Description of work (Annex 1)

PART A: CONTRACT DETAILS AND OBJECTIVES

1. **Full title:** Computational and Cognitive Vision Systems: A Training European Network

   Short title: VISIONTRAIN

2. **Proposal number:** 005439
   Contract number: MRTN-CT-2004-005439
3. **Duration of the project:** 48 months

4. **Contractors and places of implementation:**

	The coordinator		*Established in:*
1	Institut National de Recherche en Informatique et Automatique	INRIA	FR
	Other contractors		
2	Scuola Internazionale Superiore di Studi Avanzati	SISSA	IT
3	Malmo University	MAH	SE
4	IT-Universitetet i Kobenhavn	ITU	DK
5	Technion - Israel Institute of Technology	Technion	IL
6	The Chancellor, Masters and Scholars of the University of Oxford	UOXF.DF	UK
7	Univerza v Ljubljani	UL	SI
8	Ecole Polytechnique Fédérale de Lausanne	EPFL	CH
9	Czech Technical University in Prague	CTU	CZ
10	Universiteit Utrecht	UU	NL
11	University of Mannheim	U. Mannheim	DE

5. Project overview

Abstract: Visiontrain addresses the problem of understanding vision from both computational and cognitive points of view. The research approach will be based on formal mathematical models and on the thorough experimental validation of these models. Under this framework we intend to reduce the gap that exists today between biological vision (which is by large not yet understood) and computer vision (which is biologically inspired and whose flexibility, robustness, and autonomy remain to be demonstrated). In order to achieve these ambitious goals, 11 academic partners plan to work cooperatively on a number of targeted research objectives: (i) computational theories and methods for low-level vision, (ii) motion understanding from image sequences, (iii) learning and recognition of shapes, objects, and categories, (iv) cognitive modelling of the action of seeing, and (v) functional imaging for observing and modelling brain activity. In practice these objectives will be achieved under the form of 11 PhD theses, each of one achieved collaboratively between 3 teams, as well as within-network mobility of 12 person-years of experienced researchers, each one of these researchers having the opportunity to visit other teams in the network. The training programme includes participation to proof-of-concept achievements, annual thematic schools, industrial meetings, application of patents, attendance of conferences, and a women and science activity.

5.1 Overall objectives

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 the following topics will be addressed:

Topic 1 : 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, color, texture, motion, and binocular cues such as depth.

Topic 2 : One of the major goals of vision is to be able to recover **the 3-D geometric layout**, the **light field**, and the **material properties** as seen from images. This task is tremendous and one objective is to further investigate shape-from-X processes and, possibly, suggest variational approaches.

Topic 3 : **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;

Topic 4 : 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.

Topic 5 : 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.

Topic 6 : 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;

Topic 7 : **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;

Topic 8 : 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;

Topic 9 : 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.

Topic 10 : 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.**

Topic 11 : **Proof of concept** is a transversal topic which addresses the problem of practically building vision systems. We will identify applications that need computational and cognitive models integrated to the system in order to achieve the required capabilities. We will build demonstrators and develop software packages.

5.2 Overall approach and methodology

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 12 experienced researchers.

Breakdown in work-packages

WP	Work-package name	Description
WP1	Computational theories and methods for low-level vision	§1.1.1 - page 5
WP2	Motion understanding from image sequences	§1.1.2 - page 6
WP3	Learning and recognition of shapes, objects, and categories	§1.1.3 - page 7
WP4	Cognitive modelling of the action of seeing	§1.1.4 - page 8
WP5	Functional imaging for observing and modelling brain activity	§1.1.5 - page 9
WP6	Network dissemination activities	pages 23–27
WP7	Network management activities	§2 - pages 28–38

	Start/End	Partners	Responsible
WP1	6/48	1,2,3,4,5,8,9,10,11	CTU (9)
WP2	6/48	1,3,4,8,9,11	MAH (3)
WP3	6/48	1,4,6,7,11	UL (7)
WP4	6/48	1,4,5,6,7,8	Technion (5)
WP5	6/48	1,2,4,10	SISSA (2)
WP6	1/48	all	INRIA (1)
WP7	1/48	all	INRIA (1)

PART B: IMPLEMENTATION

1 Description of joint research/training project

1.1 Research

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.*

The key elements of the research methodology that will be followed is explained in detail below. In practice there will be 5 research topics that will be developed, validated and implemented.

Ethical issues. There are no ethical issues involved in this work.

