. Eventually the theories, methods, and experiments developed by the network researchers will allow the design of complete and autonomous computer vision systems. The approach advocated herein is to build proof-of-concept computer systems on top of formal mathematical and computational models and theories.*

More precisely we propose to work on the following aspects:

- At the low-level, a number of **monocular cues** are present in an image and their extraction requires the simultaneous estimation of the following scene characteristics: the geometrical layout, the light field, and the material properties;

- **Stereopsis** has been one of the most difficult process to understand and current computational models are not entirely satisfactory. We want to develop fast, accurate, and robust stereo methods that are able to acquire dense depth data from images associated with natural scenes;

- A major task of vision is to perform **motion analysis**. In the past the problem has been solved for images of rigid scenes, solutions for finding the motion field for highly deformable and articulated objects need to be investigated. Optimality criteria will be derived whenever possible and solutions based on variational approaches will be investigated.

- The vision system plays a major role in capturing motion information and in rapidly **interpreting motion stimuli in terms of higher level events**. In particular we plan to study methods for **human motion and gesture recognition**. We intend to develop sophisticated body models that can be used to guide and disambiguate the image-based motion analysis process.
- The appearance of objects in images and in image sequences comprises both shape and colour/texture within regions with the shape as its boundary. We want to develop **appearance models for shapes** which are flexible enough to account for the high variability of these components in real scenes;
- **Recognition of an object class** (such as people, cars, trees, dogs, etc.) is a challenging problem because in addition to the usual visual difficulties that affect an object's image appearance, such as lighting, shadowing, viewpoint and partial occlusion, there is also the issue of within class variability. If we are to build vision systems that can recognize, say 10,000 object categories, then **effortless learning** is a crucial step. Therefore we plan to develop unsupervised learning/recognition techniques, i.e., operator intervention is reduced to a minimum or eliminated entirely;
- One of the most fundamental aspects of communication between people is the **recognition of faces** as well as the recognition of associated facial expressions. A human face is a complex shape and currently there is no satisfactory natural mathematical representation of such shapes. Moreover, there is a discrepancy between a mathematically convenient definition of shape distance and the human perception of similarity. It is important to develop models which bridge this gap;
- The human vision system is very efficient because it is able to concentrate very quickly on informative pieces within images. Highly effective **visual search (or attention) mechanisms** in the human visual system were extensively studied in psychophysics and physiology. We intend to investigate and develop computational models of attention;
- It is well established that the visual system processes information on an intentional basis – it is goal directed. This means that higher-level information is used to control lower-level information extraction and this top-down strategy is guided by knowledge. The **representation of knowledge** typically needed to successfully accomplish a visual task is far more complex that we are used to in other domains such as, for example, natural language understanding.
- The fact that brain imagery spatio-temporal resolution is constantly improving implies that we are able to **measure the activity of cortical entities** whose size is getting within the reach of the current modelling and simulation tools. Another goal will be to **model at a higher level than the voxel the way the brain solves a visual perception task.**

1.2 Project objectives

With respect to the research topics just enumerated, this project aims at developing a number of scientific objectives which may be grouped as follows:

The first objective: Computational theories and methods for low-level vision;

The second objective: Motion understanding from image sequences;

Third objective: Learning and recognition of shapes, objects, and categories;

The fourth objective: Cognitive modelling of the action of seeing, and

The fifth objective: Functional imaging for observing and modelling brain activity.

1.2.1 Computational theories and methods for low-level vision

Monocular cues figure prominently in both computer vision and in human psychophysics. The theory of many of these **monocular cues** has been mainly developed in computer vision, taking ideas from psychophysics (for instance, **shape from shading**, **shape from texture**, **shape from contour**, and so forth). Psychophysics is using these theories increasingly in order to help understand human perception. There are many more leads to take from psychophysics that it would be most useful to develop formally. We are interested to test and develop theories of monocular cues and especially their interactions. It is not at all clear that the human observer functions along the lines of the formal theories. If not, then there must exist alternative methods and we are much interested in identifying them and developing them formally. A correct understanding of the functioning of human observers is in itself a most important topic since computer graphics is aimed at the human observer (essentially sending a message to the observer without really understanding the receiver!) and many man-machine interface depend heavily on graphical communication between man and machine.

Moreover, we want to design novel **variational approaches for solving shape-from-shading** problems. Such problems consist of recovering the 3D surface of an illuminated object from one or multiple 2D images. These schemes should differ from existing ones in a number of properties: (i) The data term in the energy functional does not have to be quadratic. This allows to incorporate methods from robust statistics. (ii) The regulariser can also be a non-quadratic function. This allows for discontinuities in the surface derivatives. (iii) The regulariser may also be anisotropic. This strategy enables e.g. a stronger smoothing along a discontinuity than across it. (iv) The regulariser may involve derivatives of orders larger than one. This leads to Euler-Lagrange equations of order greater or equal to 4. As a consequence, the surface derivatives may be linear (instead of constant). This can result in more realistic surface reconstruction algorithms. (v) Efficient numerical algorithms for the large nonlinear systems of equations will be developed. (vi) The performance of the algorithms will be evaluated using data where the ground truth is available. Such a test-bed allows to compare our novel algorithms with existing ones.

Stereoscopic matching is a long-standing problem in low-level vision. Despite a significant effort over the past decades, obtaining **dense and accurate correspondences among a set of images** without restricting the application domain still constitutes a challenging task.

Our long-term goal in this area is to achieve (1) accuracy in stereoscopic matching, i.e. to guarantee gross-error probability to be less than a pre-defined level in unrestricted scenes and (2) high precision, i.e,. to avoid usual interpolation artifacts under the presence of weakly textured objects. We intend to develop a game-theoretic approach to stereoscopic matching that allowing to meet these criteria directly.

The matching problem will be formulated as one of finding a super-stable kernel in a special graph. The principle used will allow for the construction of many matching problem variants (multi-image matching, multi-criteria image similarity, natural integration of multiple cues for stereopsis, dense or sparse tentative matches, single-valued or multi-valued matching, etc). Importantly, it is possible to formulate the problem in a way that a high-level vision process could influence the low-level matching task.

1.2.2 Motion understanding from image sequences

One main goal is to develop a completely **autonomous structure from motion system**. The purpose of the system is to provide accurate and photo-realistic 3D structure from image sequences taken by un-calibrated hand-held video cameras. Prototype systems were developed in the past within a series projects. However, significant improvements in several aspects are needed in order to reach the overall goal.

Firstly, the feature extraction and tracking needs to be improved. This will be done by using scale-space theory to extract significant and invariant features and by implementing an improved correlation tracker. Secondly, the motion estimation and auto-calibration needs to be improved. This will be done in several respects. Automatic hypothesis generation, testing and constraint enforcement will be included. This will allow the system to solve common problems, such as reappearing features, patching together different sub-sequences and detecting and utilizing planar structures. Dynamic modelling will be introduced in order to obtain more stable and robust results. The dynamic models will describe the motion of the camera and model selection between different models will be included. Thirdly, the object representation needs improvement. We will investigate different ways to parameterize surfaces, such as spline representations and Bezier surfaces. Of special interest is to use dense depth maps (using stereoscopic matching) as a starting step to a more natural representation of the scene. Finally, an important step in order to obtain photo-realism is to estimate the BRDF:s (Bi-Reflectional Distribution Function) from the image sequence. Low parameter models will be used to model the BRDF:s and estimate them. Also specularities will be used, both to infer the structure of the scene and to model the reflectance properties. Also this task will made in cooperation with the teams working on low-level vision aspects of the project.

In recent years, there has been increasing interest in modelling **human shape and motion** from image data. Such an ability has many applications, such as entertainment, sports, medicine, and training. This, however, is an inherently difficult task, both because the body is very complex and because the data that can be extracted from images is often incomplete, noisy and ambiguous.

To overcome these difficulties, we intend to develop sophisticated body models that can be used not only for animation purposes, but also to guide and disambiguate the image-analysis process. In earlier work, we have proposed and implemented an innovative framework based on implicit surfaces and have shown its ability to derive body shape and motion from stereo and silhouette data. The robustness and applicability of this framework, however, is limited by the fact that the underlying model is too simple and not anatomically correct enough. We can effectively track and model limbs but are less successful where more complex parts of the body, such as shoulders, spine and hips, are concerned. We intend to remove this limitation by: (1) Creating and integrating new and more sophisticated models that take biomedical constraints into account and (2) Developing effective optimization technique to handle the many degrees of freedom of those models.

1.2.3 Learning and recognition of shapes, objects, and categories

The main scientific objective within this proposal is to develop mathematical and computational models for the representation, acquisition and use of prior knowledge about the visual appearance of objects in images. The appearance of objects comprises both shape appearance and the appearance in terms of colour/texture within regions with the shape as its boundary. The objective is to develop appearance models which are flexible enough to account for the high variability of these components in real scenes, and for which the acquisition of knowledge about the appearance of objects by statistical learning from images and videos is feasible. The models to be developed should be applicable to the visual detection, segmentation, recognition and tracking of objects in real scenes.

The visual data, which is acquired by our retina, needs to be processed in order to infer significant descriptions, i.e., until particular parts of our surroundings are recognized to a certain extent. Recognition is hence an essential part of the human perception. Similarly, visual recognition should be an essential part of an intelligent machine system as well.

Recognition itself implies learning. Therefore, one (human or machine) has first to learn how an object looks and store its representation in the memory to be able to recognize it in the future. Learning, representation, and recognition are thus three inseparable parts of visual perception.

Visual recognition seems to be an easy task for humans. We can recognize most (previously seen) objects almost immediately without any particular effort. On the other hand, visual learning and recognition is one of the most difficult tasks in computer vision. Several methods for machine visual learning and recognition have already been proposed, however all of them have very limited functionality. Mostly they are constrained to specific conditions and domains. In the real world, however, the conditions are not always ideal. Input data may be corrupted by noise, occlusions and other undesirable artifacts. Our main scientific objective within this proposal is to develop novel robust methods for visual learning and recognition, which will be able to handle such artifacts and work under non-ideal conditions. This would lead to the applications designed for real-world conditions.

The objective is to recognize visual object classes (such as people, airplanes, cars, leopards) in an image or video, and determine their position within the image or frame. Furthermore, the object classes should be learnt directly from example images of that class, where the images are not pre-segmented and may contain visual "clutter".

Recognizing an object class is a challenging problem because, in addition to the usual visual difficulties that affect an object's imaged appearance, such as lighting, shadowing, viewpoint, and partial occlusion, there is also the issue of within class variability. For example, the difference between a jumbo jet and a glider.

If we are to build vision systems that can recognize, say, 10,000 object categories, then effortless learning is a crucial step. This means that the training sets should be small and that the operator-assisted steps that are typically required (e.g. elimination of clutter in the background of the object, scale normalization of the training images) should be reduced to a minimum or eliminated entirely. Ideally one should be able to play a video containing the object class of interest, and the model should be learnt from that alone.

1.2.4 Cognitive modelling of the action of seeing

It has long been recognized that vision is a cognitive process: knowledge representation and intelligent behavior are embedded in the vision system, is it biological or artificial (David Hubel, 1979 and David Marr 1982). For example, half of the human brain is, more or less, involved in vision. Hence, modelling vision from a cognitive point of view is crucial. The task is tremendous and it will not be possible, within the lifetime of a networked project to solve the problem. Nevertheless, we intend to investigate to important features of visual behavior: attention and top-down control.

This work comes to quantify the performance abilities of computerized attention systems for visual search. We aim to suggest search algorithms based on probabilistic methods and other quantitative approaches, and to implement a system for testing their performance. In addition, we want to quantify the inherent difficulty of a search task, both by probabilistic measures and deterministic measures, that would help in predicting the search time. In particular, we want to explore how mutual similarities (or dissimilarities) between objects in the scene can contribute to the search efficiency. In addition, relatively little was done to quantitatively characterize the inherent difficulty of search tasks, which we consider as our main goal.

A prevailing direction of research has been on methods for development of the complex models required for a cognitive system to understand a complex world. It is a firm belief that systems have to be organized in a hierarchical or pyramid manner, where lower level outputs feed into higher level operations. Similarly, higher level information will be used for control of lower level operations in the form of context. Learning has to be used to overcome the immense complexity in specification of a multiplicity of partly overlapping models in dependence upon context. This generates particular demands upon the information representation, to allow information to be communicated and usable at levels different from where it was generated.

It is argued that information must be acquired actively by the system itself, through response driven association with percept transformations. After an active training, however, the system can exhibit a reactive behaviour to passively observed percepts. From this derives our view that the development of powerful Cognitive Vision structures inevitably has to go the path over Active Vision, Learning and Robotics, even if the systems will be used for interpretation of static imagery.

1.2.5 Functional imaging for observing and modelling brain activity

Brain imagery is necessary in order to improve our knowledge of visual perception and it also provides some challenging computer vision problems. The fact that brain imagery spatio-temporal resolution is constantly improving implies that we are able to **measure the activity of cortical entities** whose size is getting within the reach of the current modelling and simulation tools. Due to their complementary in terms of resolution and information they can recover, we will mainly focus on the three main modalities Magnetic Resonance Imaging (MRI), Electro-encephalography (EEG) and Magneto-encephalography (MEG).

Another goal will be to model at a higher level than the voxel **the way the brain solves a visual perception task**. This will help us to push forward the state of the art of the knowledge of the brain mechanisms supporting visual perception and to unveil a computational architecture that may inspire us in our computer vision work.

1.3 Scientific originality of the project

1.3.1 Computational theories and methods for low-level vision

We are interested in developing and using psychophysical techniques in the context of the exact sciences and engineering. Psychophysical techniques should be important tools in man-machine interfacing and computer graphics. Yet-at this moment-they are not. Problems are the difficult communication between experimental psychologists and engineers and scientists with a strong formal background. Moreover, the "output" of typical psychophysical studies is not in a format that is immediately useful in engineering applications. We are interested in developing the tools that are useful to the engineers and can be applied to real problems, yet are fully acceptable as state of the art or beyond by the experimental psychology community.

Variational approaches to shape-from-shading have been pioneered in the eighties. In that time quadratic regularization methods were state-of-the-art. In the meantime, however, tremendous progress has been made in the field of nonlinear image processing and computer vision. Examples include nonlinear diffusion filtering, non-quadratic regularization methods such as total variation de-noising. In some computer vision areas such as optic flow estimation, these novel ideas have been rapidly incorporated into previous approaches. As a result, significantly better methods have been developed. In the area of shape-from-shading, however, this progress has not yet entered the field. Thus our goal is to transfer a number of ideas that have been successfully used in other imaging areas to the field of shape-from-shading. Such a transfer is by no means straightforward, since the shape-from-shading problem differs significantly from many other computer vision problems such as optic flow estimation. It contains a significantly more complicated data term where the unknown surface derivatives enter in a nonlinear way. Moreover, integrability constraints have to be included in order to ensure that both derivatives result from a joint function. Besides these modelling problems, it has to be noted that most current variational algorithms for shape-from-shading do not take advantage from state-of-the-art algorithms from numerical analysis: Often simple fixed point iterations are used and not much is known about their convergence. We expect that a detailed investigation of suitable numerical algorithms could lead to a significant progress in this field.

Scientifically, the application of graph kernel theory to the area of stereoscopic matching is largely unexplored. What is more, the graph kernel theory itself is still an area of active research with many potential applications in other fields. From an application point of view, the state of the art in the computer vision field will be advanced by the ability to cope with complex unknown scenes, by the guaranteed performance (obtaining matchings at a given confidence level), by the fact that the approach naturally opens a door for a high-level vision process to influence the low-level optimization task, and by the fact that the algorithms are simple and fast (at least in the most important problem types) and easy to parallelize.

1.3.2 Motion understanding from image sequences

The main innovative aspects of the proposed research lies in the development of the automatic motion estimation module. On innovative aspect is the hypothesis generation, testing and constraint enforcement. Another one is the usage of dynamical models to stabilize and robustify the whole process, including the model selection mechanism. There will also be innovative aspects in the propose surface representation techniques and the usage of dense-depth maps as a starting point to generate better object representations. We also expect to contribute to the state of the art of estimating BRDF:s. We have already made pioneering work on using specular reflections to estimate surface shape. This work will continue to deal with more complex BRDF:s. A final innovative aspect is to introduce feedback loops in the reconstruction process instead of processing each step after another. For example, this will make it possible to find new and better features, once the motion is known. It will also be possible to find new object points once a rough surface description has been obtained.

In recent years there has been much interest in capturing complex motions solely by analyzing video sequences. Single camera solutions would be ideal to process standard image sequences. However, they are not always robust, in part because image data is inherently noisy and in part because it is inherently ambiguous. By contrast, using multiple cameras leads to a considerable reduction in the size of the search space and a considerable improvement in robustness, at the cost of having to deal with a large amount of data, most of which is redundant. Most of the proposed approaches, however, rely on oversimplified body models, such as cylinders or ellipsoids attached to articulated skeletons. Such models are too crude for precise recovery of both shape and motion and it is this shortcoming that our approach addresses.

Methodologies for reconstruction of degraded and/or noise-corrupted images exist in term of Bayesian methods, PDE-based methods, and simple filtering techniques. The two first provide the most efficient methods. These methods have, however, not yet convincingly been generalized to the situation of image sequences.

Methodologies for computation of visual motion exist in term of optic flow-based methods and model-based methods. The optic flow-based methods can handle the situation of very complex motion patterns but are very sensitive to the image quality. Model-based methods can handle the situation of noise-corrupted/degraded image quality but can only handle motion patterns as complex as the model they are based on.

The state-of-the art is hence that images may be reconstructed to some degree, but high quality reconstruction of image sequences requires prior knowledge of the motion pattern. Simultaneously, estimation of complex motion patterns requires high image quality. Thus, the sequence restoration and motion determination problem is a classical "chicken and egg problem".

In the last years, a lot of attention have been devoted to the problem of optic flow and disparity estimation, the innovative aspects of the proposed research concern, firstly the approach to the problem, most of the techniques proposed in the literature are based on correlation estimations or stochastic approaches. We propose a different methodology based on sophisticated and accurate models formulated in terms of energy minimization, variational approach and partial differential equations. Secondly, the problem of video sequence analysis using multiple cameras including occlusions estimation is rather new in the literature (at least from the point of view of the proposed methodology). The combination of different video sequences taken at the same time from different cameras includes new challenging problems as how to relate the stereo information between the different cameras with the sequence of images taken from each one of the cameras.

1.3.3 Learning and recognition of shapes, objects, and categories

Our research addresses the following points corresponding to limitations of the state of the art.

Non-Euclidean nature of shape: There is no "natural" mathematical representation of shape. In particular, due to the correspondence problem and required invariance properties, shapes cannot simply be regarded as points in some Euclidean space. As a consequence, many tools of mathematical and computational modelling for the purpose of statistical modelling, comparison, classification and recognition are not directly applicable.

Similarity measures and metrics: There is a discrepancy between mathematically convenient similarity (or distance) measures and the human perception of similarity. Accordingly, it is important to develop models which bridge this gap.

Statistical learning from examples: In order to account for the high variability of object appearance, non-parametric models developed recently in the field of machine learning will be preferred in order to address the "bias-variance" dilemma of statistical learning theory.