1.1.1 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 regularizer can also be a non-quadratic function. This allows for discontinuities in the surface derivatives. (iii) The regularizer may also be anisotropic. This strategy enables e.g. a stronger smoothing along a discontinuity than across it. (iv) The regularizer 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.

1.1.2 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.

1.1.3 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.

1.1.4 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.

1.1.5 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.

1.1.6 Proofs of concept

This objective 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 human motion capture 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*. The **proof of concept** objective will be launched in conjunction with the network's work-packages WP1 to WP5. It is planned that each early-stage and experience researcher will participate to at least one proof-of-concept task during her/his network training.

1.2 Detailed description broken down into work-package tasks

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. Figure 1 is a graphical representation of the work-packages.

WP	Work-package name	Description
WP1	Computational theories and methods for low-level vision	§1.1.1 - page 5
WP2	Motion understanding from image sequences	§1.1.2 - page 6
WP3	Learning and recognition of shapes, objects, and categories	§1.1.3 - page 7
WP4	Cognitive modelling of the action of seeing	§1.1.4 - page 8
WP5	Functional imaging for observing and modelling brain activity	§1.1.5 - page 9
WP6	Network dissemination activities	pages 23–27
WP7	Network management activities	§2 - pages 28–38

	Start/End	Partners	Responsible
WP1	6/48	1,2,3,4,5,8,9,10,11	CTU (9)
WP2	6/48	1,3,4,8,9,11	MAH (3)
WP3	6/48	1,4,6,7,11	UL (7)
WP4	6/48	1,4,5,6,7,8	Technion (5)
WP5	6/48	1,2,4,10	SISSA (2)
WP6	1/48	all	INRIA (1)
WP7	1/48	all	INRIA (1)