A popular approach to visual learning and recognition is based on appearance-based methods in which objects and scenes are represented as collections of viewpoint-specific features. Often the representations are built using subspace methods, such as Principal Component Analysis (PCA). These methods have proven to be successful both for learning and recognition under constrained conditions. However, their full potential in unconstrained natural environments has not been exploited, mainly due to several shortcomings of these methods in their standard forms. The standard subspace methods are intrinsically non-robust; they are sensitive to shift, scale, rotation, illumination changes, cluttered background, and occlusions. Several extensions of the standard approach have already been proposed with the aim to overcome some of these shortcomings (i.e. methods for robust recognition, robust learning, illumination insensitive recognition etc). However, these methods are still limited to the specific problems and domains and work well only under certain assumptions. Within this project novel more general robust methods will be developed, which will consider new principles (local/global approach, the combination of the classifiers) and will be able to work in uncontrolled environments as well.

The problem of describing and recognizing object categories, as opposed to specific objects (e.g. a specific spanner), has recently gained some attention in the machine vision literature (e.g. Agarwal & Roth 2002,Amit & Geman 1999, Borenstein & Ullman 2002, Weber et al 2000). Examples include faces, cars, and horses. There is broad agreement on the issue of representation: object categories are represented as collection of features, or parts, each part has a distinctive appearance and spatial position. Different authors vary widely on the details: the number of parts (from a few to a thousand), how these parts are detected and represented, and whether the variability in part appearance and position is represented explicitly or is implicit in the details of the matching algorithm. The issue of learning is perhaps the least well understood. Most authors rely on manual steps to eliminate background clutter and normalize the pose of the training examples.

The first innovation here is to build visual models for which both learning and recognition is affine invariant. This means that objects can be learnt under far less restrictive circumstances than at present – it will not be necessary to scale normalize the images or arrange the object surfaces to be parallel to the image plane. The second innovation is to represent the object by a heterogeneous collection of parts – for example by curves and appearance patches. Currently most learnable models use a representation built on appearance patches alone. However, allowing the model to learn the representation type that is most appropriate for it again gives greater flexibility.

1.3.4 Cognitive modelling of the action of seeing

The highly effective Visual-search (or more generally, attention) mechanisms in the human visual system were extensively studied from Psychophysics and Physiology points of view. Several search mechanisms were implemented, usually in order to model the HVS. Some implementations focused on computer vision applications. Those studies usually suggest to direct attention based on bottom-up or top-down information about the target objects. i.e. - using measures of saliency in order to give a higher attention priorities to some locations in the scene, or using prior information about features of the sought for object in order to decide which locations are more likely to contain targets. The role of inner-scene similarity, which was proven as essential in the human visual system was not studied in the computer-vision context.

Usually, systems based on approaches suggest so far suggest to pre-calculate a static map of priorities before the search starts which pre-specifies the scan order. We suggest to make the search more dynamic in the sense that it allows the saliencies, or priorities, to be changed based on the results of the recognition process applied to already attended locations.

In addition, relatively little was done to quantitatively characterize the inherent difficulty of search tasks, which we consider as our main goal.

One focus of ongoing research is on the development and usage of associative learning structures for complex response generation. The past work has demonstrated that a so-called channel representation of features and parameters allow very fast and efficiently implemented methods for learning. These methods are now being refined and further developed to allow more complex behavior of the integrated system.

A second focus is on compact representations of low and medium level features for cognitive vision systems, mainly using various tensor representations. The work is based on previous results on tensor representations of local orientation made at the laboratory , but goes beyond these by considering tensors of higher order which in addition are defined on more complex spaces. One objective is to define representations which can describe complex local features, e.g., corners and closed shapes, providing robust medium level features both for 2D and 3D analysis of images and scenes.

1.3.5 Functional imaging for observing and modelling brain activity

Exploring simultaneously computer and biological vision is a great and very exciting challenge. We believe that a better knowledge of the mechanisms of human and animal visual perception can have a significant impact on algorithm design, performance evaluation and interfacing of computer vision systems with humans.

At a more general level, biological visual perception, in particular in humans and monkeys is not well understood not to mention modelled. Pushing the state of the art of this understanding is a great scientific and philosophical challenge that shapes the general framework of our research.

In revisiting the important problem of image segmentation, one would like to address also the feature integration problem from the computational and biological viewpoints. The interest of combining different cues for image segmentation has been stressed since many years but it's only recently that the variational framework has been applied to successfully integrate different cues and models. Consequently, we are currently investigating the idea to extend our previous work and propose a variational framework to integrate some of the most important cues in image segmentation, namely grey level, texture, color and motion information. Another fundamental point will be the one associated to the use of geometric priors to guide the segmentation process. Incorporating prior knowledge in the segmentation process is a very challenging problem that we would like to address too.

Regarding the area associated to the functional imaging for observing brain activity, one would like to develop processes that will help us to better understand the organization of the brain with the help of DT-MRI. Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) is an exciting new magnetic resonance technique measuring the rate of water diffusion in biological structures such as white matter tissues in the brain. This new imaging modality provides non invasive means to have a better knowledge of the structural organization of the white-matter fiber bundles in the brain and has already a large number of clinical applications It can also be used for a better understanding of the neurobiological functions of the brain by improving the localization of the brain activity sources from EEG and MEG measurements through a better estimation of the electrical diffusivity tensor. A robust estimation and regularization of diffusion tensor images from noisy diffusion gradient raw data sequences is an exciting and challenging problem that we would like to address using a new multidimensional variational framework. This could be extended and applied to the objectives of Diffusion tensor trajectography and electrical conductivity tensor estimation for the purpose of EEG and MEG localization.

Regarding the Modelling brain activity, some exciting work has already been done in theoretical biological vision where researchers propose computational models of some visual pieces of competence, of various visual areas and their interactions. However, it's still a challenging problem to propose processes that implement these theories and help comparing the results with those of more traditional computer vision algorithms. It's also a very challenging and exciting problem to try to link some of these biological theories to the PDE framework. If successful, we believe that this work will shed a new light on the underlying biological computational processes as well as allow a more formal analysis of their competence and perhaps their shortcomings.

1.4 Research method

The research objectives listed above will addressed under a number of **doctoral** and **post-doctoral** studies. A doctoral study will be carried out by a PhD student (or an early-stage researcher) under the supervision of three researchers from three VISIONTRAIN partners.

A post-doctoral study will be carried out by an experienced researcher in collaboration with at least two researchers from two VISIONTRAIN partners.

Therefore we plan to recruit 11 early-stage researchers and 16 experienced researchers. The recruitment, training, and transfer of knowledge activities are described in detail in section 2.

The following is the list of 11 PhD topics that will be addressed by this training network. Table 1 below summarizes the research objectives to be addressed within each-one of the PhD proposals. **Although the PhD topics are described in detail, a lot of freedom will be given to the early-stage researchers in order to express their own creativity and develop original approaches and methods**.

The research method for experienced researchers is somehow different. As explained in detail in section 2.3.2, each recruited researcher will write a research proposal in collaboration with two network researchers and addressing the VISIONTRAIN objectives.

PhD proposal	Objective 1	Objective 2	Objective 3	Objective 4	Objective 5
Phd #1		•	•		
Phd #2	•	•			
Phd #3	•		•		
Phd #4				•	•
Phd #5			•	•	
Phd #6	•				•
Phd #7	•		•		
Phd #8			•	•	
Phd #9			•	•	
Phd #10	•	•			
Phd #11		•	•		

Table 1: The research objectives addressed by each PhD proposal.

PhD #1: Layered representation of videos using motion-based segmentation. The research will build on recent video segmentation methods using layers. When a moving camera observes moving objects and persons, the only way to determine the camera motion is to detect the "static layer" – the image pixels that correspond to static objects in the scene. The basis for properly detecting the static layer is to estimate the camera motion. This will be done using knowledge about the camera geometry undergoing general rigid motion and based on matched features. A static layer will be associated with an image sequence corresponding to a specific camera motion. In the case of a rotating and zooming camera, the static layer is just a panoramic image. In the case of a translating, rotating, and zooming camera, the static layer will be a dense depth map, and hence, each pixel in the static layer will be described by both color and depth. Moving objects will form the dynamic layer. It will be possible to find moving objects without explicitly determining their motion parameters. This is particularly important when both deformable and articulated objects are present in the scene. Dynamic layers (sets of image pixels associated with moving objects and persons) will be associated with short image sequences thus facilitating the task of recognition.

PhD #2: Geodesic active segmentation and integration. The interest for combining different cues for image segmentation has been stressed for many years. Therefore, a first objective is to propose a variational framework to integrate some of the most important cues in image segmentation, namely grey level, texture, color, and motion information. The second objective is to understand how to acquire the information from images. The variational integration and segmentation framework which will be investigated consists of two steps: pre-segmentation based on non-linear diffusion to improve the quality of the features and a variational framework for vector-valued data using a level-set approach as well as a statistical model to describe the interior and the complement of a region.

PhD #3: Shape Priors for Detection, Segmentation and Recognition. An important focus of recent research in the area of image segmentation has been the use of geometric priors to guide the segmentation process. When segmenting medical images, for example, it is often the case that a rich body of anatomical knowledge is available concerning the average shapes of the regions of interest. Furthermore, it has been shown recently that statistically encoding the 2D-appearance of 3D-objects considerably enhances the capability of computer vision systems to detect, segment, recognize and track objects in unstructured environments. The objective of this PhD-project is to study mathematical shape representations in view of the trade-off between cognitive and computational viewpoints, the feasibility of statistical learning of shape variations from examples (e.g. appearance of humans or anatomical variability in medical imaging), and the combination of shape priors with variational approaches to image segmentation.

Phd #4: Towards a neural model-based segmentation The starting point of this thesis will be a critical study of some of the work in theoretical biological vision where researchers propose computational models of some visual competence, of various visual areas and their interactions. We would like to implement some of these theories and compare the results with those of more traditional computer vision algorithms. From the theoretical viewpoint these theories are based on the interactions of elementary neurons and result in signal processing-like formal descriptions of the corresponding computations. We would like to try to link some of these biological theories to the PDE framework. If successful we believe that this work will shed a new light on the underlying biological computational processes as well as allow a more formal analysis of their competence and perhaps their shortcomings.

PhD #5: A computerized attention system for visual search. Given an image, we plan to extract a number of sub-images which are the candidates for attention. The identity of the candidates will be modelled as a set of correlated random variables taking target/non-target values. It will be assumed that the correlation between two candidate identities depends on the similarity between them: more similar candidates (sub-images) tend to have more similar identities. A search algorithm will be proposed based on probability estimation of the identities, which relies on the above assumption. This assumption will be tested by experimentally studying the behavior of identities correlation vs. similarity for objects appearing in the same scene. In addition, measures that will quantify the difficulty of search tasks will be investigated, by analytically proving that those measures bound the performance abilities of all search algorithms.

PhD #6: The perception of three dimensional shape, the light field and material properties as seen in flat pictures. We propose to work on the perception of three dimensional shape, the light field and material properties as seen in flat pictures (e.g., photographs of real scenes, processed photographs of real scenes, or computer graphics). When the human observer is confronted with a picture a perception of what is "in" the picture requires the simultaneous estimation of the geometrical layout of the scene, the light field in the scene, and the material properties of the objects in the scene. Although this problem remains unsolved in computer vision, we are not proposing to work on a theory that might eventually lead to a solution, but we propose to study the competence of the human observer empirically. In order to do this we need to develop a number of novel psychophysical techniques that will be of value by themselves. The results will lead to a more realistic understanding of the capabilities of normal human observers. This is useful in ergonomic contexts, but it will also lead to fresh ideas for theory in computer vision and graphics.

PhD #7: Segmentation of biomedical and satellite images. In biomedical and satellite images often it is important to detect and segment small features irrespective from their meaning. In this case the number of regions forming the segmentation can be larger than 100 and possibly of the order of some thousands. This segmentation will be referred as a fine segmentation, where all detected details are kept. In other instances it is useful to have a segmentation of the image in a smaller number of distinct regions, corresponding to features to be identified, such as the sky in outdoor scenes, the eyes, nose and mouth in faces, cars and pedestrians in street scenes. In this case the number of regions of the segmented image could be around 100 or so. This segmentation will be referred as a medium segmentation, where major important features are detected. In other instances it may be useful to have a coarse segmentation of the image in a very few number of distinct regions, i.e. from 5 or so to two dozens. Therefore : we will envisage best methods for obtaining fine, medium and coarse segmentation of biomedical and satellite images; we will investigate methods for merging regions based on a variety of similarity criteria and heuristics; we will investigate self-adaptive segmentation in which the number of distinct regions in a segmented region adapts to the global image features.

PhD #8: Robust approach to subspace-based visual learning and recognition. A popular approach to visual learning and recognition is based on appearance-based representations, which are often built using subspace methods. Within the doctoral dissertation novel more general robust methods will be developed, which will be able to work in uncontrolled environments. The research will be focused on two main topics. Firstly, the robustness will be increased by combining the holistic global approach and local approaches by considering structural information as well. Secondly, the recognition and classification will be further strengthen by combining generative and discriminative subspace-based classifiers. Such combined methods should significantly improve the results of the subspace-based visual learning and recognition.

PhD #9: Learning and recognition of objects and categories The PhD will investigate learning and recognizing visual object models with a specific architecture. The object model will consist of a number of parts. Each part has a description (for example its appearance or geometric shape) and relative scale. The object shape is represented by the spatial layout of the parts. The entire model is generative and probabilistic, and appearance, scale, shape and occlusion will all be modelled by probability density functions. The process of learning an object category is one of first detecting regions and their scales, and then estimating the parameters of the above densities from these regions, such that the model gives a maximum-likelihood description of the training data. Recognition is performed on a query image by again first detecting regions and their scales, and then evaluating the regions in a Bayesian manner, using the model parameters estimated in the learning.

PhD #10: Stereoscopic matching using super-stable graph kernels. Obtaining dense and accurate correspondences among a set of images without restricting the application domain still possess a challenging task. The aim is to achieve (1) accuracy in stereoscopic matching, i.e. to guarantee gross-error probability to be less than a pre-defined level in unrestricted scenes and (2) high precision, i.e. to avoid usual interpolation artifacts under the presence of weakly textured objects. The game-theoretic allows s to meet these criteria directly, especially the (1). The matching problem is formulated as one of finding a super-stable kernel in a special graph. This formulation is suitable for several matching problem variants (multi-image matching, multi-criteria image similarity, natural integration of multiple cues for stereopsis, dense or sparse tentative matches, single-valued or multi-valued matching, etc). Importantly, it is possible to formulate the problem in a way that a high-level vision process could influence the low-level matching task.

PhD #11: Implementation of PDE-based methods for simultaneous image sequence restoration and motion estimation. Implementation of PDE-based methods for image restoration based on statistics of natural images and least committed motion models. Investigation of their numerical and mathematical properties. Application to restoration of movies that have been damaged due to inappropriate storage. Application to motion estimation in medical images where patient concern makes it impossible to record high quality sequences. This may be gated MR heart sequences or ultrasound imaging.

1.5 Work plan

The work plan for the whole duration of the network is described in detail below. There are seven work-packages. Work-packages 1 to 5 implement the five scientific objectives of the network. Work-package 6 is concerned with dissemination and work-package 7 is concerned with management.

1.5.1 Breakdown in work-packages and overview

WP	Work-package name	Description
WP1	Computational theories and methods for low-level vision	§1.2.1 - page 7
WP2	Motion understanding from image sequences	§1.2.2 - page 8
WP3	Learning and recognition of shapes, objects, and categories	§1.2.3 - page 9
WP4	Cognitive modelling of the action of seeing	§1.2.4 - page 10
WP5	Functional imaging for observing and modelling brain activity	§1.2.5 - page 10
WP6	Network dissemination activities	pages 27-30
WP7	Network management activities	§4.3 - page 66

	Start/End	Partners	Leader (Partner)	Person-months
WP1	0/48	2,3,5,8,9	V. Hlavac (9)	185
WP2	0/48	1,3,4,8,9,11	A. Heyden (3)	152
WP3	0/48	1,4,6,7,11	A. Leonardis (7)	152
WP4	0/48	4,5,6,7,8	M. Lindenbaum (5)	152
WP5	0/48	1,2,5,10,11	V. Torre (2)	152
WP6	0/48	all	R. Deriche (1)	95
WP7	0/48	all	R. Horaud (1)	56

Table 2: Overview of the network work-packages.

Work-package number	Early-stage res.	Experienced res.	Other res.	Total
WP1	99	36	50	185
WP2	66	36	50	152
WP3	66	36	50	152
WP4	66	36	50	152
WP5	66	36	50	152
WP6	33	18	44	95
WP7	-	-	56	56
Total	396	198	350	944

Table 3: Breakdown of efforts (in person-months) for each work-package and for each category of researchers.

	Number of ESR (PhD topic numbers)	Number of ER	Total
WP1	3 (#2, #7, #10)	3	6
WP2	2 (#1, #11)	3	5
WP3	2 (#3, #9)	4	6
WP4	2 (#5, #8)	3	5
WP5	2 (#4, #6)	3	5
Total	11	16	27

Table 4: Number of early-stage and experienced researchers for each workpackage.

1.5.2 Graphical presentation of the components and milestones

Figure 1 below is a graphical representation of the work-packages. The milestones correspond to coordinated network achievements and are described in detail in this paragraph.

M0.a Kick-off meeting. Election of committees. Official start of the network and of work packages. Take official decisions concerning the recruitment of early-stage researchers (*The recruitment procedure starts during the negociation phase*).

M0.b Start the PhD programme. Finalize contracts with 6 early-stage researchers. Officially allocate them to partners and to work-packages.

M1.a Start the experienced-researcher recruitment procedure. Publicize the ER opportunities on VISIONTRAIN's website.

M1.b Start the training and transfer of knowledge activities.

M6.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M6.b Internal audit of early-stage researchers

M12.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M12.b The first annual industrial meeting.

M12.c The first annual thematic school.

M12.d Finalize contracts with 5 early-stage researchers. Officially allocate them to partners and to work-packages.

M12.e Annual PhD reports by early-stage researchers.

M12.f The first network review meeting.

M12.g Update the network technical annex.

M12.h Internal audits of early-stage and exp. researchers.

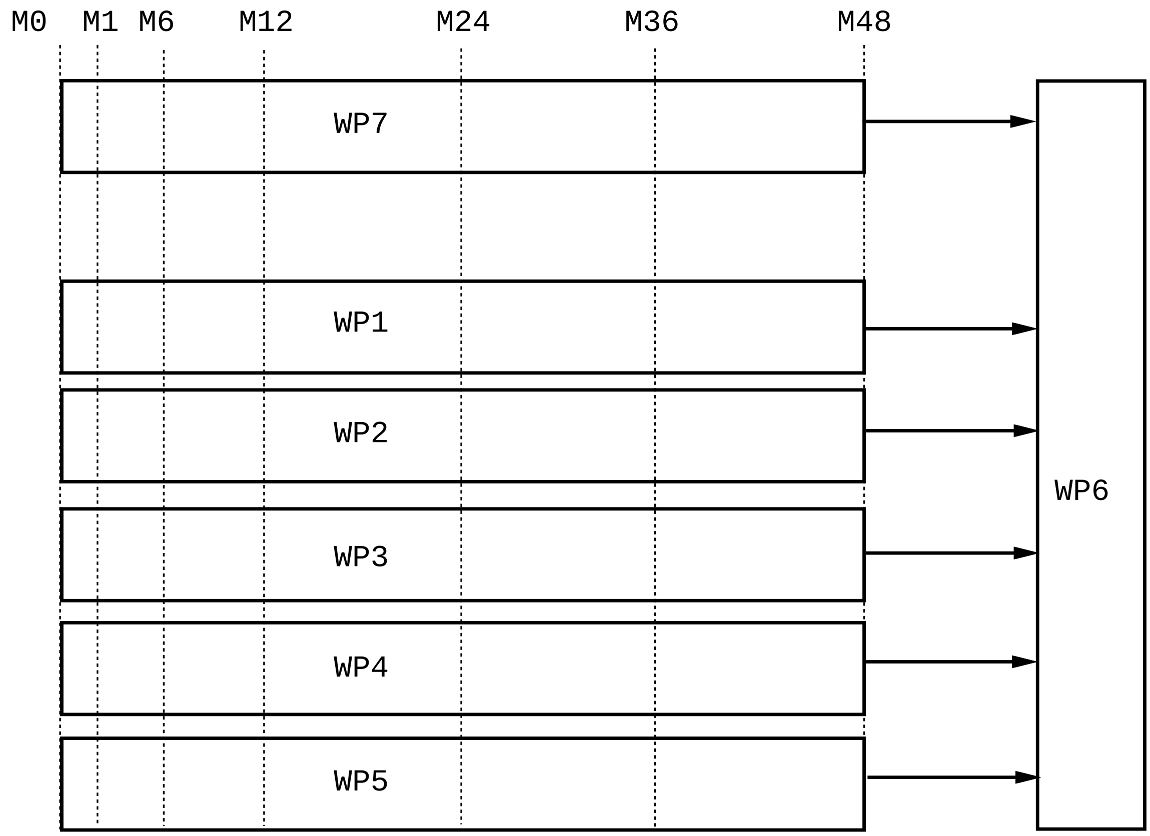

Figure 1: Graphical presentation of work packages.

M18.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M24.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M24.b The second annual industrial meeting.

M24.c The second annual thematic school.

M24.e Annual PhD reports by early-stage researchers.

M24.f The second network review meeting.

M24.g Update the network technical annex.

M24.h Internal audits of early-stage and exp. researchers.

M30.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M36.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M36.b The third annual industrial meeting.

M36.c The third annual thematic school.

M36.d Final PhD examination of 6 early-stage researchers

M36.e Final (6) and annual (5) PhD reports by early-stage researchers.

M36.f The third network review meeting.

M36.g Update the network technical annex.

M36.h Internal audits of early-stage and exp. researchers.

M42.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M48.a Bi-annual meetings of Executive Committee, Management Board, Scientific Council and IPR Council.

M48.b The fourth annual industrial meeting.

M48.c The fourth annual thematic school.

M48.d Final PhD examination of 5 early-stage researchers

M48.e Final (5) PhD reports by early-stage researchers.

M48.f The final network review meeting.

M48.g Internal audit of experienced researchers.

1.5.3 Detailed description broken down into work-package tasks

Work-packages 1, 2, 3, 4, and 5. These work-packages are broken down along the same structure wich is described in detailed below:

Task I.1: Individual work of early-stage researchers towards the completion of their PhD. This task is done in coordination between the three host partners involved in each one of the PhD topics. The task includes the preparation of internal reports and audits as well as the preparation of the final PhD dissertations.

Task I.2: Individual work of experienced researchers. The task is done in coordination between the partners hosting the experienced researchers. The task includes the preparation of internal reports and audits as well as the preparation of the final reports.

Task I.3: Collaborative work between ESR, ER, and the other researchers through the various training and transfer of knowledge activities: intra-network working groups and visits, proof-of-concept activities, etc.

Work-package 6: Dissemination. The following tasks will be active during the lifetime of the project, in accordance with the training and transfer of knowledge activities:

Task 6.1: Organisation and participation to the annual thematic schools.

Task 6.2: Organisation and participation to the annual industrial meetings.

Task 6.3: Organisation and participation to the *Women and science* programme.

Task 6.4: Participation to workshops, conferences, and tutorials.

Task 6.5: Patent application activities.

Work-package 7: Management. The following tasks will be active during the lifetime of the project. They correspond to the management activities described in section 4.3 at page 66:

Task 7.1: Setup and maintain a network infrastructure

Task 7.2: Communication with the EC and coordination of reporting

Task 7.3: Organisation of external audits

Task 7.4: Organisation of internal evaluation

Task 7.5: Financial management

Task 7.6: Technical management

Task 7.7: Legal and knowledge management

Task 7.8: Human resources management

1.5.4 List of deliverables

The network will produce the following sets of deliverables:

- Four annual progress reports and one final report – PR1, PR2, PR3, PR4, and FR. These reports will be based on the following deliverables;
- Each early-stage researcher will deliver a progress report at months 6, 12, 24 and a final report at month 36. The final report corresponds to the PhD dissertation – TIR (thesis interim report) and TFR (thesis final report).
- Each early-stage researcher will deliver a proof-of-concept demonstration and its associated software at month 24 and 36 of her/his PhD;
- Each experienced researcher will deliver a progress report (ERPR) at month 6 and a final report (ERFR) at month 12;
- Each work-package (1, 2, 3, 4, and 5) will deliver a progress report at months 12, 24, 36, and 48 as well as a final report – WPPR (work package progress report) and WPFR (work package final report).

- Each work-pacakge (1, 2, 3, 4, and 5) will deliver a proof-of-concept demonstration and its associated software as a result of collaborative work between network researchers and associated partner researchers;
- Work-package 6 (dissemination) will deliver progress reports at months 12, 24, 36, and 48 as well as a final report – DPR (dissemination progress report) and DFR (dissemination final report). The dissemination reports will contain the followings:
  - Description of the thematic schools, their programme, courses, number and list of network researchers who attended the schools, number and list of non-network researchers who attended the schools.
  - Description of the annual industrial meetings, abstracts of presentations, number and list of companies who attended the meetings, number and list of network researchers who presented their work at the meetings.
  - Description of the "Women and science" activity. Percentage of women recruited as early-stage and experienced researchers, names of women who co-authored publications, description of special forums organized by the network or forums attended by network researchers.
  - List of publications by network researchers to international conferences and workshops.
  - List of tutorials organized by network researchers.
  - List and description of patents applied by network researchers.
- Work-package 7 (management) will deliver management progress reports (MPR) every 6 months (following the bi-annual meeting of the executive committee) as well as a management final report (MFR).

2 Training and transfer of knowledge activities

2.1 Transfer of knowledge programme

The transfer of knowledge activities of this network are to carry out 11 PhD theses as well as to provide to 16 experienced researchers the opportunity to develop their own methods and approaches in collaboration with network researchers.

Each early-stage researcher to be recruited during the first and second years of the project will have a main partner of attachment and two secondary partners. Therefore, out of 36 months, she/he is expected to spend 30 months with the main partner and 3 months with each one of the secondary partners.

In parallel, the network will provide to 16 experienced researchers the opportunity to spend 12 months with two network partners (on the basis of 6 month periods) and this activity will be synchronized with the training of the early-stage researchers.

The proposed assignment of the main and secondary partners for each PhD thesis preparation is given in Table 5. This assignment maximizes exposure of each early-stage researcher to three different European groups with complimentary expertise.

Each early-stage researcher will have a PhD advisor associated with the main partner and two secondary advisors.

Each experienced researcher will work in collaboration with at least two researchers from two different partners.

The cross-visits of both early-stage and experienced researchers will be synchronized such that small groups of people are hosted by one partner and work together for periods of time of approximatively three months.

PhD proposal	Main partner	Sec. partner	Sec. Partner
Phd #1	8 (EPFL)	3	1
Phd #2	3 (U. Malmo)	8	2
Phd #3	11(U. Mannheim)	1	7
Phd #4	1 (INRIA)	10	11
Phd #5	5 (Technion)	4	6
Phd #6	10(U. Utrecht)	2	5
Phd #7	2 (SISSA)	9	3
Phd #8	7 (UOL)	6	8
Phd #9	6 (U. Oxford)	7	4
Phd #10	9 (CTU)	5	2
Phd #11	4 (IT-UC)	11	9

Table 5: The main and secondary partners associated with each PhD proposal

In more detail, the transfer of knowledge programme will feature the following activities:

2.1.1 Intra-network working groups and visits

We firmly believe that fine-scale person-to-person interaction is indispensable for building a sense of collaboration and community, so one of the most important activities of the network is to favor interactions and small working groups. This activity is structured around groups of the following categories: one or two early-stage researchers, their supervisors, one or two experienced researchers, and a few other researchers.

Such a group of researchers will be hosted by one partner for periods of time of typically three months.

2.1.2 Proof of concept

This activity addresses the problem of practically building vision systems. One can easily identify applications that need computational and cognitive models integrated to the system in order to achieve the required capabilities. Nevertheless, rather than attacking a specific application, this activity will focus on the development of **generic platforms**. The latter can be split into four broad categories:

- *Real-time systems*: Typically such a system will address the case of one or two cameras (plus other types of sensors) linked to a single PC. The computing power available today can potentially process 10 to 25 images per second. Nevertheless, the real-time implementation of a complete intelligent system remains an open issue.

- *Distributed systems*: Many applications will demand a large number of sensors – a traffic monitoring system, for example, may use up to 20 cameras – and therefore a cluster of PC's or a distributed PC network will be required. The development of a distributed vision system is a challenging task in itself involving knowledge from many scientific and technological fields. Therefore it requires expertise from different research teams, such as multiple-camera acquisition and processing, distributed control, parallel and distributed algorithms, etc.

- *Embedded systems*: The latest technical developments have made it possible to integrate (even on the same chip) a processor together with a sensor. This opens a wide range of new applications, where a lot of computations can take place inside the sensor instead of off-line. Examples are surveillance and face recognition systems as well as text recognition systems embedded in portable devices. Of special interest is to investigate how performance of the computer vision and cognitive vision algorithms can be optimized given the constraints on processing power and memory.

- *Bench-marking*: The network will systematically work towards identifying the best solutions for implementing real-world vision systems. Bench-marking will be carried out for several categories of techniques and systems built upon them, from low level-visual cues to cognitive modelling, with increasing degree of difficulty while the work progresses. We recognize the challenge of designing test cases in a manner that show in a fair and reliable manner the potential and capabilities of cognitive and computational models, but also see this as a research problem itself.

Therefore, **proof of concept activities** will focus on *software development*, *system integration*, and *bench-marking*. In practice the following types of network actions will be made available:

- A group of early-stage and experience researchers together with a number of network researchers from different partners decide to work **cooperatively** and **remotely** towards the development of a **unique** system. They plan to integrate pieces of software from various partners, develop new software, integrate their contributions within a single hardware/software platform, etc.
- Early-stage and/or experienced researchers from one or several network partners wishing to visit a network team in order to undertake collaborative work that needs **day-to-day presence** of **all researchers** involved in the action. In particular this type of action will be adopted whenever the early-stage and experienced researchers need to use sophisticated hardware/software equipment and the associated human expertise available with one network team.

The **proof of concept activities** will be launched in conjunction with the network's work-packages. It is planned that each early-stage and experience researcher will participate to at least one proof-of-concept action during her/his network training.

2.1.3 Spread of excellence

During the lifetime of the project a number of conferences, workshops, tutorials and schools will be organized through Europe. In particular, one of the partners (Czech Technical University, partner 9) will be in charge of organizing the Eighth European Conference on Computer Vision in 2004 (ECCV'04) and another partner (INRIA, partner 1) will be engaged with organizing the Tenth IEEE International Conference on Computer Vision in 2005 (ICCV'05) to be held in China. Both ICCV and ECCV are the most prestigious conferences in computer vision. Around 10 workshops and tutorials are organized in conjunction with these main events.

VISIONTRAIN will actively participate to these events and to forthcoming ECCV'06, ECCV'08 and ICCV'07 since the partners involved in the network actively participate to these conferences, workshops, and tutorials.

Both early-stage and experienced researchers will be encouraged and trained to submit papers and to participate to these prestigious conferences as well as to other European and international conferences.

2.1.4 Thematic schools

We plan to launch a new event named the *European Computer Vision School* or ECVS. ECVS will take place once a year in conjunction with another major event such as ECCV, ICCV (whenever the latter is hold in Europe) or any other major European conference in the fields of computational and cognitive vision.

ECVS will function on the following grounds. There will be one session per year. The expected duration of each session is of 5 to 10 working days. The programme will be a combination of tutorial courses provided by experienced researchers and seminars provided by early-stage researchers involved in the network. There will be at least 5 tutorial courses and 10 seminars per session. Since we expect 11 early-stage researchers and 16 experienced researchers to be trained, each one will have the opportunity to present her/his work once or twice during the lifetime of the project. In parallel with these technical programme, other tutorials will addressed issues such as: intellectual property rights, software commercialization and dissemination, patent applications protocoles, success stories, etc.

Currently the following ECVS sessions are already planned:

- 2004 – Prague (Czech Republic), in conjunction with ECCV'04, professor Vasek Hlavac and Dr. Martin Urban
  Topic: Low-level image formation and understanding, crossroad of biological and computational methods .
- 2005 – Sophia-Antipolis (France), Dr. Rachid Deriche
  Topic: Variational methods and their usefulness for vision.
- 2006 – Ljubljana (Slovenia), possibly in conjunction with ECCV'06, professor Ales Leonardis
  Topic: Learning and recognition of objects and categories.
- 2007 – Malmo/Copenhagen (Sweden/Denmark), professors Anders Heyden and Mads Nielsen.
  Topic: Geometrical computer vision.

NB. The whole **Thematic school activity** may be shifted forward by one year to match the starting of the VISIONTRAIN network.

2.1.5 Women and science

This activity will offer to VISIONTRAIN early-stage and experienced researchers the opportunity to organize and to participate to forums, seminars, and workshops dedicated to the role of women in the production and dissemination of scientific and technological knowledge. These activities will be coordinated in collaboration with TWOWS, the *Third World Organization for Women in Science* located in Trieste, Italy – the hometown of SISSA (partner 2).

The VISIONTRAIN researchers and management also plan to establish an observatory in order to measure the role of women in the particular scientific field addressed by the network. One major feature of VISIONTRAIN is that equal opportunities will be given to women and men: **An equal number of women and men will be recruited by the network**. The network management will be responsible of advertising all the VISIONTRAIN job opportunities to specialized *women and science* international organizations, web-sites and newsletters.

2.1.6 European patent application

Recent experience in image processing, video coding, storage and transmission, multi-media technology, virtual and augmented reality, and so forth has shown that an ambitious patent application policy represents a true added value to companies, but also to research institutes and universities.

We believe that we are at a stage of technological development when the differences between hardware and software become fuzzier and fuzzier. This is particularly true for cognitive and computational vision systems which will integrate knowledge from many disciplines.

Therefore, VISIONTRAIN will launch an activity specially designed to prepare and submit patents to the *European Patent Office* – EPO – in Munich. The latter includes 27 European member states. In 2001 the EPO received 150,000 patent applications out of 700,000 patent applications received worldwide.

Formal relationships will be established with the European Patent Office such that immediate feedback and advice can be obtained from this Office. In particular, a former INRIA PhD student – Andreas Ruf who was granted a Marie Curie TMR20 fellowship in the period 1998-2001 – is now a patent examiner for the EPO in the fields of computer vision and computer graphics.

Early-stage and experienced researchers wishing to apply for a patent will proceed as follows. A short description of the "invention" is forwarded to her/his advisor or network researcher. The latter will play an interface role between the researcher(s) wishing to apply for a patent, the *Intellectual Property Rights Council*, and if needed, an external patent consulting company. When the above cited parties come to the consensus that the invention should be patented, a request for funding is forwarded to the *Executive committee*. The latter will take a decision and will allocate specific budget for covering the expenses associated with the patent preparation and application.

2.1.7 Technology transfer to companies

One important aspect of the network training and transfer of knowledge programme is to give to both early-stage and experienced researchers the opportunity to establish formal relationships with companies interested in vision systems. Indeed, past and recent experience of the partners revealed that there is a tremendous technological potential associated with the VISIONTRAIN topics of research.

The partners will prepare an extensive list of European companies potentially interested in technologies and applications associated with vision systems. These companies will be regularly informed of the VISIONTRAIN activities. In particular, the list of early-stage and experienced researchers with their research objectives will be made available to the companies' contact points.

At the end of each year the companies will be invited to attend an *Industrial annual meeting*. During these meetings, the network will be given the opportunity to present their results.

Each ESR and ER will be granted with a travel budget allowing her/him to visit one or several companies during her/his training and establish formal work relationships.

2.1.8 Summary of the transfer of knowledge programme

To conclude, the transfer of knowledge programme will have the following characteristics:

- The network plans to finance 11 early-stage researchers and 16 experienced researchers
- Each early-stage researcher will spend 30 month with her/his main partner of attachment and 3+3 months with two other partners.

- Each early-stage researcher will have an advisory committee of 3 researchers from the 3 partners.
- Each early-stage researcher will meet her/his advisors at months 0, 6, 12, 24, and 36 (for the PhD defence). This group of people will have the mission to make sure that the scientific work advances according to a pre-established programme, to insure that the researcher is well integrated within the host teams and that all the resources are made available to her/him, and to envision the career of the PhD candidate after the thesis defence.

  In particular they will advise the candidate to attend courses and seminars directly related to her/his integration in the European research community and to help the candidate plan her/his career, orient her/him towards the academia or towards industry, etc.
- Each experienced researcher will spend 6+6 months with two network partners and will work collaboratively with at least two network researchers.
- Each early-stage and experience researcher will be assigned a tutor among the members of the *Scientific board* or the *Management board* of the network. Each researcher will have the opportunity to refer to her/his tutor whenever she/he feels that there is necessary.
- In addition to the main research activity, each early-stage researcher will attend courses from a list of courses established with her/his advisors.
- Each early-stage and experienced researcher will be assigned a *teaching assistantship* in coordination with her/his main advisor.
- Each early-stage researcher will present her/his research at months 6, 12, and 24 in addition to her/his final PhD examination at month 36. These presentations will be attended by the 3 advisors as well as by the tutor.
- Each experienced researcher will present her/his research at month 6 and at month 12. These presentations will be attended by the 2 advisors as well as by the tutor.
- Each early-stage researcher will attend ALL sessions of the *European Computer Vision School – ECVS* and will present her/his work twice.
- Each early-stage and experienced researcher will be expected to publish her/his work at major journals and conferences.
- Early-stage and experienced researchers will be provided with support to submit patents to the European patent office.
- Early-stage and experienced researchers will work collaboratively with network researchers within an intra-network mobility activity.
- Both early-stage and experienced researchers will be encouraged and provided with support to participate to proof-of-concept collaborative activities.

- Each early-stage and experienced researcher will have the opportunity to present her/his results to companies both at the network *Industrial annual meetings* or through individual company visits.
- A women and science activity will be supported.