Table 1: Summary of project work-packages

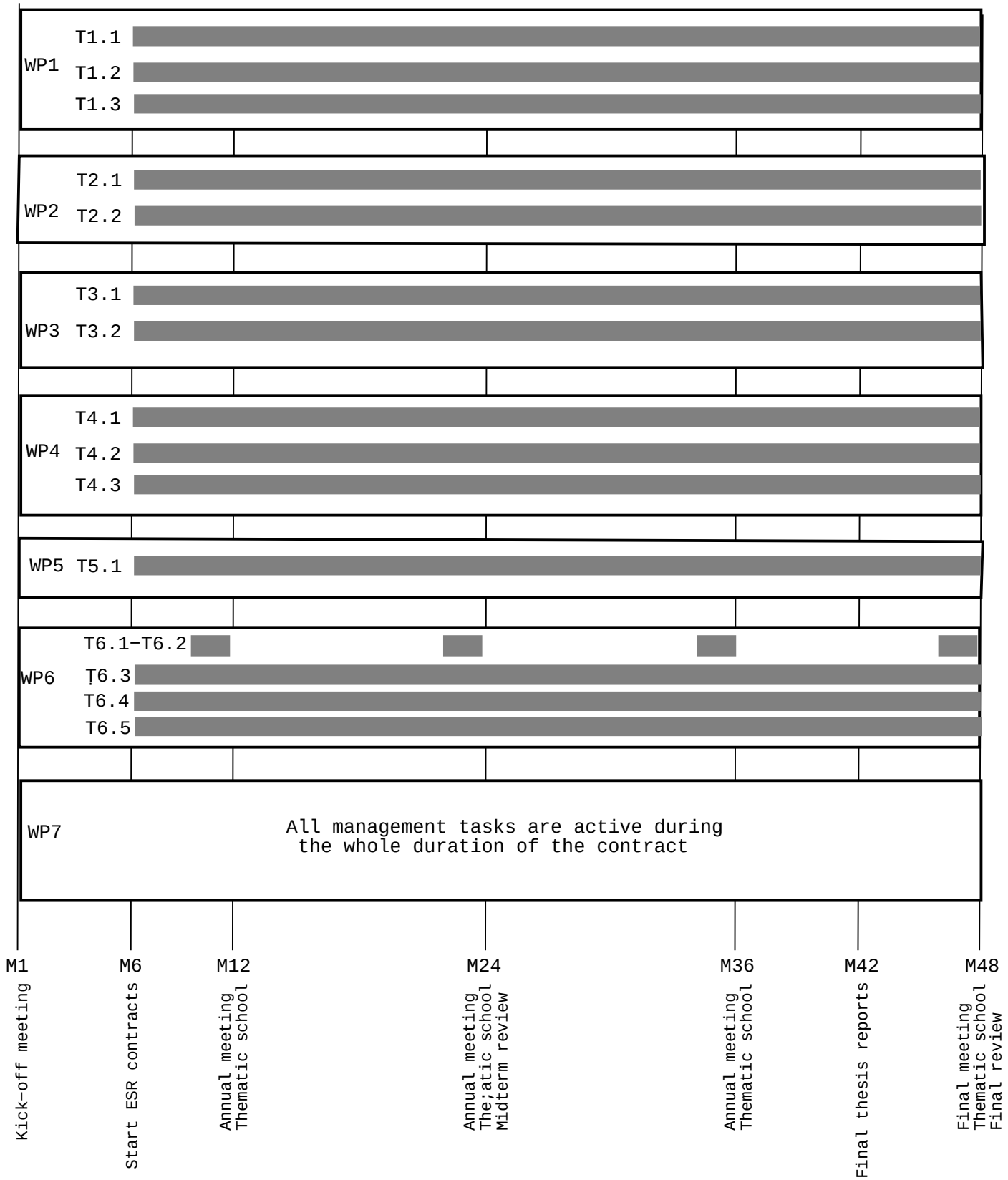

Figure 1: Graphical presentation of work packages.

Work-package 1

Work-package name: Computational theories and methods for low-level vision.

Work-package leader: Partner 9.

Description of work: §1.1.1 - page 5.

Task 1.1 : Monocular visual cues and their interactions.

- Partner 10 will investigate formal theories and test whether human vision models may or may not function along these theories.
- Partner 3 will develop methods for recovering three-dimensional shape from monocular cues.
- Partner 5 will integrate monocular cue extraction with top-down attention mechanisms.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.10, M36.1, M36.2, M36.6, M42.1, M42.2, M48.1, M48.4, M48.5, M48.10.

Adresses: Topic 1, Topic 2, Topic 8, Topic 10, Topic 11, pages 2–3.

Task 1.2 : Variational approaches to low-level vision

- Partner 2 will investigate methods to segment images into regions irrespective from their meaning.
- Partner 4 will explore a variational framework for the integration of cues such as color, texture, and motion.
- Partner 11 will design novel variational approaches for solving shape from shading problems.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.10, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.10.

Adresses: Topic 1, Topic 2, Topic 10, Topic 11, pages 2–3.

Task 1.3 : Stereoscopic matching

- Partner 9 will formulate the stereo matching problem as a super-stable kernel in a special graph.
- Partner 1 will investigate the integration of depth (stereo) cues with monocular cues (color and texture).
- Partner 8 will provide methods for modelling the integration of depth with image contours and silhouettes.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.11, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.11.

Adresses: Topic 3, Topic 4, Topic 1, Topic 11, pages 2–3.

Work-package 2

Work-package name: Motion understanding from image sequences.

Work-package leader: Partner 3.

Description of work: §1.1.2 - page 6

Task 2.1 : Robust feature tracking in image sequences.

- Partner 3 will provide methods for simulteanously solving for motion estimation and sensor auto-calibration.
- Partner 4 will use scale-space theories to extract significant and invariant featuers.
- Partner 11 will investigate methods to parameterize the surface of the 3-D shapes to be extracted.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.11, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.10

Adresses: Topic 4, Topic 5, Topic 11, pages 2–3.

Task 2.2 : Human shape and motion from multiple image sequences.

- Partner 8 will combine visual data (shape and silhouettes) with sophisticated bio-mechanical models for articulated shapes.
- Partner 1 will provide methods for recovering 3-D surface patches from several videos.
- Partner 9 will develop fast 3-D shape extraction methods.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.11, M36.1, M36.2, M36,5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.11.

Adresses: Topic 5, Topic 6, Topic 11, pages 2–3.

Work-package 3

Work-package name: Learning and recognition of shapes, objects, and categories

Work-package leader: Partner 7.

Description of work: §1.1.3 - page 7

Task 3.1 : Shape priors for detection, segmentation, and recognition.

- Partner 11 will study mathematical shape representations and the feasibility of statistical learning of shape variations.
- Partner 1 will study methods for combining shape priors with variational approaches to image segmentation.
- Partner 7 will bring its expertise in sub-space methods for representing shapes.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.12, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.12.

Adresses: Topic 6, Topic 7, Topic 11, pages 2–3.

Task 3.2 : Learning and recognition of objects classes and categories.

- Partner 6 will develop generative and probabilistic models for objects and their parts
- Partner 4 will bring its expertise in scale-space analysis for the representation and recognition of classes of objects.
- Partner 7 will develop methods to reduce the dimensionality of visual shape representations.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.12, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.12.