2.2 Impact of the training and transfer of knowledge

Systems for handling and understanding perceptual information are expected to have as great impact on society over the next decades, as what conventional computers and telecommunication have on today's society. They promise to relieve humans of many burdens in the use and the communication with increasingly complex systems, be they technical or deriving from an increasingly complex society. They will make many new applications possible, ranging from autonomous home appliances to intelligent assistants keeping track of the operations in an office.

Nowadays, scientific findings, technological innovations, and society progress are intertwined. We believe that the development of vision systems is one prominent example of such a coupling between science and technology on one side, and society on the other side. More precisely, the understanding and design of computational and cognitive models for vision systems is an excellent opportunity to propose a global cognitive architecture with many theoretical and practical implications.

Let us remark that images and videos are everywhere: There are tens of thousands of video cameras used for security-, traffic-, medical-, sport-, land-monitoring, and so forth. Digital cameras, video recorders, and web cameras are sold by millions every year. More recently, cameras became available with portable and mobile devices, robots of various kinds, cars, homes, training sites, conference rooms, etc.

The advantages of cameras in order to gather information about the surrounding physical world over other sensing devices are numerous: they are compact, robust, flexible, light-weight, high-resolution, wide-range (from a few centimeters to hundreds of meters), and cheaply produced. Moreover, they capture incredibly potentially rich data in terms of geometry, light, color, and semantics.

However, the technologies widely available today, enabling to process this vast amount of information, are currently limited to data compression and storage, data telecommunication, and data visualization. For example, video compression made tremendous progress in the last decade (often in collaboration with cognitive and computational vision teams). **It is now time to enable the extraction and representation of meaning from images and videos in particular, and from sensorial input in general**.

VISIONTRAIN will address a number of challenging and unsolved research topics through cross-fertilization between brain, cognitive, information, and mathematical sciences. Multidisciplinary research is more difficult to achieve than single-topic research but, if successful, will have a great impact on scientific findings because it brings fresh blood from one topic to another.

The overall result will benefit all these disciplines. Thanks to the sophisticated methodologies that have emerged in information and mathematical disciplines, better measurements of the activity of the human brain will be available for analysis in neuroscience and psychophysics. Theories of various cognitive processes developed in computer science can be used to chart possibly related brain operations. Experimental findings in neuroscience can suggest new theories, models and architectures in artificial cognitive system, an interaction very much alike the one between theoretical and experimental physics.

We foresee that this synergy will foster the emergence of a new scientific domain and considerably advance the state of the art at the theoretical and practical levels, a fine example of the tight coupling between science and technology.

To conclude the impact of VISIONTRAIN's programme is twofold:

(i) **Achieve inter-disciplinarity**. *Vision research is and must be an interdisciplinary topic exploring synergies between computer science, mathematical modelling, numerical analysis, information theory, control theory, physics, neurophysiology, psychophysics, and so forth. The cross-fertilization between these disciplines will be achieved at the European level.*

(ii) **Raise and maintain European vision research at a top level:** *Increase the number of researchers specialized in the field and give them opportunities to be recruited by European institutes and universities.*

	Person-months for financed researchers			Other researchers	
Partner	Early-stage (a)	Experienced (b)	Total (c)=(a)+(b)	Number (d)	Person-month (e)
1	36	18	54	15	88
2	36	18	54	3	30
3	36	18	54	9	40
4	36	18	54	2	24
5	36	18	54	4	20
6	36	18	54	3	36
7	36	18	54	5	24
8	36	18	54	2	18
9	36	18	54	6	22
10	36	18	54	3	12
11	36	18	54	5	36
Total	396	198	594	56	350

Table 6: Total number of researchers and person-months involved in the project

2.3 Planned recruitment

The VISIONTRAIN network plans to recruit 11 early-stage researchers (ESR) and 16 experienced researchers (ER). Each ESR will prepare a PhD along the lines described above and for a duration of 3 years. Each ER will have a 1 year appointment with two partners.

2.3.1 Recruitment of early-stage researchers

Upon acceptance of this RTN by the European commission, the *Executive committee* will meet in order to prepare the launch of the network. One priority among the priorities is to launch the recruitment of candidates for early-stage training. This recruitment will be performed on the following grounds:

- A VISIONTRAIN website will become active;
- A list of PhD topics will be available on this website;
- For each topic, a list 3 partners that are likely to host the candidate will be available (the main host partner and the two secondary partners).
- A researcher in-charge for each topic will be designated and this contact person's name, address, telephone and e-mail will be associated with each topic for further information and inquiries;
- Potential candidates are encouraged to, first get in contact with the researcher in charge and second to submit their application.
- Each candidate will be invited to submit a complete CV as well as more formal data. The candidates are encouraged to apply for several topics, up to a maximum of 3, and to give an order of preference for these topics.
- There will be two fixed deadlines to apply for an ESR position, the first deadline at the start of the project, and the second deadline after one year.
- Once the call for applications is closed, the *Management board – MB* and the *Scientific and transfer of knowledge board – STKB* will meet and verify the eligibility of the candidates. They will check the adequacy of the applicants' CV with the scientific scope of the network and thus they will draw a list of *admissible applicants.*
- Three referees, one among the VISIONTRAIN researchers and two from the computer vision community, worldwide, will be appointed for each applicant.Clear and specific instructions will be given to the referees such that they use a common set of criteria in order to measure the quality of the applicants.
- After a 2 month period, the MB and STKB will meet again in order to produce a classified list of applicants who are admitted for recruitment. The final recruitment decision is taken by the Executive committee. The following criteria will be used:
  1. The scientific quality of the applicant according to the referees' reports
  2. Up to 30% of the retained candidates could be from third countries;
  3. The female/male ratio should be as close to 1:1 as possible;
  4. No priority will be given to internal mobility – ESR applicants previously trained at one partner wishing to complete a PhD at another partner will have to apply like any other applicants. However, special cases, where such an internal mobility is proven to be necessary, will be taken into account.

- An ordered list of candidates retained for recruitment will be published and a recruitment negotiation phase will start. The ordered list (One woman followed by one men, etc.) will be by 50% larger than the number of open positions.

2.3.2 Recruitment of experienced researchers

The recruitment of experienced researchers will be meant both for internal mobility within the network and for attracting external researchers. Around 50 researchers will be active in the network and 80% of these researchers are entitled to be recruited as experienced researchers. External applications of an excellent quality will be equally treated, especially from candidate and third countries.

There will be 3 deadlines for applying for a VISIONTRAIN experienced researcher position: at the beginning of the first, second, and third years.

Candidates for an experience researcher position will have access to a website describing the research activities of the teams involved in the network, the research objectives of VISIONTRAIN, as well as the list of teams likely to host them.

Each potential candidate will prepare an application with the following items:

- A detailed CV;
- A complete list of publications;
- A description of her/his scientific achievements;
- *A programme of work* describing her/his research objectives and how she/he plans to implement these objectives. Applications will be in conjunction and in association with the PhD topics;
- A list of 2 to 4 VISIONTRAIN researchers from at least 2 different VISIONTRAIN teams with whom the candidate wishes to collaborate.

Once the call for applications is closed, the *Management board – MB* and the *Scientific and transfer of knowledge board – STKB* will meet and verify the eligibility of the candidates. They will check the adequacy of the applicants' programme of work with the scientific scope of the network and thus they will draw a list of *admissible applicants*.

Three referees, one from the STKB and two from the computer vision community, worldwide, will be appointed for each applicant. Clear and specific instructions will be given to the referees such that they use a common set of criteria in order to measure the quality of the applicants.

After a 2 month period, the MB and STKB will meet ago in order to produce a classified list of applicants who are admitted for recruitment. The following criteria will be used:

1. The scientific quality of the applicant according to the referees' reports;
2. The adequacy between the programme of work of the candidate and the VISIONTRAIN objectives;
3. The cross disciplinary nature of the application;

4. Maintain a balance in terms of resource allocation among the partners;
5. Up to 30% of the retained candidates could be from third countries;
6. The female/male ratio should be 1:1, that is the VISIONTRAIN network plans to recruit 8 women and 8 men as experienced researchers.

2.3.3 Publicity

The open positions for both early-stage and experienced researchers will be jointly published by the network. If this proposal is successful, the VISIONTRAIN network will arrange special presentation sessions in conjunction with European and international conferences and workshops in computer vision and related topics.

The communications department of the coordinating partner (INRIA) will build and maintain a recruitment website in coordination with the partners. Links to this website will be made available to other websites specialized in scientific careers.

VISIONTRAIN work opportunities will also be publicized on websites specialized in promoting the role of women such as:

- *The Association of Women in Science*
  http://www.awis.org/
- *The Third World Organization for Women in Science*
  http://www.ictp.trieste.it/ twas/TWOWS.html

All the partners involved in VISIONTRAIN will join their efforts and know-how in order to publicize the opportunities to third country universities and institutes, more specifically the far East (China and India), South and Central America, and so forth.

3 Quality and capacity of the network partnership

3.1 Collective expertise of the teams

The composition of VISIONTRAIN is driven by three main factors: (i) excellence in research, training, and education, (ii) cross-disciplinarity, and (iii) collaboration and coordination of the network activities. VISIONTRAIN is built around 11 partners and 56 researchers. Among these researchers, 52 hold permanent positions with their universities/institutes which is clear evidence that the teams involved in VISIONTRAIN are committed on the long term, beyond the EC funding period.

The collective expertise of the teams is detailed in the table below. The VISIONTRAIN partners have expertise in computer vision, image understanding, cognitive vision, statistical methods and tools, variational and PDE methods, inverse methods, neurophysiology, augmented and mixed reality systems, robotics and control.

Partner	Broad expertise	Specialized expertise
1 INRIA	Computer Science	geometry, variational methods, sensors, optical flow, Markov random fields
2 SISSA	Neurophysiology & Physics	statistical and biological physiscs, neuroscience, machine vision
3 Malmo	Applied mathematics	structure-from-motion, level-set methods, kernel methods, support vector machines, scale-space theory
4 IT-UC	Image and Signal Processing	image analysis, human-computer interfaces film/TV production, medical images
5 Technion	Computer Science	differential geometry, statistcal models, feature grouping, object recogntion
6 Oxford	Engineering science	multiple camera geometry, rendering, invariants learning, category recognition
7 UOL	Comp. & Information Science	visual learning, recognition, range imagery panoramic imaging, multimedia
8 EPFL	Computer Science	human motion capture, computer animation, implicit surfaces, optimization, augmented reality
9 CTU	El. Eng. & Cybernetics	geometry, stereo vision, photometry, physics-based vision
10 Utrecth	Physics and Astronomy	psychophysical vision research, haptics
11 Mannheim	Math. & Comp. Science	statistical modelling, optimization partial derivative equations

Table 7: Collective expertise of the teams.

3.2 Description of partners

3.2.1 Partner 1: INRIA.

The Institut National de Recherche en Informatique et Automatique is one of the largest institutes in the world whose research activity is fully dedicated to computer science, applied mathematics, and control theory. INRIA's headquarters are in Rocquencourt near Paris. INRIA is composed of six research units spread over the French territory. Each research unit has strong relationships and collaborations with universities and their associated research laboratories and PhD programs. Three INRIA research teams will be involved in this network: The MOVI group associated with INRIA Rhône-Alpes in Montbonnot (near Grenoble), the Odyssée group associated with INRIA Sophia-Antipolis in Valbonne (near Nice) and the VISTA group associated with IRISA in Rennes.

The MOVI group

The research team. The MOVI group (Models for computer vision) carries out research in computer vision with four main topics of interest: geometric modelling, feature extraction and matching, learning and recognition, and visual-based robot control. The general philosophy is to (i) establish solid theoretical foundations, (ii) develop methods and algorithms based on mathematical modelling and numerical analysis, and (iii) validate the results through experiments and software development. In particular, the MOVI group has the following contributions: multiple-camera vision, 3D point and surface reconstruction, spatio-temporal tracking based on invariant features, image-based rendering, human motion analysis, indexing large image databases, statistical pattern, texture, and shape learning, visual servoing, and so forth.

Currently the MOVI group is composed of approximatively 18 people (4 full time researchers, 1 assistant professor, 2 development engineers, and 10 PhD students). MOVI is an international group of researchers (1 permanent researcher and 1 development engineer are from Germany, 1 PhD student is from Spain, 1 from Algeria, and 1 from Poland).

Research is done cooperatively within European and National projects. Currently MOVI is involved in 4 IST projects under FP5 (VISIRE, EVENTS, VIBES, and LAVA) as well as in a number of government grants.

In the recent past MOVI achieved a number of successful technology transfers with the following companies: Aérospatiale, Sextant-Avionique, Tecnospazio, The European Space Agency, Eptron, Odense Steel Shipyard, Trivision, Image Systems, Imetric, Heyman-Systems, Plettac Electronics, Siemens, to cite the most important ones.

The associated PhD programme: PhD students working in the MOVI group are enrolled either at the Institut National Polytechnique de Grenoble (INPG) or at the Université Joseph Fourier (UJF). These are both large, renowned universities and important actors on the European level. INPG and UJF have joined their forces to propose a single programme for doctoral studies in the field of "Imagerie, Vision, Robotique" (IVR), as one out of 8 fields grouped together in the "Ecole Doctorale Mathématiques Informatique, Sciences et Technologies de l'Information" (EDMIST). The doctoral programme is organised for a duration of 3 years. As a complement to the research activities, doctoral students must follow a certain number of courses, for a total of 32 ECTS units (The European Credit Transfer System defines academic credits on the transnational level) during the entire three year period. Mandatory courses include:

- Specialized scientific training during the first and second years.
- Training for professional insertion
- General scientific training: scientific writing, language courses, etc.

As part of the proposed program, European fellows will be enrolled in **French language courses** in order to enable them to work and to live in a French language environment.

Specialized scientific training: Courses are chosen with the thesis supervisor, they may be taken from any of the curricula offered in the various fields of the EDMIST. They also may include summer schools or similar scientific events, agreed by the EDMIST. In our particular domains of interest, the courses offered by IVR include: Numerical methods: interpolation approximation, optimisation; Geometry: position, orientation, motion, projective geometry; Computational geometry; Computer vision; Robotics; Virtual imagery; Surfaces and meshes; Reactive and real-time programming Advanced image synthesis; Computational learning; Medical imagery; Motion planning; Physical modelling.

As a complement to these research training activities, fellows will be given the occasion to teach classes to university students; the teaching activity will be carried out in connexion with the CIES (Centre Interuniversitaire d' Enseignement Supérieur).

The proposed activities are designed to prepare early stage researchers not only to their immediate research activity, but also to prepare them to a future career as researcher or academic staff in industrial companies, or in academic institutions on the international level.

Radu Horaud (20%) holds a position of director of research at INRIA, is the scientific leader of the MOVI group and is in charge of European relationships at INRIA Rhône-Alpes. Radu Horaud is born in Rumania and became a French citizen in 1975. After obtaining a PhD in 1981, he started his career at SRI International (Menlo Park, CA USA) where he spent two years and pioneered work on object recognition using range data. In 1984 he joined the Centre National de la Recherche Scientifique. In 1998 he joined INRIA. Radu Horaud has 100 scientific publications mostly in international journals and conferences. His research interests cover both computer vision and robotics. Currently he is an area editor for the "Computer Vision and Image Understanding" journal and a member of the editorial board of the "International Journal of Robotics Research". So far he directed 20 PhD students.

Edmond Boyer (10%) is assistant professor at "Université Joseph Fourier" since 1998. He obtained his PhD from the Institut National Polytechnique de Lorraine in 1996. He started his professional career as a research assistant at the University of Cambridge's Department of Engineering. His fields of competence cover both computer vision and computational geometry. He has research contribution in the domain of 3D surface reconstruction from multiple cameras. Edmond Boyer has 20 scientific publications.

Peter Sturm (10%) holds a position of associated researcher at INRIA. Peter Sturm graduated from the University of Karlsruhe in 1994 and obtained his PhD in 1997 from the "Institut National Polytechnique de Grenoble". He started his professional career at the University of Reading where he spent two years. He joined INRIA in 1999. In 1998 Peter Sturm got the SPECIF PhD award (he was the first one to receive this prize). His research interests cover camera calibration, 3D reconstruction with a moving camera, and photometric-based reconstruction. Peter Sturm has 30 scientific publications.

Frédéric Devernay (10%) holds a position of associated researcher at INRIA. Frédéric Devernay is a former student and PhD student of Ecole Polytechnique (1992 and 1996). He started his professional career with Istar Ltd. He joined INRIA in 2000. His research interests cover computer vision, computer-assisted robots for surgery, augmented reality systems, and motion capturing systems. Frédéric Devernay has 10 scientific publications.

Rémi Ronfard (10%) holds a position of associated researcher at INRIA. Remi Ronfard graduated from Ecole des Mines de Paris in 1986 and obtained his PhD in 1991, also from Ecole des Mines de Paris. He has worked as a software engineer for Dassault Systemes (CATIA) and Computervision (CADDS) and as a researcher at Institut National de l'Audiovisuel. He was invited twice at the IBM T.J Watson Research Center in Yorktown Height, New York as a post-doc in 1992 and a visiting scientist in 2000. His research interests have evolved over the years from deformable contours and geometric modelling, to video processing, analysis and indexing.

- Peter Sturm and Bill Triggs. A Factorization Based Algorithm for multi-Image Projective Structure and Motion. Proceedings of the 4th European Conference on Computer Vision, Cambridge, England, LNCS volume 1065 pages 709–720, April 1996.
- R. Horaud, G. Csurka, and D. Demirdjian. Stereo calibration from rigid motions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1446–1452, December 2000.

The Odyssée group

The research team. The Odyssee research team is supported jointly by INRIA, the Ecole Normale Supérieure in Paris and the Ecole Nationale des Ponts et Chaussées (CERMICS Laboratory). It is located in Sophia-Antipolis. The laboratory is exploring simultaneously computer and biological vision under the direction of Olivier Faugeras. The team includes 3 research directors, 4 senior research scientists and confirmed researchers, 10 PhD students, 1 post-doc, and 2 software and hardware engineers.

The associated PhD programme: The Sophia-Antipolis site encompasses one of the highest concentrations of National and Industrial Research Centers (e.g. I3S lab. of CNRS, INRIA, EURECOM) and leading Graduate Engineering Schools in addition to leading-edge telecommunications and technology-based companies. The Doctoral School *Sciences et Technologies de l'Information et de la Communication* with high scientific standard and a competitive admission procedure, is a fundamental actor of this network of excellence with this Telecom Valley.

This Doctoral School, with high scientific standard and a competitive admission procedure, is a fundamental actor of this network of excellence with this Telecom Valley. This Doctoral School proposes DEAs for graduate students (in France, PhD programs start with a DEA) as well as Doctoral Theses in the fields of Computer Science (referred to as Communication and Information Sciences in France), including Telecommunication. This Doctoral School is linked with the University of Nice-Sophia Antipolis and the local and national European Research Centers and Industrial Partners.

Olivier Faugeras (10%) is Research Director at INRIA in the Sophia Antipolis Research Unit where he leads the ODYSSEE Laboratory. O. Faugeras is member of the French Academy of Science and member of the French Academy of Technologies.He currently teaches a class on Computational and Biological Vision at ENS Cachan and Nice University. He contributed to start Noesis and Realviz companies. He is the author of two books and more than two hundred journal papers. He directed 30 PhD students. Many of them are internationally re-known researchers.