Adresses: Topic 7, Topic 6, Topic 11, pages 2–3.

Work-package 4

Work-package name: Cognitive modelling of the action of seeing.

Work-package leader: Partner 5.

Description of work: §1.1.4 - page 8.

Task 4.1 : Computational attention for visual search.

- Partner 5 will develop search algorithms able to implement the purposive nature of visual perception.
- Partner 4 will develop a multiple-level representation for image sequences.
- Partner 6 will investigate bottom-up shape learning methods in conjunction with top-down search mechanisms.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.10, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.10.

Adresses: Topic 8, Topic 9, Topic 1, Topic 7, Topic 11, pages 2–3.

Task 4.2 : Robust approaches to sub-space based visual processes.

- Partner 7 will develop robust sub-space methods able to work in uncontrolled environments.
- Partner 6 will develop general-purpose classifiers based on sub-space methods.
- Partner 8 will develop representations of highly articulated objects in terms of sub-space statistical descriptors.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.12, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.12.

Adresses: Topic 7, Topic 5, Topic 8, Topic 11, pages 2–3.

Task 4.3 : Human action and gesture recognition and interpretation.

- Partner 1 will develop methods to recognize complex human gestures and to interpret them within a cognitive framework.
- Partner 8 will investigate the link between the use of motion capture data both for animation and for recognition.
- Partner 5 will model the process of human action recognition from the viewpoint of an active observer.

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.11, M36.1, M36.2, M36.5, M42.1, M42.2, M48.1, M48.4, M48.5, M48.11.

Adresses: Topic 9, Topic 5, Topic 11, pages 2–3.

Work-package 5

Work-package name: Functional imaging for observing and modelling brain activity.

Work-package leader: Partner 2.

Description of work: §1.1.5 - page 9.

Task 5.1 : Quantification and interpretation of the activity of cortical entities.

- Partner 4 will develop variational and scale-space methods for analyzing and measuring the brain activity.
- Partner 10 will provide a framework to compare formal mathematical model with psychophysical data.
- Partner 1 will model the way the brain solves a visual task at a level higher than the voxel representation.
- Partner 2 will provide variational and PDE-based methods for the analysis of various brain imagery modalities (MRI, EEG, MEG).

Milestones: M12.1, M12.2, M12.5, M24.1, M24.2, M24.5, M24.6, M24.10, M36.1, M36.2, M36.6, M42.1, M42.2, M48.1, M48.4, M48.5, M48.10.

Adresses: Topic 10, Topic 2, and Topic 9, pages 2–3.

Work-package 6

Work-package name: Dissemination and exploitation

Work-package leader: INRIA (partner 1)

Description of work: Organisation and implementation of dissemination and transfer of knowledge activities.

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

Milestones: M12.6, M24.13, M36.7, M48.13.

Task 6.2: Organisation and participation to network workshops.

Milestones: M18.1, M42.3.

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

Milestones: M12.6, M24.13, M36.7, M48.13, M18.1, M42.3, M24.8, M48.7.

Task 6.4: Participations to international workshops, conferences, and tutorials.

Milestones: M12.1, M24.1, M24.8, M36.1, M48.1, M48.7.

Task 6.5: Patent application activities.

Milestones: M24.8, M48.7.

Work-package 7

Work-package name: Management

Work-package leader: INRIA (partner 1)

Description of work: The following tasks will be active during the lifetime of the project. They correspond to the management activities described in section 2 at pages 28–38.

Task 7.1: Setup and maintain a network infrastructure

Milestones: M1.2

Task 7.2: Communication with the EC and coordination of reporting

Milestones: M1.1, M12.3, M12.4, M24.3, M24.4, M24.8, M24.9, M36.3, M36.4, M48.2, M48.3, M48.7, M48.8, M48.9.

Task 7.3: Organisation of external audits and reviews

Milestones: M1.1, M12.3, M24.3, M24.7, M24.9, M36.3, M48.2, M48.6, M48.9.

Task 7.4: Organisation of internal meetings and evaluations

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

Task 7.5: Financial management

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

Task 7.6: Technical management

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

Task 7.7: Legal and knowledge management

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

Task 7.8: Human resources management

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

Task 7.9: Strengthen the role of women

Milestones: M1.1, M12.3, M24.3, M36.3, M48.2.

1.3 Scientific milestones and deliverables

Table 2: List of scientific milestones and deliverables (continues on next page).

Milestone	Month	Description	Comment
M1.1	1	Kick-off meeting	Tasks 7.2 to 7.9
M1.2	1	Website	Task 7.1
M12.1	12	Activity reports	Tasks 1.1 to 5.1
M12.2	12	Thesis reports	Tasks 1.1 to 5.1
M12.3	12	Annual meeting	Tasks 7.2 to 7.9
M12.4	12	Management report	Task 7.2
M12.5	12	Scientific deliverable	Completion state-of-the-art Tasks 1.1 to 5.1
M12.6	12	Thematic school	Tasks 6.1 and 6.3
M18.1	18	Workshop	Tasks 6.2 and 6.3
M24.1	24	Activity reports	Tasks 1.1 to 5.1
M24.2	24	Thesis reports	Tasks 1.1 to 5.1
M24.3	24	Annual meeting	Tasks 7.2 to 7.9
M24.4	24	Management report	Task 7.2
M24.5	24	Midterm demonstration	Tasks 1.1 to 5.1
M24.6	24	Software packages	Tasks 1.1 to 5.1
M24.7	24	Audit certificates	Task 7.3
M24.8	24	Midterm report	Task 7.2
M24.9	24	Midterm review	Tasks 7.2, 7.3
M24.10	24	Scientific deliverable	Joint intermediate results Tasks 1.1, 1.2, 4.1, 5.1
M24.11	24	Scientific deliverable	Joint intermediate results Tasks 1.3,2.1,2.2,4.3
M24.12	24	Scientific deliverable	Joint intermediate results Tasks 3.1, 3.2, 4.2
M24.13	24	Thematic school	Tasks 6.1 and 6.3

Table 3: Continued from previous page: list of scientific milestones and deliverables.

Milestone	**Month**	**Description**	**Comment**
M36.1	36	Activity reports	Tasks 1.1 to 5.1
M36.2	36	Thesis reports	Tasks 1.1 to 5.1
M36.3	36	Annual meeting	Tasks 7.2 to 7.9
M36.4	36	Management report	Task 7.2
M36.5	36	Scientific deliverable	Completion theoretical studies Tasks 1.2, 1.3, 2.1, 2.2, 3.1, 3.2, 4.1, 4.2, 4.3
M36.6	36	Performance comparisons	Tasks 1.1 and 5.1
M36.7	36	Thematic school	Tasks 6.1 and 6.3
M42.1	42	Final thesis reports	Tasks 1.1 to 5.1
M42.2	42	Experimental validation	Tasks 1.1 to 5.1
M42.3	42	Workshop	Tasks 6.2 and 6.3
M48.1	48	Activity reports	Tasks 1.1 to 5.1
M48.2	48	Annual meeting	Tasks 7.2 to 7.9
M48.3	48	Management report	Task 7.2
M48.4	48	Final demonstration	Tasks 1.1 to 5.1
M48.5	48	Software packages	Tasks 1.1 to 5.1
M48.6	48	Audit certificates	Task 7.3
M48.7	48	Final report	Task 7.2
M48.8	48	Final Man. report	Task 7.2
M48.9	48	Final review	Tasks 7.2, 7.3
M48.10	48	Scientific deliverable	Final joint results Tasks 1.1, 1.2, 4.1, 5.1
M48.11	48	Scientific deliverable	Final joint results Tasks 1.3,2.1,2.2,4.3
M48.12	48	Scientific deliverable	Final joint results Tasks 3.1, 3.2, 4.2
M48.13	48	Thematic school	Tasks 6.1 and 6.3

1.4 Training and transfer of knowledge (ToK)

The network as a whole undertakes to provide a minimum of 540 person-months of Early Stage and Experienced Researchers whose appointment will be financed by the contract. Quantitative progress on this, with reference to the table contained in part C and in conformance with relevant contractual provisions, will be regularly monitored at the consortium level.

1.4.1 General training and ToK

The backbone of the transfer of knowledge programme is the successful implementation **at the network level** of 11 PhD theses undertaken by the recruited early stage researchers as well as collaborative work between 12 experienced researchers and network researchers.

Each early-stage researcher to be recruited 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.

The proposed assignment of the main and secondary partners for each PhD thesis preparation will be decided at the project kick-off meeting. This assignment will maximize exposure of each early-stage researcher to three different European groups with complementary 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.

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

1.4.2 Training list at local and network level

Spread of excellence. During the lifetime of the project a number of conferences, workshops, and tutorials will be organized through Europe.