Rachid DERICHE (10%) graduated from Ecole Nationale Supérieure des Télécommunications, Paris, in 1979 and received the Ph.D degree in Mathematics from the University of Paris IX, Dauphine in 1982. He is currently a Research Director at INRIA Sophia-Antipolis the ODYSSEE Laboratory. His research interests are in Computer Vision and Biological Vision and include Partial Differential Equations applied to Image Processing and Computer Vision. More generally, he is very interested by the application of mathematics to Computer Vision and Image Processing. He has authored and co-authored over 120 scientific papers.

Maureen Clerc (10%) research scientist

Thierry Vieville (10%) research director

Theo Papadopoulo (10%) research scientist

Renaud Keriven (10%) research scientist

Pierre Kornprobst (10%) research scientist

- C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras. Constrained flows on matrix-valued functions : application to diffusion tensor regularization. In *Proceedings of ECCV'02*, June 2002.
- N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *The International Journal of Computer Vision*, 46(3):223, 2002.

The VISTA group

The research team. Vista involves 7 full-time researchers (5 Inria researchers and 2 CNRS researchers) and 1 assistant professor from University of Rennes 1, 1 Inria engineer, 10 Ph-D students, 3 post-docs. Vista research work is concerned with image sequence analysis and active vision. More precisely, we address two main topics : - analysis of physical phenomena (in a broad sense) to provide image-based or scene-based motion-related measurements, to achieve motion modelling and learning along with recognition of temporal events; - perception and control of automated or robot systems, to handle manipulation, tracking, navigation, surveillance or exploration tasks. In that context, we are interested in several types of spatio-temporal images, optical images (video, infra-red) as well as acoustics ones (sonar, echography). We mainly explore statistical approaches (Markovian models, Bayesian methods, robust estimators, statistical learning techniques, non linear filtering,...) to properly handle motion and deformation analysis and tracking issues. We also investigate active perception, when the sensor can be controlled. This involves the design of visual servoing techniques and of higher level exploration strategies. We are dealing with several application domains: video processing and indexing, meteorological imagery, medical imagery, experimental visualization in fluid mechanics, vision-based robotics and augmented reality, transportation.

The associated PhD programme: University of Rennes 1, MATISSE doctoral school, Research master STIR (Signal, Telecommunications, Image, Radar)

Patrick Bouthémy, Inria senior researcher, Head of Vista team, 10% involvement Motion modelling and learning, statistical image sequence processing, image sequence restoration and super-resolution, motion analysis, tracking, motion recognition, video indexing.

Etienne Mémin, Assistant professor, University of Rennes 1, 20% involvement Fluid motion analysis, optic flow, non linear tracking.

Frédéric Cao, Inria researcher, 20% involvement Motion detection, curve matching, EDP-based image processing

- R. Fablet, P. Bouthemy, Non-parametric scene activity analysis for statistical retrieval with partial query, Journal of Mathematical Imaging and Vision, 14(3):257-270, May 2001.
- T. Corpetti, E. Mémin, P. Pérez, Dense estimation of fluid flows, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(3):365-380, March 2002.

3.2.2 Partner 2: The International School for Advanced Studies.

The research team. The International School for Advanced Studies (SISSA), founded in 1978, is a centre for research and postgraduate studies leading to a PhD degree (equivalent to the Italian "dottore di ricerca"). It is the only institute of its kind in the Italian university system. The school is funded by the Ministry for Education, University and Research: about 15% of its budget comes from research contracts with Italian, European and International Institutions. Initially concentrated around the so-called "hard sciences", SISSA's Sectors have recently raised to comprise new highly promising and dynamic fields such as Neuroscience, Structural and Functional Genomics. Each Sector is scientifically and didactically autonomous, but exhibits flexible lines of research that often lead to fruitful synergies with other institutions in the Trieste Area (University, Elettra Synchrotron, ICTP, ICGEB). SISSA's permanent teaching staff relies on 52 full and associate professors and researchers, but this number increases regularly due to a large number of visiting scientists and post-doctorate fellows. Besides, both permanent and temporary staff includes many foreign scientists from all over the world.

The associated PhD programme: From an official point of view SISSA, in accordance with articles n. 14, 15 and 16 of the Statutes, published in the "Gazzetta Ufficiale" n. 62 dated 15.03.2001, in compliance with the decision of the governing bodies of the School, the International School for Advanced Studies in Trieste offers post-graduate training with the possibility of obtaining the "Doctor Philosophiae" (Ph.D.) degree in the following fields: Applied Mathematics, Astrophysics, Functional and Structural Genomics, Mathematical Analysis, Mathematical Physics, Statistical and Biological Physics, Neuroscience, Theory and Computational Physics of Condensed Matter.

The Ph.D. degree is equivalent to the Italian "Dottore di Ricerca" in accordance with article n. 18 of the Statutes, published in the "Gazzetta Ufficiale" n. 62 dated 15.03.2001. The Ph.D. programme will last for three years and will be based on teaching and research activities, with examinations to determine admission from one year to the next and the attainment of the Ph.D. degree. The course of study may be extended for a fourth year, subject to approval by the School authorities. In accordance with law n. 398/89 and Ministry Decree dated 11.08.1998 published in the "Gazzetta Ufficiale" n. 293 dated 16.12.1998, the School offers fellowships to students who have made a specific request in their application, and who meet the requirements concerning personal income as stipulated in the regulations. The fellowship is awarded for 3 years, extendable for a fourth year subject to availability of funds. Candidates may also hold fellowships from other sources, however these cannot be held in addition to those awarded by the School. The School may supplement external fellowships if they are of an amount less than those awarded by the School. Positions are advertised in the Italian "Gazzetta Ufficiale" and recently also in "Nature" Magazine and web site. All lectures, seminars, classes at SISSA are held in English so to make the whole environment fully international. One of the major goals of SISSA is to operate a fruitful integration among different lines of research and disciplines, and to favour the natural osmosis existing in its pluri-disciplinary environment. Special emphasis is given to young students' preparation: only highly motivated students enter the School, after a competitive exam. SISSA's courses offer specific features as an added value for the students: 1) basic and applied research techniques 2) small groups and close supervisor interaction 3) very frequent and often daily interactions with the tutor/supervisor 4) concentrated 3-4 years programs 5) top class and international external examiner for PhD exams 6) excellent track-record for job placement in top-quality institutions

The PhD curriculum is strictly linked to research activity. Initially there is an intense and concentrated period of courses aiming at bringing students coming from different academic backgrounds up to the same level of preparation. At S.I.S.S.A., all students are required to attend the introductory courses. In addition, we offer a number of short courses, divided into three areas: Cognitive Neuroscience, System and Computational Neuroscience, Cellular and Molecular Neuroscience, for a total of 80 hours. Students will choose among the available courses and are requested to attend lectures for at least 40 hours. After this " tutorial " stage training involves the full immersion of the student in a research group in which his/her specific work is guided by a member of staff (supervisor).

Vincent Torre (Full Professor): percentage of time 25 %. Vincent Torre has a strong background in physics and mathematics. He has contributed to the foundation of Computational Neuroscience and has given important contribution to Machine Vision, in particular to edge detection and motion analysis.

Anna Menini (Full Professor): percentage of time 15 %. Anna Menini has started her scientific carrier as a solid state scientist before becoming involved in the analysis of sensory processing. She has provided major contributions to the understanding of olfaction.

Alessandro Treves (Associate Professor): percentage of time 20 %. Alessandor Treves is a theoretical physicist primarily interested in the analysis of neuronal network and of their properties. He has significantly contributed to our understanding of the parallel processing of the brain.

1. Zoccolan D., Torre V., J Neurosci., 2002, Mar. 15, 22(6): 2283-98 Using optical flow to characterize sensory-motor interactions in a segment of the medicinal leech.

2. Pellegrino F.A., Vanzella W., Torre V., BMCV 2002, Tubingen, Germany (pag. 38-49) How the spatial filters of area V1 can be used for a nearly ideal edge detection.

3.2.3 Partner 3: Malmo University and Lund University.

The research team. The partner Malmo consists of two different teams in very close co-operation. One team is the Applied Mathematics Group (AMG) at the newly started Malmo University and the other team is the well-known Mathematical Imaging Group (MIG) at Lund University.

AMG consists of one professor, 4 readers (one in image analysis), 5 associate professors (2 in computer vision), 1 research assistant and 6 PhD-students (3 in computer vision, 2 in medical image analysis, 1 in mathematical modelling) and 1 software engineer. The laboratory infrastructure consists of digital video cameras, digital cameras, machine vision cameras, machine vision equipment (lighting, conveyor belt, robot) and computers. The topics of interests are: compute vision, especially structure and motion estimation, 3D-reconstruction, auto-calibration, affine reconstructions, reconstruction from specularities, medical image analysis, especially segmentation methods, splines, level-set methods, fast marching methods, applications, classification methods, artificial neural networks, support vector machines, kernel methods, foundation of image analysis, especially feature extraction, tracking, scale-space theory, PDE-methods.

MIG consists of two professors, 1 reader, 1 associate professor, 1 research assistant and 6 PhD-students (3 in computer vision, 2 in medical image analysis, 1 in mathematical modelling). The laboratory infrastructure consists of digital video cameras, digital cameras and an industrial robot available at the department of automatic control. The topics of interests are: compute vision, especially structure and motion estimation, 3D-reconstruction, auto-calibration, affine reconstructions, degenerate configurations and motions, medical image analysis, especially active shape methods, active appearance models, applications, cognitive vision, especially support vector machines, kernel methods, foundation of image analysis, especially feature extraction, tracking, scale-space theory, PDE-methods.

The associated PhD programme: PhD-programme in applied mathematics at Lund University, Centre for Mathematical Sciences.

Anders Heyden, Professor in Applied Mathematics, 20% involvement (M) Expertise: Computer Vision and Image Analysis, especially, structure and motion estimation, 3D-reconstruction, auto-calibration, robust methods, medical image analysis,

Gunnar Sparr, Professor in Mathematics, 10% involvement (L) Expertise: Computer Vision and Image Analysis, especially, structure and motion estimation, 3D-reconstruction, medical image analysis, computer tomography, handwritten character recognition.

Kalle Åström, Professor in Mathematics, 10% involvement (L) Expertise: Computer Vision and Image Analysis, especially, structure and motion estimation, 3D-reconstruction, auto-calibration, cognitive vision, medical image analysis.

Peter Linde, Associate Professor in Image Analysis, 5% involvement (M) Expertise: Image Analysis, especially applications in astronomy, non-linear point-spread functions, optimization methods.

Alexei Iantchenko, Associate Professor in Mathematics, 5% involvement (M) Expertise: Partial Differential Equations, especially foundations of image analysis, scale space theory, level-set methods.

Stefan Diehl, Associate Professor in Applied Mathematics, 5% involvement (M) Expertise: Partial Differential Equations, especially foundations of image analysis, scale space theory, level-set methods, viscosity solutions, non-linear conservation laws.

Amiran Ambroladze, Associate Professor in Mathematics, 5% involvment (L) Expertise: Wavelets, Partial differential equations, Cognitive Vision, especially Support Vector Machines and Kernel Methods.

Niels Christian Overgaard, Research Assistant, 10% involvement (M) Expertise: Partial Differential Equations, especially feature extraction and tracking, scale space theory, level-set methods.

Fredrik Kahl, Research Assistant, 10% involvement (L) Expertise: Computer Vision, especially, structure and motion estimation, 3D-reconstruction, auto-calibration, affine structure and motion, degenerate configurations/motions.

1. Heyden, A., Reduced Multilinear Constraints: Theory and Experiments, International Journal of Computer Vision, Vol. 30, No. 1, 1998, pp. 5-26.

2. Heyden, A., Tensorial Properties of Multilinear Constraints, Mathematical Methods in the Applied Sciences, Vol. 23, 2000, pp. 169-202.

3.2.4 Partner 4: IT University of Copenhagen.

The research team. The Image Group is located at the Department of Innovation, IT University of Copenhagen (IT-C), Denmark. The IT-C is founded on April 1st, 1999, and is still for the next many year expected to grow. Currently, the Image Group consists of 2 permanent staff, 1 post. doc. and 12 PhD students. One more tenure track is being filled, and one will be advertised for 20004. 2 post.doc. positions will be filled in 2003.

The group conducts research in the foundations of image analysis and its applications to medicine, human-computer interfaces, and film/TV production.

The group uses networked portable PCs, an image server with 2 Tbyte HD, a Linux cluster of 10 closely interlinked and 256 Ethernet linked PCs, 2SGI workstations including a professional film post-production machine with Discreet Logic Flint and real time playback from raided HDs of 2Kx2K movies.

The associated PhD programme: The IT University of Copenhagen has its own PhD program according to the regulations of the IT University of Copenhagen ("Vedtægter for IT-højskolen i København") approved by the Ministry of Science, Technology, and Innovation. This is managed according to "The PhD Curriculum at the IT University of Copenhagen". Further information may be obtained at www.it-c.dk/Internet/research/phd/.

The image group is founding member of the PhD school CISP: Copenhagen Image and Signal Processing Graduate School (www.imm.dtu.dk/cisp) in collaboration with the Technical University of Denmark and University of Copenhagen.

The IT University and CISP offer in collaboration approximately 10 PhD courses a year relevant to students in image and signal processing.

The PhD program at the IT University of Copenhagen is regulated according to "Ministerial Order on the PhD Course of Study and on the PhD Degree" implying 3 years of PhD studies including 30 ECTS of course work and work for the department including teaching at the amount of up to 840 hours.

Current items on the list of activities are: 1) Advanced Image Analysis, 10 ECTS, every spring semester 2) Foundation of Image Analysis, 7.5 ECTS every autumn semester 3) Medical Image Analysis, 10 ECTS, every autumn semester 4) Image Analysis for post-production, every spring semester and a number of ad hoc courses offered in topics reflecting the interests of the PhD students in CISP, being currently a population of 34 PhD students.

We foresee a number of additional PhD courses in collaboration with the partners under this program: 1) Nonlinear diffusion, 2) PDE-based computer vision 3) Scale-space for computer vision.

Mads Nielsen, 37 years, Full professor, 10% involvement. Expertise in image analysis and its application to medicine, human-computer interfaces, and film/TV production. Major research topic is in the statistical and geometrical foundation of image analysis striving toward a proper theory of handling geometric inference.He has published more than 50 papers in peer reviewed journals and conferences.

Ole F Olsen, 33 years, Assistant professor, 10% involvement. Expertise in the geometrical foundation of image analysis and its application to medicine and film/TH production. Major research topic is in singularity theory and its application to PDEs and image feature detection. He has published more than 10 papers in peer reviewed journals and conferences.

1. M. Lillholm, M. Nielsen, L. Griffin, "Feature-Based Image Analysis", to appear in International Journal of Computer Vision, 2003.
2. P. R. Andresen, M. Nielsen, "Non-rigid registration by geometry-constrained diffusion", Medical Image Analysis, vol 5(2), pp. 81–88, 2001.

3.2.5 Partner 5: The Technion.

The research team. We are working within the intelligent systems lab. in the CS department at the Technion, Israel. The vision group include 4 faculty members, one engineer, about 2-3 postdocs, and about 20 graduate students. The topics are relatively wide and include: geometry of images, differential geometry methods, 3D measurement and recognition, robotic guidance, statistical models for recognition and grouping tasks, and multi-robotics.

Students who choose to specialize in image related issues can choose from a variety of courses. Every year we teach at least 5 courses. This year for example we taught: introduction to signal and image processing (introductory), (advanced) image processing, computer vision, numerical geometry and Theoretical issues in object recognition.

We hold a weekly seminar: the "pixel club" where guests, faculty and graduate students describe their research.

The associated PhD programme: Our PhD project is research oriented. It is specifically stated that the prime goal of the doctoral studies is the development of original research work, which must advance the state of knowledge in the student's field and be acceptable for publication in professional journals.

Usually the PhD comes after completing a master thesis, and therefore, not much formal studying is required. Still, it is expected that the student will study at least one course every semester. There is an option for a "direct path" (PhD without completing a master first) which now becomes more popular. Choosing this option the student should take the full load of the master studies as well (about 12 courses).

After finding a supervisor, the student has to submit a detailed research proposal within 11 months from admittance and to pass a qualifying examination after submitting the proposal. After the student successfully passed the qualifying examination, he has to give talks on his progress once a year in a special seminar called the "Candidates' Day".

Prof. Alfred Bruckstein (10%) - image geometry, differential methods, distributed robotics

Prof. Ron Kimmel (10%) - differential methods for segmentation, path planning etc. recognition from 3D data

Prof. Michael Lindenbaum (10%) - Statistical analysis and algorithms for recognition and grouping.

Prof. Ehud Rivlin (10%) - function based recognition, vision for robot navigation

1. A. Amir and M. Lindenbaum. A generic grouping algorithm and its quantitative analysis. IEEE Trans. Pattern Analysis and Machine Intelligence, 20(2) pp. 168-185, 1998.

2. M. Elad and A.M. Bruckstein, "A Generalized Uncertainty Principle and Sparse Representation in Pairs of R^N Bases", IEEE Trans. On Information Theory, Vol. 48, no. 9, pp. 2558-2567 2002.

3.2.6 Partner 6: University of Oxford.

The research team. The research will be carried out within the Robotics Research Group (RRG) of the Department of Engineering Science, University of Oxford. The RRG is one of the largest and best known in its field in Europe, with five faculty and around fifty researchers in total. A substantial part of the group's effort has been invested in the recovery of 3D scene information from multiple images and rendering novel views. Emphasis has been placed on the use of projective geometry and geometrical invariants, as well the use of real-time visual processing. More recently the focus is changing to the use of learning for visual category recognition. Other interests include navigation for mobile robots and wearable computing.

The associated PhD programme: A student would be enrolled as a Probationary Research Student in the Department of Engineering Science. A PhD typically takes 3-4 years. During the first year the student attends a number of courses run by the department. There are also weekly seminars and reading groups. The student would meet with his/her supervisor weekly.

Dr. Ian Reid is a University Lecturer in Engineering Science and Fellow of Exeter College. His research concerns various aspects of Computer Vision. He obtained BSc from the University of Western Australia in 1987, and came to Oxford University on a Rhodes Scholarship in 1988 where he completed a D.Phil. in 1991. His research interests concern algorithms for visual control of active head/eye robotic platforms (for surveillance and navigation), visual geometry and camera self-calibration (applications of these to measurement, AR and VR, particularly for sporting events), and human motion capture. Together with Prof. David Murray, he runs the Active Vision Laboratory, but also has close ties with the Visual Geometry Group and the Visual Dynamics Group.

Professor Andrew Zisserman (10%) has researched in the Computer Vision area since 1984. He has written/edited eight books and has co-authored over a hundred papers. In particular he has extensive experience in object recognition (and received the IEEE Marr prize in 1993 for work in this area) and multiple view geometry.

Dr Andrew Fitzgibbon (10%) is a Royal Society Research Fellow. He has researched in Computer Vision for over 10 years, and also has extensive experience in multiple view geometry.

1. R. Fergus, P. Perona and A. Zisserman, 'Object Recognition by Unsupervised Scale-Invariant Learning', Computer Vision and Pattern Recognition, 2003.
2. R. I. Hartley and A.Zisserman. 'Multiple View Geometry in Computer Vision', Cambridge University Press, ISBN: 0521623049, 2000.

3.2.7 Partner 7: University of Ljubljana

The research team. Computer Vision Laboratory is a part of the Faculty of Computer and Information Science, University of Ljubljana, Slovenia. Currently, the research group consists of 3 professors, 3 experienced researchers, 3 PhD students, and 4 Master students.

The laboratory is involved in basic research in computer vision, with emphasis on visual learning and recognition, range image interpretation, panoramic imaging, and multimedia applications. Research in the area of visual learning and recognition focuses on subspace methods, which enable direct view-based building of visual representations and subsequent visual recognition of objects and scenes. The main research achievements in this area are development of robust approaches to both learning and recognition. Applications include recognition of objects, scenes, and activities, in visual cognitive tasks. Range image interpretation includes range image acquisition, segmentation of range images using the "recover-and-select" paradigm and modelling of shapes using different types of parametric models (i.e., superquadrics, surface patches). Panoramic imagery is mainly used for mobile robot navigation, as well as for designing efficient user interfaces for remote camera manipulation and for visual surveillance applications.

The laboratory infrastructure is based on the network of about 20 Linux/Windows/Macintosh computers. Two mobile robots (indoor IS Robotics and outdoor iRobot) are used for visual navigation experiments. The other special vision equipment consists of a structured light range scanner with a translational (x, y and z direction) and rotational (360o) computer controlled table of Arrick Robotics, pan/tilt units (Directed Perception) and several cameras, including web cams and panoramic sensors.

The associated PhD programme: Postgraduate programmes at the Faculty of Computer and Information Science, University of Ljubljana, consist of two years (four semesters) of course work. This can be followed by a Masters thesis leading to a M.Sc. degree and subsequently by a Doctoral thesis leading to a Ph.D. degree in Computer and Information Science. The Masters thesis can also be avoided and students can go straight to work on their Doctoral thesis. The faculty offers two postgraduate programmes: Computer and Information science, and Information Systems and Decision Making.

A student holding M.Sc. degree (or its equivalent) in computer and information science (or in related disciplines) can skip the course work and go directly on the PhD programme, working on the doctoral thesis without having to pass additional exams. He has to finish his thesis in four years after the beginning of the doctoral study at the Faculty of Computer and Information Science.

During the research, a PhD student will have a supervisor (one of the faculty professors), who will guide him through the research work and arrange him necessary facilities. Additional co-supervisors from other institutions are also possible. The PhD student will have a working place in the Laboratory for computer vision; he will have at disposal all the equipment, library, and infrastructure. He will be able to collaborate with the laboratory staff and to use their software and other facilities. He will also be involved in the projects related to his work. The supervisor will monitor his work, which should at the end result in a doctoral dissertation

ALES LEONARDIS (10%) is an Associate Professor of computer and information science at the Faculty of Computer and Information Science at the University of Ljubljana. He is author or co-author of over 130 papers published in journals and major conferences, and he co-authored the book Segmentation and Recovery of Superquadrics (Kluwer, 2000). He has been involved in national, international (bilateral), and EU-supported projects. His interests include robust and adaptive methods for computer vision, object and scene recognition, learning and 3-D object modelling.

FRANC SOLINA (10%) is a Professor of computer and information science. His research interests lie in the area of 3D volumetric modelling (superquadrics), range image segmentation, part-level object recognition, face detection, and use of computer vision in the fine arts.

JASNA MAVER (10%) is an Assistant Professor of computer and information science at the University of Ljubljana. From 1989-1991 she was a visiting researcher in the General Robotics and Active Sensory Perception Laboratory at the University of Pennsylvania. She is author or co-author of over 20 papers published in journals and conferences. Her research interests include active vision and methods for learning and objects recognition

ALES JAKLIC (10%) is an Assistant at the Faculty of Computer and Information Science. He received his PhD in 1997 from the same institution. His main research interests include range image segmentation, recovery of superquadrics, and construction of CAD models from range images.

DANIJEL SKOCAJ (10%) is a researcher at the Faculty of Computer and Information Science, University of Ljubljana. He received his PhD degree from the same institution in February 2003. He is currently involved in EU 5th-FP project Cognitive Vision Systems - CogVis. His main research interests include automatic modelling of objects from visual information with the emphasis on the appearance-based visual learning and recognition.

1. A. Leonardis, A. Jaklic, and F. Solina. Superquadrics for segmentation and modelling range data. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(11):1289-1295, November 1997.

2. A. Leonardis and H. Bischof. Robust recognition using eigenimages. Computer Vision and Image Understanding, 78(1):99-118, 2000.

3.2.8 Partner 8: Ecole Polytechnique Fédérale de Lausanne.

The research team. Our activities focus on modelling people and their motion from images and video sequences. Deriving human body shape and motion from image-data is an inherently difficult task. The body is very complex. The data is often incomplete, imprecise or, even, full of errors. We therefore develop techniques that use sophisticated human animation models to fit such noisy and incomplete data. We use facial and body animation models, not only to represent the data, but also to guide the fitting process, thereby substantially improving performance.

We are also interested in Augmented Reality and have developed robust real-time registration techniques that can handle large camera displacements, drastic aspect changes, and partial occlusions. We use them to create Augmented Environments in which real and virtual humans coexist and interact.

The current staff includes one professor, one post-doctoral fellow and eight doctoral students. They are funded in part by the Swiss Federal Institute of Technology, the National Swiss Research Foundation, and the Federal Office for Education and Science and in part by several European Union projects.

The associated PhD programme: EPFL offers an interdisciplinary program for graduate studies in computer, communication and information sciences. The study areas presently include systems, communications, signal processing, and all aspects of computer science. We currently offer only a doctoral degree program.

The study in the program has three aspects. First, a number of courses are offered to let the students develop strong theoretical background. Second, more advanced courses, projects and a variety of seminars introduce the research problems of current interest in many fields as well as equipping the student with the advanced techniques that may be useful in attacking them. The third and major part of the study is original research conducted under the direct supervision of a faculty member that results in a PhD thesis.

For the purposes of administration, research is considered to be carried out in four orientations given below. A student admitted to the program is registered in one of these orientations:

- Communication Systems: An inter-disciplinary program, at the crossroads of electrical engineering, computer science and mathematics, extending from physical layer to application aspects. Keywords: Stochastic Models, Communication/Information Theory, Signal Processing, Networking, Mobility, Security.

- Computer Science / Global Computing: A focused computer science program, oriented towards foundations of distributed and global computing. Keywords: Concurrency, Distributed Algorithms, Distributed Systems, Information Systems, Computer Networking, Programming Languages, Software Agents.

- Computer Science / Computing and Multimedia Systems: From multimedia information to virtual environments, an inter-disciplinary program focused on the new methods to interact with computers using multimodal interfaces. Keywords: Virtual Reality and Tele-Presence, Computer Vision, Neural Networks, Natural Languages, Object-Oriented Technologies, Distributed Systems, Software Agents.

- Signals and Systems: A program focused on signal analysis, system and control theory, with a special emphasis on mathematical description, properties and design aspects. Keywords: Linear Systems, State Estimation, Kalman Filtering, Digital Image Processing, Nonlinear Signal Modelling, Robust Control, Optimization of Dynamic Processes.

Professor Pascal Fua (involvement 10%) received a degree from Ecole Polytechnique, Paris, in 1984 and a Ph.D. in Computer Science from the University of Orsay in 1989. He joined EPFL (Swiss Federal Institute of Technology) in 1996 where he is now a Professor in the School of Computer and Communication Science. Before that, he worked at SRI International and at INRIA Sophia-Antipolis as a computer scientist. His research interests include human body modelling from images, optimization-based techniques for image analysis and synthesis, and the use of information theory in the area of model-based vision. He has (co)authored over 100 publications in refereed journals and conferences.

Dr. Vincent Lepetit (involvement 25%) joined EPFL in 2001 as a Research Assistant. Before that, he worked at INRIA Lorraine in the ISA research group as a Ph.D. Candidate. His research interests include camera motion recovery and occlusions handling for Augmented Reality applications and 3D tracking in monocular sequences.

1. R. Plänkers and P. Fua. Articulated Soft Objects for Multi-View Shape and Motion Capture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003.

2. L. Herda, R. Urtasun, A. Hanson, and P. Fua. Automatic Determination of Shoulder Joint Limits using Experimentally Determined Quaternion Field Boundaries. *International Journal of Robotics Research*, 2003. In Press.

3.2.9 Partner 9: Czech Technical University.

The research team. The Center for Machine Perception (CMP) will be the partner in VISIONET-TRAIN network. The CMP is part of the Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics. The CMP consists of 15 full time staff people (researchers plus one professor, one associate professor and three assistant professors). The group supervises 10 full time PhD students.

The CMP topics of interests are: computer vision, pattern recognition and mathematics for treating uncertainty. We intend to contribute to the network team in 3D vision related to modelling the 3D geometry of complex scenes, both static and moving. We have expertise in 3D projective and metric 3D reconstruction, solving the correspondence problem (both narrow and wide baseline), omni-directional vision, shape from shading, physics-based vision, e.g., shape from polarization, photometric stereo. The other experience of the CMP is in theoretical pattern recognition, both statistical and structural one as a theoretical basis of learning.

The associated PhD programme: PhD programme is named Artificial Intelligence and Biocybernetics. The head of the programme board is Prof. Vladimir Marik. It runs at the Czech Technical University, Faculty of Electrical Engineering.

The students of the programme are obliged to follow courses in the first three semesters of their PhD study. They can select their courses from quite large offer. Courses are given in English if there is one person in the audience who does not speak Czech. This courses will be available to the VISIONET-TRAIN PhD students.

Besides that the CMP runs a regular scientific seminar which takes place every week during the term. There are also one-day colloquia (a scientific event running the whole day with about 6 invited speakers, most of them from abroad). Check http://cmp.felk.cvut.cz/cmp/events/ for foreseen and past events.

The experience researchers visiting CMP will give a talk at a seminar or several talks within standard PhD or MSc level courses (in English).

Vaclav Hlavac, male, born 1956, professor, PhD (Prague) head of the Center for Machine Perception, interests: 3D vision, statistical and structural pattern recognition, industrial applications, 10% involvement

Jiri Matas, male, born 1964, PhD (U of Surrey, UK), full time researcher, interests: image-based recognition, retrieval, biometrics, face recognition, pattern recognition, industrial applications, 10% involvement

Mirko Navara, male, born 1959, associate professor, mathematician, PhD (Prague) interests: fuzzy and quantum logic, theoretical aspects of statistical pattern recognition, 5% involvement

Tomas Pajdla, male, born 1969, assistant professor, PhD (Prague), interests: 3D vision geometry, omni-directional vision, 10% involvement

Radim Sara, male, born 1963, PhD (Linz, Austria), full time researcher, interests: computational stereo, physics-based vision, shape from shading, 15% involvement

Martin Urban, male, PhD (Prague), born 1971, full time researcher, interests: 3D vision, geometry, reconstruction, motion analysis, structure from motion, prepares foundation of a spin-off company, 5% involvement

1. Michail I. Schlesinger and Vaclav Hlavac. Ten lectures on statistical and structural pattern recognition, volume 24 of Computational Imaging and Vision. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

2. Tomas Pajdla. Stereo with oblique cameras. International Journal of Computer Vision, 47(1-3):161-170, May 2002.

3.2.10 Partner 10: Universiteit Utrecht.

The research team. We are a small unit that forms (one of the twelve) programs of the department of Physics and Astronomy of the Universiteit Utrecht. The unit consists of a chair holder (full professor), an associate professor and a part time assistant professor. We hope to appoint an assistant professor later this year. A variable number of postdoctoral fellows (one to three) and graduate students (four to five on the average) participate in the research on a temporary basis. Our laboratory is equipped for psychophysical research in vision and haptics. We have special equipment for psychophysical work in the natural environment and for the study of light fields and the optical properties of materials from the natural environment. Topics of research are the perception of form and spatial relations in both haptics and vision and the perception of material properties.

The associated PhD programme: Helmholtz Research School, Universiteit Utrecht, Program "Physics of Man: Human Perception" of the department of Physics and Astronomy. The curriculum is subdivided into four categories: 1. The common and obligatory part. a. Attendance of the PhD student days, which will be held three times a year at different locations, and will consist of an introductory lecture by a senior scientist, a lab tour and an afternoon of PhD student presentations; b. Regular attendance of the Helmholtz colloquia, where guest researchers will present their work, A small number of PhD students can actively participate during the ÔPhD student lunchÕ with these researchers; c. Regular attendance of at least one of the seminars, held weekly or two-weekly by most of the participant research groups; 2. Courses. a. General methodological courses (e.g. English writing, presenting, statistics, lab animals etc.); b. Specialised courses (e.g. Special Helmholtz Courses or Regular Courses, listed elsewhere on the website). Both courses from within or from outside the Helmholtz Institute can be included; 3. Attendance of Winter- or Summer schools and Workshops. In concurrence with the supervisor, PhDs can attend Winter- or Summer schools as part of their curriculum; 4. Conference attendance. In concurrence with the supervisor, each PhD student is supposed to participate actively in one or more international conferences in the relevant fields of science

Jan J. Koenderink full professor, background in physics, experience in psychophysics, theoretical issues in computer vision and optics. Involvement 10%.

Astrid M. L. Kappers, associate professor, background in physics, experience in psychophysics. Involvement 10%.

Andrea J. van Doorn, assistant professor, background in physics, experience in psychophysics, Involvement 5%

1. J. J. Koenderink, A. J. van Doorn, A. M. L. Kappers, J. T. Todd, Ambiguity and the "mental eye" in pictorial relief. Perception, 30, 431-448, 2001.

2. S. C. Pont, J. J. Koenderink, Bidirectional reflectance distribution function of specular surfaces with hemispherical pits. Journal of the Optical Society of America A, 19, 2456-2466, 2002.

3.2.11 Partner 11: University of Mannheim.

The research team. Our research group is member of the Faculty of Mathematics and Computer Science and currently consists of the chair, 12 PhD-students and several graduate students. The laboratory infrastructure consists of devices for image (sequence) acquisition and Linux-PCs.

Basic research focuses on variational methods for segmentation and motion estimation, statistical learning and models for appearance and shape, and combinatorial optimization in connection with mid-level vision tasks (feature selection, grouping, matching). Applied research addresses problems in medical imaging, image measurements in fluid dynamics, and parallel variational processing on PC-clusters. Furthermore, there is a tight collaboration with mathematical research groups of our faculty, in particular in the fields of stochastic modelling (stochastic PDE's, MCMC sampling), and approximation theory (Fourier and wavelet analysis).

The associated PhD programme: The PhD-programme at the Faculty of Mathematics and Computer Science is organized by the three institutes of the faculty: Mathematics, Computer Science, and Computer Engineering. Besides courses on image processing, statistical pattern recognition, computer vision and computer graphics, a wide range of further courses from various mathematical topics up to those of computer engineering (e.g. special architectures for image processing) may be visited. Seminars and colloquia with respect to advanced and more specialized topics are regularly offered. The faculty organizes a two-week summer school each year including external guests.

Students enrolled in the PhD-programme have to participate in (at least) three advanced courses or seminars related to their research topic and in agreement with their supervisor. Colloquia with talks given by guests and visiting scientists are regularly offered. Students are expected to report regularly about progress and problems related to their own work (both by oral presentations and in written form). They are encouraged to publish results in international journals and conferences. There are weekly meetings with the supervisor.

Christoph Schnörr (10%), full professor of applied mathematics and computer science, has been working in the field of computer vision since 12 years and published about 70 papers in corresponding journals and conferences.
Research interests: image processing and computer vision, pattern recognition and related aspects of variational analysis, and convex and combinatorial programming.

Gabriele Steidl (5%), full professor of applied mathematics and computer science.
Research interests: approximation theory (Fourier and wavelet analysis), numerical linear algebra (Krylov space methods, Toeplitz and related systems).

Jürgen Potthoff (5%), full professor of mathematics.
Research interests: stochastic analysis, and filtering, MCMC-methods, image sequence processing.

Wolfgang K. Seiler (5%), apl. professor of mathematics.
Research interests: algebraic geometry and applications, invariant theory, algebraic methods in experimental design.

Ernst Binz (5%), full professor of mathematics.
Research interests: differential geometry, mathematical physics, (geometric) quantization, relativity, signal theory, (quantum) information theory.

1. D. Cremers, T. Kohlberger, and C. Schnörr. Shape statistics in kernel space for variational image segmentation. Pattern Recognition, 36(9):1929–1943, 2003.
2. J. Keuchel, C. Schnörr, C. Schellewald, and D. Cremers. Binary partitioning, perceptual grouping, and restoration with semidefinite programming. IEEE Trans. Patt. Anal. Mach. Intell., 25(11):1364–1379, 2003.

3.3 Intensity and quality of networking

We are aware of the fact that some partners are likely to attract early-stage researchers more than others. This is particularly true for partners belonging to Member states. On one hand partners from Great Britain, Germany, Sweden, France, Italy, The Netherlands, and Denmark do attract students from all over the world. This is also true for Switzerland. On another hand partners candidate countries are likely to be less attractive. European wise, the 11 partners are disseminated as follows:

- 7 partners belong to Member states;
- 2 partners belong to Associated states (Switzerland and Israel);
- 2 partners belong to Candidate states (Czech Republic and Slovenia).

In spite of discrepancies outlined above, the VISIONTRAIN proposal offers an **equal** number of person-months to all partners.

The recruitment strategy of the network, outlined in section 2.3 at page 33 plans to hire early-stage and experienced researchers **at the network level and not at the partner level**. This will enforce and encourage candidates to prepare their PhDs and their training with the network and not with one partner. We firmly believe that such opportunities will be attractive, especially for candidates from third countries.

3.4 Relevance of partnership composition

This partnership stems from the achievement of a large number of formal collaborations and exchanges. The 56 researchers involved in the network cover a large number of disciplines (see table 7, page 37) that are necessary for the successful implementation of the research objectives. The table below gives an indication of the intensity of existing collaborations, formal and informal.

Partner	1	2	3	4	5	6	7	8	9	10	11
1 (INRIA)		yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
2 (SISSA)	yes										
3 (Malmo)	yes			yes		yes	yes				
4 (IT-UC)	yes		yes			yes					
5 (Technion)	yes										
6 (Oxford)	yes		yes	yes				yes			
7 (UOL)	yes									yes	
8 (EPFL)	yes					yes					
9 (CTU)	yes									yes	yes
10 (Utrecht)	yes						yes		yes		
11 (Manheim)	yes		yes	yes			yes		yes		

Table 8: Existing collaborations between the network teams

4 Management and feasibility

Shaping and integrating training by research at the European level in the field of computational and cognitive vision systems is one of our main objectives. This can only durably succeed if there is an appropriate management structure which organises and monitors the cooperation within VISIONTRAIN. The network management structure, the role of the different bodies that compose it, and the management activities are described hereafter.

4.1 Network management structure

The following global tasks are identified for handling the network management:

- Decision making, i.e. how decisions are agreed between the partners;
- Operational management, i.e. how decisions are implemented within the network, and
- Advisory and assessment, i.e., recommendations for decisions and how implementation of decisions is controlled and should evolve.