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

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

Thematic schools. We plan to launch a new event named the *European Computer Vision School* or ECVS. ECVS will take place once a year, possibly 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

Month	Event	Milestones	Responsible	Participating
1	Kick-off meeting	M1.1	Partner 1	all other partners
12	Thesis reports	M12.2	all partners	all other partners
12	Thematic school	M12.6	Partner 9	all other partners
12	Annual meeting	M12.3	Partner 1	all other partners
18	Workshop	M18.1	Partner 1	all other partners
24	Thematic school	M24.13	Partner 1	all other partners
24	Thesis reports	M24.2	all partners	all other partners
24	Annual meeting	M24.3	Partner 1	all other partners
24	Midterm report	M24.8	Partner 1	all other partners
24	Midterm review	M24.9	Partner 1	all other partners
36	Thematic school	M36.7	Partner 7	all other partners
36	Thesis reports	M36.2	all partners	all other partners
36	Annual meeting	M36.3	Partner 1	all other partners
42	Workshop	M42.3	Partner 1	all other partners
42	Final thesis reports	M42.1	all partners	all other partners
48	Thematic school	M48.13	Partner 3&4	all other partners
48	Annual meeting	M48.2	Partner 1	all other partners
48	Final review	M48.9	Partner 1	all other partners

Table 4: Training and Transfer of Knowledge Schedule

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 12 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 address issues such as: intellectual property rights, software commercialization and dissemination, patent applications protocols, success stories, etc.

Currently the following ECVS sessions are already planned:

- 2006 – Prague (Czech Republic), organized by partner 9: professor Vasek Hlavac and Dr. Martin Urban
  Topic: Low-level image formation and understanding, crossroad of biological and computational methods .

- 2007 – Sophia-Antipolis (France), organized by partner 1: Dr. Rachid Deriche
  Topic: Variational and PDE methods and their usefulness for vision.

- 2008 – Ljubljana (Slovenia), partner 7: professor Ales Leonardis
  Topic: Learning and recognition of objects and categories.

- 2009 – Malmo/Copenhagen (Sweden/Denmark), partners 3 and 4: professors Anders Heyden and Mads Nielsen.
  Topic: Geometrical computer vision.

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.

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.

Workshops. The network will organize two workshops. During the first workshop the ESR will present their preliminary and intermediate results. During the second workshop both the ESR and the ER will present their final results. The workshop will be open to participants outside the network (both academic and industrial).

1.4.3 Multidisciplinary and intersectorial knowledge

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.

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.

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.

1.4.4 Integration of young researchers

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.

Training concerning non-scientific issues. Both the early-stage researchers and experienced researchers will be kept informed about non-scientific matters such as: their contractual rights and obligations, intellectual property rights, confidential and non-confidential policies of their organism, career development plans within their organism and, on a broader basis, within the host country.

They will also be continuously informed about the network training events (the VISIONTRAIN website will maintain information about the courses offered by the universities and institutes involved in the network). The website will also contain information about various training events complimentary to purely scientific training.

1.5 Training and transfer of knowledge summary

Mobility of ESR – Each early-stage researcher will spend 30 month with her/his main partner of attachment and 3+3 months with two other partners.

Networked advisory of ESR – Each early-stage researcher will have an advisory committee of 3 researchers from the 3 partners.

Mobility of ER – Each experienced researcher having a 12 month contract with one partner will spend 3 months with another network partner and will work collaboratively with at least two network researchers.

Tutoring of ESR and ER – 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.

Graduate courses – In addition to the main research activity, each early-stage researcher will attend courses from a list of courses established with her/his advisors.

Teaching assistantship – Each early-stage researcher will be assigned a *teaching assistantship* in coordination with her/his main advisor and with the partner's doctoral school.

Thematic school – Each early-stage researcher will attend ALL four sessions of the *European Computer Vision School – ECVS* and will present her/his work twice.

Spread of excellence – Each early-stage and experienced researcher will be expected to publish their work at major journals and conferences. They will be assisted by the network researchers.

Workshops – There will be two network workshops allowing the ESR and the ER to present their work to a large audience.

Patent application – Early-stage and experienced researchers will be provided with support to submit patents to the European patent office.

Intra-network short visits/meetings – Early-stage and experienced researchers will work collaboratively with network researchers within an intra-network mobility activity.

Industrial collaboration – Each early-stage and experienced researcher will have the opportunity to present her/his results to companies both at the network level, i.e., workshops and through individual company visits.