Decision Making will mainly aim at handling contractual issues (approving changes) regarding Consortium agreement, changes in the Network specification, budget and fund distribution policy, quality assurance policy, Consortium structure, IPR principles, publications and confidentiality issues. VISIONTRAIN includes in the decision making bodies a Governing Board and an Executive Committee. Each body has very specific roles described in the next section.

Operational Management will aim at preparing and implementing decisions taken by the decision bodies. This implementation will encompass several aspects: implementation of the work, follow-up of work done, providing logistics for all coordination tasks, supporting and coordinating reporting, financial and administrative management, etc. It will also provide the platform from which the dissemination and IPR policy are programmed for the Network. VISIONTRAIN's operational management is carried out at the level of the **Management Board – MB**.

Advisory and assessment will be ensured by the **Scientific and Transfer of Knowledge board – STKB** which further decomposes into two councils:

- A *Scientific Council*, of experts in the field, which are both from inside and outside of the Network, and which, being stakeholders within the scientific field, have an interest of maintining the excellence represented by VISIONET.
- An *Intellectual Property Rights Council*, with one representative per partner which advises the Governing Board and the Executive Committee regarding publication, protection and dissemination of IPRs.

The *Scientific and Transfer of Knowledge Board* (STKB) will be chaired by Dr. Olivier Faugeras (INRIA). Dr. O. Faugeras is one of the most prominent researchers in the field of computer vision and since 2000 he is a full member of the prestigious French "Académie des Sciences".

For all bodies in the network structure, rules of quorum will be detailed in the Consortium agreement as well as voting rules in particular for decisions requiring unanimity.

4.2 Role of the different bodies

4.2.1 Decision Making bodies

Governing Board. The Governing Board is the Consortium's main decision-making and arbitration body. The Governing Board has one voting representative from each contractor. The Governing Board is chaired by a representative of the coordinator. The following list summarizes the role of the Governing Board that will be formalized in the Consortium agreement:

- decide upon the allocation of the network's budget
- decide about major changes in work packages, in particular for termination
- decide to suspend or to terminate the participation of a contractor
- take actions in case a contractor makes default
- anticipate and suggest solutions in case of a conflict between several partners and/or researchers
- agree upon a new entity to join the network
- decide the rules regarding the depository and the management of the funds received from the EC
- set-up, manage and co-ordinate budgets rules

Executive Committee. The Executive Committee's head is the project coordinator and has one member from each contractor who make proposals for decisions to be taken by the Governing Board, report to the Governing Board on how the decisions have been implemented, interact with both the Management Board and the Scientific and Transfer of Knowledge Board. The head of the Executive Committee chairs the committee. In emergency situations the head of the Executive Committee is authorised to take any decision required by the circumstances, which must then be validated by the Executive Committee The Executive Committee is in particular in charge of:

- making proposals to the Governing Board for the allocation and re-allocation of the budget based on the needs of the training and transfer of knowledge activities aiming at a strong integration, for the management of the funds received from the EC, for suspending or terminating part of the contracts or participation of contractors, for handling actions regarding a defaulting partner, including the coordinator;
- reviewing and/or amending the programme of activities contract and or the Consortium agreement
- selecting new contractors and proposing them to the Governing Board
- deciding about technical road-maps and selection of sub-contractors to get additional expertise

- supervising the work of the network management support team including quality assurance and preparing meetings with the EC, including related data and deliverables
- collecting and integrating recommendations from the network advisory councils
- approving press releases done by the network
- deciding on IPR management issues as detailed in IPR management subsection.

4.2.2 Operational Management Bodies

Operational Management is performed by the Management Board. However, for the smooth day-to-day management of the network, a **Management support team** will be appointed.

Management Support Team. The Management Support Team is composed of the network secretariat located at the coordinator site as well as one representant from each partner. The Management Support Team has the following responsabilities:

- manages the administrative, legal, financial and other aspects of the network;
- supplies project steering support, including follow-up of planning schedule, issue reminders for task initiation or completion, etc.
- supports the Coordinator in duties related to the interfacing with the European Commission
- supports the Coordinator in preparing Network deliverables;
- supports the Executive Committee in implementing the competitive selection procedure for new contractors.
- carries out secretariat tasks for the Governing Board and for the Executive Committee. The head of the Management Support Team is designated by the Executive Committee.

The Management Board. The MB is composed of the managers of the network Work Packages. The MB elects a leader and has the following responsibilities towards the Executive Committee:

- presenting progress reports on the state of activities and making proposals on programmes to be conducted and the arrangements for performance, the orientations of the activity and of the programme of activities;
- making proposals on the allocation of tasks, evaluating the financial needs and proposing the allocation among the contractors, the need to bring in new contractors, and the need for subcontractors for the establishment and the fulfilment of the Programme of Activity;
- drafting and validating Network deliverables to be submitted to the Commission for the relevant activity;

- identifying and informing on contractors presenting financial or technical risks within an activity;
- informing on any difficulty arising in connection with the running of the activity.

4.2.3 Advisory bodies

A Scientific Council and an Intellectual Property Rights Council (IPRC) are established to handle specific issue detailed here after. They make recommendations to the Executive Committee in relation to these issues.

Scientific Council. The Scientific Council is composed of 12 experts:

- 6 experts from inside the network;
- 6 experts from outside the network: 3 from Europe and 3 from other countries.

The Scientific Council is appointed by the Governing Board for two years and is renewed by one third at the beginning of the third and fourth year. The renewing procedure will maintain the balance between network and non-network experts. The Governing Board appoints the chairperson of the Scientific Council from among the members of this council. The head of the Executive Committee and the IPRC attend Scientific Council meetings in an advisory capacity. Prior to the beginning of his activity, each member of the Scientific Council shall enter into a non disclosure agreement. The Scientific Council is a scientific evaluation consultative body:

- it advises the Governing Board on Network orientations and the implementation by the Network of its mission for training by research and mobility;
- it is responsible of organizing the sessions for recruiting the early-stage and experienced researchers;
- it is responsible of periodically organizing the sessions for auditing and evaluating the research performed by the early-stage and experienced researchers.
- it validates and evaluates the training and transfer of knowledge activities as well as the results obtained;
- it may be consulted by the Governing Board on any scientific issues;
- it may make any proposal or transmit any information it deems useful to the Governing Board. For meeting organisation purposes, the Scientific Council shall be assisted by the head of the management board.

Intellectual Property Rights Council (IPRC) The Intellectual Property Rights Council is composed of IPR experts appointed by the Governing Board. The Governing Board appoints a chairperson of the IPR council from among the members of this council. The head of the Executive Committee attends IPRC meetings in an advisory capacity. Together with the Executive Committee the IPRC elaborates the rules governing IPR issue for the Network. Their tasks are described in the next sub-section.

4.2.4 Management of IPRs

INRIA as the coordinator of the Network has made everything possible so that partners express their willingness with regards to financial compensation due for access rights granted either on the pre-existing know how or on the knowledge. However intellectual property rights and their management will be detailed in the Consortium agreement. The management of IPRs is organized around the Executive Committee and the IPRC. The Executive Committee will:

- decide on licencing projects pursuant to the terms of the Consortium agreement;
- decide on terms and conditions of access rights to pre-existing know-how not listed prior to the signature of the EC Contract;
- decide on terms and conditions of access to IPRs by subsidiaries and affiliates not listed prior to the signature of the EC Contract;
- give instructions to the Management Support Team concerning the management of the Network knowledge portfolio upon consultation of the IPRC (application for patent, extension, withdrawal, etc.);
- in collaboration with the IPRC, ensure, the review of Network knowledge and take measures in connection with their industrial protection, defence and valorisation;
- decide on the acquisition of rights from third parties.

The Intellectual Property Rights Council (IPRC) will:

- propose to the Executive Committee the updating of the pre-existing know-how list,
- establish and review the Plan for use and dissemination of the Network to be submitted to the Governing Board for approval,
- identify knowledge that could be the subject matter of protection, use or dissemination by decision of the Executive Committee, based on contemplated publications and activity reports issued by the Management Board,
- assist the Executive Committee in the protection of Knowledge,
- submit proposals to the Executive Committee for the allocation of co-ownership shares over knowledge obtained by several contractors for decision by the Executive Committee

For a new contractor joining the Consortium, access to the pre-existing know-how of the other contractors for the purposes of research, use or dissemination is granted upon written request pursuant to the terms set forth in the Consortium agreement. Nevertheless, any contractor pursuant of the provisions of EC Regulation article 25.3 has the right to exclude some of its pre-existing know-how from the new contractor access rights. A new contractor has access to the knowledge produced in the scope of the Network prior to its arrival for the purposes of research, use or dissemination at market conditions. The specific consequences of the withdrawal of a contractor on IPRs and their access rights will also be detailed in the Consortium agreement.

4.3 Management activities

4.3.1 Setup and maintain a network infrastructure

This activity will be carried out by the **Management support team**. The coordinating partner will establish a **secretariat** (office space, office equipment, and human resources) to assist in the smooth running of the network. This will be the home of the Executive Committee of the Network. The secretariat will provide administrative assistance to all of the training and transfer of knowledge activities, though some may require additional investment at other sites as for example in the provision of additional web-site support.

Functionalities that will be considered for management support and for successfully achieving the objectives of the network include:

- Information sharing
  - Document authoring and versioning tools,
  - Code authoring and versioning tools,
  - Support for shared bookmarking, recommendation and webpage monitoring,
- Communication support
  - Video-conferencing,
  - Support for remote meetings,

Given the importance of the educational aspects for the network, these tools will be configured to take into account both peer to peer and teacher/student interactions.

The infrastructure will further include the maintenance of the constellation of web pages associated with the Network. These will be organized around a central VISIONTRAIN web-site maintained at INRIA and giving an entry point to information about the project including:

- links to the partners,
- web pages of thematic programmes and individual events organized by the project,
- the archive of technical reports arising from the research of the project,
- pages giving pointers to software and benchmark data sets,
- pages describing past and current challenges,
- pages giving pointers to demonstration systems,
- brokerage of expertise web-page,
- web pages concerned with any standardisations of format and associated formulations of 'best practice',
- links to web pages of associated projects.
- Web page giving a central location for advertising job openings in the areas covered by VISIONTRAIN.

4.3.2 Communication with the EC and coordination of reporting

This activity will consist of:

- Reporting to the EC services and contacting them for administrative purposes;
- Coordinating the yearly activity reports;
- Coordinating the final report;

4.3.3 Organisation of external audits

The project coordinator (Radu Horaud), assisted by the network secretariat will be responsible for organizing, commissioning, and monitoring the audit regime. This concerns all aspects of the network:

- Technical audits, including the annual project reviews,
- Financial audits, and
- All other audits that the EC wishes to organize.

Four technical audits will be organized after 12 months, 24 months, 36 months, as well as a final audit. The secretarial will organize these meetings, communicate all the reports needed for the audit both to the EC and to the external experts.

4.3.4 Organisation of internal evaluation

The structure of almost all the activities within the *training and transfer of knowledge programme* calls for an internal evaluation organisation and management. Therefore, this activity plays a crucial role in the network.

The Managing Board and the Scientific Council will be responsible of this activity. They are in charge of evaluating the individual applications for the recruitment of early-stage and experienced researchers, and of evaluating and allocating budget to various training activities (thematic schools, patent application, proof-of-concept, etc.).

They are also in charge of organizing internal audits in order to measure progress of the PhD theses. In particular:

- Each PhD student will be audited at month 6, 12, 24, and 36 of her/his PhD work. The PhD dissertation will be submitted to the *Scientific Council* which will recommend appropriate PhD examiners from outside the network.
- Each granted experienced researcher will be audited at the end of the granted period.
- Each group of network researchers receiving a grant within the **Proof of concept programme** will be audited at the end of the granted period. This audit will typically consist in a proof-of-concept demonstration.

4.3.5 Financial management

The project coordinator, the *Governing Board*, and the *Executive Committee* will be in charge of the project financial management, receive all payments from the EC, keep advance payments on a well identified account, and transfer payments to the partners.

This activity also includes:

- Preparation of a provisional budget for each one of the network activities, assignment of specific budget to the effective activities, and estimate the costs incurred by each partner by a specific activity;
- Reporting costs to the European commission: consolidated cost reports of participants, summary of cost statements, cost certificates, justification of the detailed expenses (labour, travel, subsistence, consumables, etc.), summary financial report by the co-ordinator;
- Obtaining the audit certificates for each partner.

4.3.6 Technical management

The *Management Board* will be responsible of technical management.

Technical management will be in charge of the monitoring of all the network scientific and technical, training, transfer of knowledge and dissemination activities. It will have the important task of promoting the integration of all these activities at all levels.

The monitoring, coordinating, and controlling the scientific and technical progress of the project will be achieved by means of bi-annual meetings and intermediate teleconferences.

In particular the technical management will have the following responsabilities:

- Produce the yearly technical report of the network. This report will be consolidated from the activity contributions. Corresponding deliverables will be referenced and annexed to these reports. The activity report will be matched against the training and transfer of knowledge programme as well as against the current scientific and technological state of the art in order to help measure the advancement of the work, and
- Suggesting changes and updating the training and transfer of knowledge programme. Detailed aspects of this activity are prospective since they depend on the evolution of the scientific and technological contexts.

4.3.7 Legal and knowledge management

The project coordinator, the *Governing board*, and the *Intellectual Property Rights Council* are responsible of this activity.

The activity concerns all legal aspects of the network:

- Negociations and monitoring of the network contract and of the consortium agreement;
- Relations, communications, and negociations with other legal bodies, such as the companies collaborating with the network, the European Patent Office, spin-off companies, etc.

- Intellectual property management: The network will produce intellectual properties whose rights must be properly stated and defended. This task will define the intellectual property regime that will be part of the consortium agreement. It will have to monitor it especially when other legal bodies wish to exploit the intellectual property provided by the network.

4.3.8 Human resources management

The Management board, i.e., the work-package leaders will be responsible for smoothly carrying out the work during the duration of the work-packages. They will have an important role to play concerning the management of human resources. Whenever needed, they will organize management meetings at the work-package level and they will try to anticipate any problems related to the achievement of the work by the concerned researchers, such as: end of contract of an early-stage or of an experienced researcher, reallocation of human resources by a partner, changes in personnel, leave of absence of a researcher, temporarily or permanent replacement of a researcher, unexpected delays encountered, conflicts between researchers, conflicts between a student and her/his advisor, etc. If the problem cannot be solved on a cordial basis and if the conflit still remains, the work-package leader will refer to the Executive Committee which will ask the Governing Board for a decision.

Since the network plans to recruit an equal number of women and men, the management board is responsible to anticipate long-term leaves of absence due to maternity. Modern communications tools will be offered to women researchers concerned by maternity leaves such that they will be able to follow-up on the research carried out by their colleagues while they are at home.

4.4 Management know-how and experience of network coordinator

Ever since the beginning of European programs, INRIA has been one of the most active actors in the field of information technologies, by coordinating and participating to a very large number of Esprit and IST projects as well as Marie Curie actions and networks.

Within FP5 INRIA has been one of the major EC contractors. This activity is supported by competent staff within the Institute's Finance, Administrative, Law, and Human Resources offices. The managerial support also includes advice and guidance on preparation of project proposals, negociations of contracts and grants, as well as patenting and IPR (intellectual property rights) issues.

In 2002 INRIA hired specialized personnel (management, intellectual property rights, etc.) for the successful implementation of FP6. More precisely, INRIA researchers appointed to coordinate European proposals and projects receive direct financial and human-resources support.

Dr. Radu Horaud will be the network coordinator. He is born in Rumania and was educated both in Rumania and in France. He spent two years in North America. He has a deep knowledge of society, culture, and education in European countries and in North America. He is married and he is the father of three children.

Radu Horaud is both a scientific and a mountain leader. As a mountaineer, he guided and climbed several hundreds of peaks in Europe (Matterhorn, Monte Rosa, ...) and in North America (Mount Whitney, Mount Rainier, ...) and organized several trips to Nepal, India, and Peru.

R. Horaud was in charge of international relationships at INRIA Rhône-Alpes from 1999 to 2002 and currently he is the INRIA Rhone-Alpes correspondent for the successful implementation of FP6.

Radu Horaud is a well known and respected scientist in the fields of computer vision and robotics. After completing a PhD at the Institut National Polytechnique de Grenoble, he started his professional career as an International Fellow at SRI International, Menlo Park, CA. where he had a two year appointment (1982-1984) with both the Robotics Department and the Artificial Intelligence Center. In 1984 he joined the Centre National de la Recherche Scientifique as a *chargé de recherche* and in 1992 he was appointed *directeur de recherche*. In 1998 he joined INRIA and in 1999 he became the head of the MOVI research group at INRIA Rhone-Alpes.

R. Horaud participated to the organization of a large number of international conferences and workshops. In particular he acted as programme co-chair for the Eighth IEEE International Conference on Computer Vision (ICCV), Vancouver, Canada, in 2001. ICCV is the most prestigious and largest conference in the field of computer vision. He also acted as local arrangements chair for the IEEE International Conference on Robotics and Systems (IROS) Grenoble, France, in 1997.

Radu Horaud directed 20 PhD students. Among these students many continue an European or international career. One of his former PhD students, Thomas Skordas (from Greece), is now in charge of an FP6/FET/Proactive program. Another of his former students, Andreas Ruf (from Germany), is a staff member of the European Patent Office in Munich.

Radu Horaud participated to a large number of European projects under Esprit-I, Esprit-II, Esprit-III, Esprit-IV, and IST programs. He was the coordinator of the VIGOR project (a Reactive LTR project) and of several TMR actions.

Radu Horaud has already collaborative experience with the following VISIONTRAIN partners: parter 3, partner 5, partner 6, and partner 8. He personally knows 30 among the 56 researchers involved in the network.

4.5 Management know-how and experience of network teams

Partner 2: The International School for Advanced Studies. SISSA has a its major mission the training of PhD students in Physics, Mathematics and Neuroscience. In 24 years of activity, S.I.S.S.A. has given 495 PhD: 360 to Italian citizens, 135 to foreigners from a large number of countries : Albania 1, Algeria 3, Argentina 9, Austria 1, Benin 1, Brazil 3, Bulgaria 2, Czech Republic 2, Chile 2, China 33, Croatia 7, Cuba 1, Egypt 1, France 3, Jordan 2, Great Britain 4, India 8, Iraq 1, Lithuania 2, Morocco 1, Mexico 4, New Zealand 1, Nigeria 1, Peru 2, Poland 11, Portugal 1, Rumania 7, Russia 3, Syria 1, Slovakia 2, Slovenia 1, Spain 1, United States 2, South Africa 1, Switzerland 2, Thailand 1, Turkey 3, Venezuela 3, Yugoslavia 1. Therefore SISSA knows how train PhD students and is very familiar in dealing with students from all over the world.

Partner 3: Malmo University and Lund University. Malmo/Lund has experience in participating in several successful EU/IST projects since 1992; VIVA 1992-1996, CUMULI 1996-1999, VISIRE, 2000-2003, INVIEW, 2002-2004, LAWA, 2002-2005. Malmo/Lund has also experience in arranging international conferences and workshops, including ECCV 2002, jointly organized with Copenhagen University and the IT-University of Copenhagen. Malmo/Lund has experience in giving PhD-courses aimed at an international audience, especially in computer vision/multiple view geometry. Malmo/Lund will assist in organizational work by organizing at least 5 PhD courses within the area of computer vision and image analysis. Anders Heyden is currently chair of the board for natural sciences and technology at Malmo University.

Partner 4: IT University of Copenhagen. The ITU is a new university, but already know it has considerable managerial and organizational experience regarding PhD studies and research projects. The Image Group is founding member of the national PhD-school Copenhagen Image and Signal Processing graduate school, where Mads Nielsen has a position in the board. CISP work for coordinating training activities in terms of courses and summer-schools, and so as to guaranty student optimal supervision. The Image Group is coordinating the FET open IST project DSSCV including 5 PhD students. The Image Group has organised 8 PhD courses of which 6 had international lecturers. The Image Group has co-organized the European Conference on Computer Vision 2002 in Copenhagen with more than 600 attendees. Mads Nielsen is head of the board of PhD studies at ITU and has as such a central position in management and organisation of the PhD studies at ITU. The Image group will assist in organizational work by organizing at least 5 PhD courses in the area of computer vision.

Partner 5: The Technion. In operation since 1924, the Technion - Israel Institute of Technology is the oldest university in Israel. Since its founding, the institute has educated three generations of men and women who have played a key role in laying the country's infrastructure and establishing its crucial defence and high-tech industries. The university offers degrees in science and engineering, and related fields such as architecture, medicine, industrial management and education in an intellectually invigorating environment. Great emphasis is also placed on its humanities and social science programs, the incorporation of which take on ever-increasing importance in today's multi-faceted workplace. But Technion's goals go beyond providing a well-rounded technical education. At the institute, scientific instruction is interwoven with professional ethics, producing leaders sensitive to social and environmental issues. The Technion occupies about 1,325,000 square meters and includes 100 buildings. There are about 40 research centers, 11 research institutes and 10 Centers of Excellence. At present there are about 9,200 undergraduate students, 3,470 Graduate Students, 2,850 M.Sc. students and 700 Ph.D., M.D. and D.Sc. students - a total of 12,800 students. Since 1929 40,000 students have graduated. There are over 700 faculty members and 58 spin-off companies. Since Israel has signed the Association Agreement with the European Union and became eligible to participate in the Framework Programs in 1996, the Technion successfully joined over 80 consortia and has participated in many of the EU specific programs, such as BRITE-EURAM, JOULE, LIFE, ESPRIT, FAIR, INCO, INNOVATION, COST, and INTAS.

Partner 6: University of Oxford. The Robotics Research Group has been a partner in a number of previous EC projects including Esprit BRA's (VIVA,IMPACT,IMPROOFS) as well as ACTS Project AC074 VANGUARD. All of these projects completed successfully. The group is currently a partner in the FET projects VIBES and CogViSys. The group has also hosted several Marie Curie postdoctoral students. Over the past 15 years about 100 hundred students have successfully completed their D.Phils, so there is considerable management and research supervision experience.

Partner 7: University of Ljubljana Management capabilities originate in past experiences related to international activities of the members of our lab: International coordination of CEEPUS network. Experiences in collaboration in bilateral (SLO-A, SLO-CZ, SLO-FRA, SLO-HR, SLO-US), and multilateral projects (Copernicus, Socrates, FP5-IST). XSManaging the exchange of researchers and students.

Partner 8: Ecole Polytechnique Fédérale de Lausanne. We have been and are involved as core partners in several IST projects and research networks. In addition to EU projects, our school's administration has extensive experience with all issues related to international collaborations. It has an Industrial Liaison Program that runs programs, aimed at fostering the exchanges between academy and economy on an EPFL-wide basis.

Partner 9: Czech Technical University. The Center for Machine Perception (CMP) of the Czech Technical University in Prague, Faculty of Electrical Engineering, Department of Cybernetics is the unit which will be involved in Visionet. This team and its head Prof. Vaclav Hlavac have experience in managing many research and industrial project, including the IST ones of European Union (e.g., ISAAC IST-2001-33266, ActIPret IST-2001-32184, BeNoGo IST-2001-39184). The EU project ICA1-CT-2000-70002: MIRACLE (Machine Intelligence Research and Application Centre for Learning Excellence) within the context of the "Centres of Excellence" running from 2000 till 2004, we have been hosting several early stage and experienced researchers in CMP for several month long stays. This means that we have demonstrated necessary managerial skills for attracting visitors and organizing their stays in Prague. Needed administrative support on the CMP and Czech Technical University level is available too. The CMP has experience in transferring academic knowledge into industry (e.g. to Rockwell Automation, Boeing, Robert Bosch, Samsung).

Partner 10: Universiteit Utrecht. We have participated in European programs for many years and have greatly benefitted from that. We have had (and still have) numerous interactions within (and outside) Europe with groups of various disciplines (physics, mathematics, experimental psychology, computer vision, computer graphics, neurophysiology, philosophy of mind). We regularly have postdocoral fellows from other countries. The university has started international masterprogrammes (in the English language), thus we expect an increase of foreign students.

Partner 11: University of Mannheim. The University of Mannheim is currently coordinating 1 EU-network, and is involved in 3 other EU-networks as core partner. Besides

EU-networks, the administration has experience with all issues related to a wide range of international collaborations. The list of partnerships for the exchange of scientists and students currently involves 34 countries worldwide. Provision is made for spending some time abroad in most course of studies. The campus includes two international centers for accommodating visiting scientists. In summary, the University of Mannheim is well prepared to serve as a partner in VISIONTRAIN.

5 Relevance to the objectives of the activity

Major universities, institutes, organizations, and companies established research groups actively working in the field of computer vision. Computer vision research in North America and especially in the USA has been sponsored since the 1970's through various NSF, DARPA, and ONR programs and grants. These funding efforts helped the establishment of a North American scientific community led by some of the major universities: Stanford, UC Berkeley, U Maryland, MIT, CMU, USC, U Illinois, Cornell, etc. In particular, DARPA funded the Image Understanding (IU) program allowing hundreds of academia to carry out computer vision work. Notably enough, a large number of American, Japanese, AND European companies established their research centers in the USA and an important "computer vision job market" is thus born across the Atlantic.

For example, in 1977 Siemens established SCR (Siemens Corporate Research) as a US research facility in Princeton, New-Jersey. SCR serves as an advanced technology resource in selected areas of R&D for both US and worldwide Siemens operating companies. At SCR in Princeton, scientists and engineers from many countries with diverse cultural backgrounds compose a unique research team with expertise in six core technologies: software engineering, imaging and visualization, **intelligent vision and reasoning**, **real-time vision and modelling**, **multimedia/video technology**, and multimedia documentation.

It is worthwhile to notice that this US-established Siemens research facility generates profit for the European-located Siemens companies and subsidiaries.

The objective of VISIONTRAIN is to launch, establish, and maintain an European center of competence in the field of computational and cognitive vision.

VISIONTRAIN will address a number of challenging and unsolved research topics through cross-fertilization between brain, cognitive, information, and mathematical sciences. Multidisciplinary research is more difficult to achieve than single-topic research but, if successful, will have a great impact on scientific findings because it brings fresh blood from one topic to another.

The overall result will benefit all these disciplines. Thanks to the sophisticated methodologies that have emerged in information and mathematical disciplines, better measurements of the activity of the human brain will be available for analysis in neuroscience and psychophysics. Theories of various cognitive processes developed in computer science can be used to chart possibly related brain operations. Experimental findings in neuroscience can suggest new theories, models and architectures in artificial cognitive system, an interaction very much alike the one between theoretical and experimental physics.

We foresee that this synergy will foster the emergence of a new scientific domain and considerably advance the state of the art at the theoretical and practical levels, a fine example of the tight coupling between science and technology.

Although Europe has excellent scientists both in computer and in biological vision research, none of the existing European universities or institutes is able to hire a critical mass of researchers. Moreover, the efforts of these researchers should be coordinated both in terms of **production of knowledge** and of **training and education**. An European job market exists in the field of vision and the ambition of VISIONTRAIN is to improve the competitiveness of European researchers as well as to attract researchers from non-European countries.

It is only around 1985 that vision research started to be funded at the European level mainly through the ESPRIT/IST and HCM/Marie Curie programmes. The European computer vision community started a number of European projects which rapidly pushed Europe as a main actor in the field.

One important vector for disseminating European research in this field is the organization of conferences and workshops. In 1990 the First European Conference on Computer Vision (ECCV'90) took place in Antibes France and gathered 200 participants. The Seventh ECCV (ECCV'02) which took place in Copenhagen in 2002 gathered 660 participants from almost all countries of Europe, as well as from USA, Canada, Japan, Australia, Korea, India, etc. It is worthwhile to notice that the First, Seventh, as well as the forthcoming Eighth ECCV have been chaired by researchers belonging to this research training network and organized by the University/Institute with which they are affiliated – INRIA (partner 1), IT-U Copenhagen (partner 4), and Czech Technical University (partner 9).

The instruments planned by VISIONTRAIN in order to (i) increase the number of researchers, (ii) maximize their intra European mobility, (iii) reduce the fragmentation of resources, and (iv) disseminate the results worldwide are as follows:

1. The VISIONTRAIN network includes 11 partners from 11 countries. The research teams involved in the network have interdisciplinary expertise (see table 7, page 37).

2. The PhD students recruited by the network as early-stage researchers will be offered the opportunity to "rotate" between at least 3 teams, therefore strong supervision of early-stage researchers within the network will be an everyday reality;

3. The experienced researchers from one network team (or from outside the network) will have the opportunity to visit at least two other teams from two different countries;

4. The total/global network expertise available will be much larger than the sum of the partners' expertise;

5. During their training period the researchers will be given the possibility to attend conferences and workshops as well as the VISIONTRAIN thematic schools and industrial meetings. They will also be given the possibility to apply for European patents and to visit one or several companies in order to establish formal contacts.

6. During their training the researchers will be exposed to different schools of thought, different social and cultural environments – a unique European feature;

7. A large number of the research teams involved in the network are international. Therefore, the experience of participating to such multi-cultural teams will be beneficial and will encourage mobility of young researchers, and

8. Cross-fertilization between information, mathematical, and neuro sciences is one keystone of VISIONTRAIN in terms of training and transfer of knowledge.

6 Added value to the community

European universities must face the internationalisation of education and research. It is notoriously known that European universities attract fewer students and fewer researchers from other countries than their American counterparts. A recent survey revealed that during the year 2000 Europe attracted some 450 000 students from other countries while North America attracted 540 000 students mostly from Asia. Moreover, the USA are more successful in keeping more persons with doctorate qualifications. Among Europeans who obtained their PhD in an American university, 50% stayed there for several years and many of them remained permanently.

Beyond the scientific training, within VISIONTRAIN we are going to implement a number of tools and instruments such that the trainees become aware of the true opportunities offered by the Member, associate, and candidate states in particular and by Europe in general. We firmly believe that European training networks like this one can better compete with North American universities and are likely to attract more researchers than a single group from a single university.

The European dimension of this Marie-Curie research training network is consistent with the application of article 169, as approved by the European Council in Lisbon in March 2000. In order to increase the role of European universities and place them at the heart of the factory of knowledge, we are going to implement a number of actions along the lines stated by the *European Research Area – ERA*. In particular the following actions will be taken.

Diversified and improved training conditions. The early-stage researchers will learn at an early stage of their career to co-operate with people from different countries and different academic backgrounds. The social skills thus acquired will lead to improvement of the social cohesion across the European Union. **Each ESR will visit at least 3 partners and each ER will visit at least 2 partners for long and medium periods of time – 3 to 30 months.**

The researchers from 11 different European countries will complement each other in this interdisciplinary scientific field in order to train their students and researchers. They will have a chance to develop lasting collaborations outside their country and their home university/institute.

Increase the role and involvement of the member states. Recently it has been recognized that a substantial increase in member state involvement must be achieved in order to retain young researchers in Europe. In particular, legal and administrative obstacles of various kinds and nature must be removed in order to allow true involvement of researchers undertaking an early stage training.

The employment contracts of the trainees, in accordance with member-state and European laws and regulations, will offer **equal opportunities** to both nationals and non-nationals, **women and men**, weather they belong to a Member state, associate state, candidate state or a third country (health insurance, maternity leave of absence, unemployment stipends, family aids, retirement plan, etc.) The early stage researchers hired and employed within this network will also be given access to cultural and linguistic training through intensive courses specially designed for graduate students. Indeed, the acquisition of linguistic, social, and cultural skills will lead to improvement of cohesion accross the European Union.

All along the completion of their PhD the trainees will be assisted to facilitate their understanding of the national academic system, how to apply for a faculty position in the host state, how to apply for government funding for starting a company, how to write an European proposal, etc.

Facilitate access to early-stage and experienced researchers from associated states and candidate countries. The partnership of this network will facilitate exchange of students between the member states, on one side, and associate and candidate states, on the other side. Namely Israel, Switzerland, Czech Republic, and Slovenia are represented in VISIONTRAIN.

It is important to outline the special case of Israel. The Israeli partner involved in this project (The Technion) counts among the best universities in the world. Currently, most of the students graduating from the Technion prefer to continue their studies over the Atlantic. It is crucial to establish strong connections between Member states and the Israeli universities. Therefore the VISIONTRAIN proposal will fully include this partner in the transfer of knowledge program.

Facilitate access to early-stage researchers from third countries. It is planned that 30% of the early-stage researchers will be recruited from third countries with special emphasis on Asia and South America. In recent years all the member-state partners of VISIONTRAIN established internship programs with major universities and institutes from third countries in order attract students. In particular there are formal collaborations with the Indian Institutes of Technology (IIT) from Delhi, Kanpur, and Kharagpur, with the Chinese Academy of Sciences, as well as with major universities from Brasil, Mexico, Chile, Uruguay, and Argentina.

The VISIONTRAIN network will benefit from formal contacts with US universities in order to recruit students there. Past experience reveals that US students interested by completing a PhD with an European university are very often from either Member or candidate states. Therefore this is one instrument that we intend to implement in order to attract these students back to Europe.

Integration of non-technical issues as part of the training programme. INRIA as a network coordinator together with its partners already recognized the importance of non-technical issues as part of a PhD training program. In particular the following issues will be addressed: Ethical issues are very relevant since the film, television, video, and video-game industries are among the potential users of the technologies developed under this research programme; Legal issues related to distribution and dissemination of software, Commercialisation of software, Copyright, European and national patenting, etc. **Intellectual property rights will be addressed explicitly by the IPR council.**

Strengthen the role of women. One of the most important goals of this research training network is to attract women from all countries.

Formal contacts will be established with "women and science" international organisations, such as *The Association of Women in Science* and *The Third World Organization for Women in Science.*

The VISIONTRAIN recruitment programme will give equal opportunity to both women and men. The vacant positions will be publicized in special newsletters and websites.

We will also offer to women early-stage and experienced researchers the opportunity to actively participate to forums, seminars, and workshops dedicated to the promotion of the role of women in the production and dissemination of scientific knowledge.

7 Indicative financial information

VISIONTRAIN plans the recruitment of both early-stage and experienced researchers. The estimated costs for each one of these categories are:

38000 euros per year for an early-researcher and

58000 euros per year for an experienced researcher.

Table 6 at page 33 indicates the total number of planned person-months for each category: 396 person-months of early-stage researchers and 198 person-months of experienced researchers. The table below indicates the total amount needed to cover these personnel expenses as well as the total requested budget.

Researchers	person-year (euros)	Number person-years	Total (euros)
Early-stage	38 000	33.0	1 254 000
Experienced	58 000	16.5	958 000
Total for ESR & ER (65%)			2 212 000
Total for hosts (35%)			1 188 000
Total requested budget			**3 400 000**
Expenses (A)			440 000
Expenses (B)			198 000
Management (C)			238 000
Overheads			309 000

	Contribution to expenses		Management	Other expenses
Partner	(A)	(B)	(C)	(D)
1	40000	18000	88700	-
2	40000	18000	15000	-
3	40000	18000	15000	-
4	40000	18000	15000	-
5	40000	18000	15000	-
6	40000	18000	15000	-
7	40000	18000	15000	-
8	40000	18000	15000	-
9	40000	18000	15000	-
10	40000	18000	15000	-
11	40000	18000	15000	-
Total	440000	198000	238000	-

Table 9: Indicative financial information on the network project (excluding expenses related to the recruitment of early-stage and experienced researchers).

8 Previous proposals and contracts

8.1 Proposals under evaluation in 2003

This proposal is a resubmission of a similar network previously rejected:

VISIONET-TRAIN — proposal number FP6-504829.

The following changes have been made:

- The number of partners was reduced from 17 to 11,
- The number of early-stage researchers was reduced from 20 to 11,
- The number of experienced researchers was increased from 7 to 16, i.e., 198 person-months,
- A detailed work plan is provided at section 1.5,
- More details are provided about the recruitment and network involvement of experience researchers,
- A detailed management structure is described in section 4 including an in-depth description of all management activities,
- A detailed break down of the requested budget is given in section 7 thus allowing the allocation of financial resources to the various network activities (internal meetings, dissemination, management activities, etc.).
- An in-depth description of the impact is provided in section 2.2.
- The relevance of the proposed research, training, and transfer of knowledge is described in section 5.

A network of excellence was submitted under the FP6-2003-IST-2, Objective 2.3.2.4 (Cognitive systems):

VISIONET — proposal number FP6-004309.

Partner 1 (INRIA) is the coordinator of both the **VISIONET network of excellence** proposal and of the **VISIONTRAIN research training network** proposal. VISIONTRAIN has 11 partners and VISIONET has 15 partners. 9 partners participate to both networks.

The two proposals compliment each other and the activities of VISIONTRAIN and VISIONET will be harmonized if the two proposals are accepted. A number of activities such as **infrastructure**, **spread of excellence and dissemination**, as well as some of the **management activities** will be coordinated.

Nevertheless, the two proposals ARE NOT inter-dependent: the rejection of VISIONET will have no implication onto the successful achievement of VISIONTRAIN, and the other way around.

8.2 Existing EC contracts

Similar **existing sources** of community support:

- IMAVIS is an European project run by the Odyssee team at INRIA (partner 1) under the *Marie Curie Fellowships – Training Sites Programme*, the project started in 2001 for a duration of 4 years, the project number is HPMT-CT-2000-00040. Within this project, visits from partners 2, 3, and 6 are expected during 2003 and 2004.
- ECVision - European Research Network for Cognitive AI-enabled Computer Vision Systems (Thematic Network), 01.03.2002 - 28.02.2005, Contract number: IST-2002-35454, partners 1,6,7, and 9 are involved in this project.

9 Other issues

Does the research presented in this proposal raise sensitive ethical questions related to:	**YES**	**NO**
Human Beings		×
Human biological sample		×
Personal data (whether identified by name or not)		×
Genetic information		×
Animals		×

Table 10: Sensitive etchical questions.

We confirm that the research presented in this proposal does not involve any of the followings:

- research activity aimed at cloning for reproductive purposes;
- research activity intended to modify the genetic heritage of human beings which could make such changes heritable;
- research activity intended to create human embryos solely for the purpose of research or for the purpose of stem cell procurement, including by means of somatic cell nuclear transfer;
- research involving the use of human embryos or embryonic stems cells.