Women and science – Organisation of forums and seminars, recruitment of an equal number of women and men, establish an observatory to measure the role of women in Visiontrain's research field.

2 Management

2.1 Organizational architecture

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 stake holders within the scientific field, have an interest of maintaining the excellence represented by VISIONTRAIN.

  Experts from outside the network will not be supported by network funds for their participation to it.

- 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.

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.

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,

2.2 Coordination and decision making process

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 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

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 authorized 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 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
- 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 subsection.

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;
- 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.

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

- 6 experts from inside the network;
- 6 experts from outside the network not supported by network funds: 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.

2.3 Communication method

The infrastructure will 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.

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;

2.4 Task delegation policy

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.

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.

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. 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.

2.5 Vacancies strategy

Early-stage researchers The recruitment of ESR will be performed on the following grounds:

- A VISIONTRAIN website will become active at month 1;
- 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 one fixed deadline to apply for an ESR position, and this deadline will be fixed such that the work contracts start at month 6;
- Once the call for applications is closed, the *Management board – MB* will verify the eligibility of the candidates with respect to the Marie-Curie criteria. 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, two among the VISIONTRAIN researchers and one 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.

- An ordered list of applicants who are admitted for recruitment will be made available by the Management board. 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 will be by 50% larger than the number of open positions.

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. However, internal and external applicants will be treated on an equal opportunity basis.

There will be 3 deadlines for applying for a VISIONTRAIN experienced researcher position: at months 6, 12, and 18.

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 who the candidate wishes to collaborate.

Once the call for applications is closed, the *Management board – MB* will 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, two from the network and one 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 will publish an ordered 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. Up to 30% of the retained candidates could be from third countries;
5. The female/male ratio should be as close as possible to 1:1.

Publicity The open positions for both early-stage and experienced researchers will be jointly published by the network.

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 on their websites.

2.6 Equal opportunity measures

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.

2.7 Monitor and reporting method

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.

2.8 Dissemination of results policy

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.

The consortium agreement details articles II.29, II.30, II.31, II.32, II.33, III.9, and III.10 of the contract with the European commission (reference number MRTN-CT-2004-5439), and it cannot contradict with these articles under any circumstances. Articles II.29, II.30, II.31, II.32, II.33, III.9, and III.10 of the above mentioned contract with the European commission always prevail to all terms laid out in this Annex I that describes the implementation of the IPR issues.

2.9 Financial management and control

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, 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 every 24 months and submitting them to the EC.

The project coordinator, 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 project reviews,
- Financial audits, and
- All other audits that the EC wishes to organize.

3 Indicators of progress and success

3.1 Quantitative indicators of progress and success to be used to monitor the project

3.1.1 Research activities

- The list and number of publications authored by network researchers from different partners directly related to the work undertaken;
- The list and number of publications co-authored by network researchers from different disciplines directly related to the work undertaken;
- The list and number of publications authored and co-authored by Visiontrain's early stage researchers directly related to the work undertaken;
- The list and number of publications authored and co-authored by Visiontrain's experienced researchers directly related to the work undertaken.
- List of scientific awards and prizes obtained from the work directly connected to the network.
- List of libraries, packages, and prototypes produced by the ESR and ER.
- List of workshops, seminars and conferences at which network researchers were invited to present their work.
- The list and names of network patents;

3.1.2 Training/Transfer of knowledge activities

- The actual number and names of ESR and ER hired by each partner;
- For each ESR and ER actually hired, the number of person-months spent with each network partner.
- The number and names of other researchers (not financed by the network) holding a permanent position with their institution and being involved in the research;
- The number and names of other researchers holding a temporarily work contract with their institution and paid on other contracts than Visiontrain, which are involved in the research;
- The number and names of PhD students holding a temporarily work contract with their institution and paid on other grants and contracts than Visiontrain, and which will participate to the research and transfer of knowledge activities;
- The number and names of network PhD students (ESR) hired by the partners after completion of their degree.

- The number, periods of time, and names of early stage researchers visiting other partners and being seconded, sorted by disciplines;
- The number, periods of time, and names of experience researchers visiting one or several network sites and being seconded;
- The number and names of other researchers participating to the intra-network working groups and visits.
- The number and names of network PhD students attending the *Thematic schools*;
- The number and names of PhD students from outside the network attending the *Thematic schools*;
- The number and percentage of women researchers in the network;
- The number and percentage of women hired as early-stage researchers;
- The number and percentage of women hired as experienced researchers;
- The number and percentage of women involved in the **Technology transfer to companies** activity;
- The number and percentage of women who applied for a patent;
- The number and percentage of articles authored and co-authored by women;
- The number and percentage of women participating to the network's events to spread excellence, such as *annual industrial meetings*, *annual thematic schools*, *tutorials*, *workshops*, etc.
- Organisation of training events (e.g. schools, training workshop/seminar, hands-on training session on specialized instrument/techniques) at individual participant sites (number, attendees, names, place, date);
- Organisation of network-wide training events (number, attendees, names, place, date).

3.2 Qualitative indicators of progress and success to be used to monitor the project

3.2.1 Research activities

- Effectiveness of interdisciplinary nature of the work carried out.
- Highlights of innovative developments by ESR and ER.
- Scientific/technological breakthroughs.
- Nature and justification of adjustments with respect to the initially planned work.
- Interactions between partners and potential industrial users.

- Highlights of the recognition of the network activities by the international scientific community.

3.2.2 Training/Transfer of knowledge activities

- Career development plans as elaborated by ESR and ER involved in the network.
- Career opportunities for ESR and ER.
- Level of satisfaction of the trainees.
- Innovative approaches to training within the particular discipline of the network.
- Exploitation of complimentarity and interdisciplinary skills of partners for the training and transfer of knowledge.

3.2.3 Management

- Effectiveness of internal communication between the management bodies, the ESR, and the ER including feedback processes.
- Effectiveness of the communication between the network and the Commission Services, particularly regarding the conformance with contractual provisions.
- Effectiveness of network communication with industrial and other stakeholders.
- Network self-assessment through benchmarking activities (exchange of best practices among participants and/or development of ad hoc performance indicators regarding cost management, staff selection, measurement of research/training/ToK outputs, young researchers involvement, etc.).
- Overall quality and efficiency of the external communication strategy of the network (Cordis; personal, team and network web sites updates; newsletters; etc.).
- Effectiveness of the recruitment strategy of the network in terms of equal opportunites (including gender balance) and open competition at international level.
- Development of any specific planning and management tool(s) and databases.
- Management of intellectual property of network research output.

PART C: CONTRACT DELIVERABLES (from A4b of the CPF forms)

Proposal Number	005439	Proposal Acronym	VISIONTRAIN

OVERALL INDICATIVE PROJECT DELIVERABLES BY PARTICIPANT						
Participant No.	**Early Stage Researchers**			**Experienced Researchers (4-10 years – MCRTN only)**		
	Full-time Person Months	Indicative number of researchers	Stipend (%)	Full-time Person Months	Indicative number of researchers	Stipend (%)
1	36	1	0%	24	2	0%
2	36	1	0%	12	1	0%
3	36	1	0%	12	1	0%
4	36	1	0%	12	1	0%
5	36	1	0%	12	1	0%
6	36	1	0%	12	1	0%
7	36	1	0%	12	1	0%
8	36	1	0%	12	1	0%
9	36	1	0%	12	1	0%
10	36	1	0%	12	1	0%
11	36	1	0%	12	1	0%
Sub-Total	396	11		144	12	

PART D: COMMUNITY CONTRIBUTION (from A5b of the CPF forms)

Proposal Number	**005439**	Proposal Acronym	**VISIONTRAIN**

OVERALL MAXIMUM COMMUNITY CONTRIBUTION									
	Eligible expenses for the activities carried out by the researchers					Eligible expenses related to the activities of the host organisations				Maximum EC contribution
Year	-A- Monthly Living Allowance	Transnational Mobility -B- Travel Allowance	Transnational Mobility -C- Mobility Allowance	-D- Career Exploratory Allowance	-E- Participation expenses of the eligible researchers	-F- Research/ training/ transfer of knowledge	-G- Management and Audit Certification	-H- Overheads	-I- Other types of eligible expenses	
	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	Costs (in euros)	(in euros)
1	204666,00	12000,00	50999,00	24000,00	28800,00	59704,00	17648,00	38017,00	0,00	435834,00
2	661746,00	22000,00	153554,00	22000,00	81600,00	208994,00	61789,00	114990,00	0,00	1326673,00
3	637141,00	11000,00	148529,00	0,00	79200,00	208994,00	61789,00	108486,00	0,00	1255139,00
4	180062,00	0,00	45973,00	0,00	26400,00	119418,00	35307,00	37185,00	0,00	444345,00
5	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
6	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Total	1683615,00	45000,00	399055,00	46000,00	216000,00	597110,00	176533,00	298678,00	0,00	3461991,00