Information Society Technologies (IST) Programme

Contract for:

Shared Cost RTD

Annex 1 - "Description of Work"

Project Acronym: ECOVISION
Project full title: Artificial vision systems based on early-cognitive cortical processing.
Proposal/Contract no.: IST-2001-32114
Related to other contract no.: n/a

Date of preparation of Annex 1: 06. 09. 2001

Operative commencement date of contract: see Article 2.1 of the contract

Contents

1. Project Summary

Objectives:
ECOVISION will build a visual scene analysis system with neuronal architecture guided by cognitive context, which operates in video-real time, thereby providing the first entry point to a new generation of artifacts with a more human-like pattern of intention driven behavior. ECOVISION: 1)contains dynamic feedback loops leading to self-organization of neuronal activity, 2)embeds early cognitive analysis at the level of context adaptive visual receptive fields ("visual cortical RFs") 3) defines a dynamically changing "map-like" representational structure of these RFs weighted by their momentarily existing task-relevance, 4) implements predictive mechanism to generate a self-organizing level of analysis, 5) front-end pre-processing steps are studied, evaluated and adapted to suit in FPGA technology in order to operate in video real-time.

Description of Work:
ECOVISION consists of three main project parts and focuses on a self-navigating system (like a car). Project Parts: A) Feasibility assessment of an FPGA-based front-end for the computation of the motion structure in the visual scenes in video real-time. (The consortium possesses already such a device for the extraction of stereo-depth information). Study of how can FPGA computation primitives be suited to neuron-like processing in order to implement low-level motion-based structure extraction. Project Part B) Design of a high level processing software device with neuronal structure, which analyses the front-end information with respect to its intrinsically existing visual context. *Visual context* is the embedding of Gestalt-laws, established by psychology and also in cortical neuronal processing, in a computer vision system. "Gestalts", thereby, represent pixel groups which have shared statistical properties in space and time. Thus, they define segments in space and time and can, therefore, be captured by highly-adaptive receptive fields (RFs) of ECOVISION's processing units. *Control is intention driven* in that every scene will be analysed with respect to an intentional context within which the momentarily existing Center of Interest (CoI) is determined by the dynamic structure of the visual scene. To this end, the RFs of the units shall be arranged on a map-structure which has an enhanced resolution at the CoI. This map, thus, represents, the momentarily existing visual relevance and the CoI can be regarded as a flexible mental fovea.
To get a deeper insight, these concepts will also be acertained by biophysically realistic models of visual cortical function. Project Part C) Design of the complete ECOVISION system by joining the components from A and B.

Milestones and expected results
M1: is a feasibility assessment of a neuronal stereo- and motion real-time processing unit (FPGA) to be used as front-end in ECOVISION M2: describes the pre-cognitive sensor side of the system where context-sensitive filters are defined. M3: describes the early-cognitive sensor side of the system. It captures visual task-relevance and it defines the Center of Interest leading to an intention driven scene analysis. M4: is the complete system.

2. Program objectives

Currently existing autonomous machines mostly behave in a stereotyped, reflex-like manner with little predictive planning. The behavioral pattern of higher vertebrates, on the other hand, is more sophisticated because visual analysis and motor planning rely heavily on the early cognitive contextual analysis of the scene. In order to come closer to this level of complexity ECOVISION addresses two challenges:

(1) to transfer the basic *goal oriented cognitive properties related to the Gestalt Principles known from psychology* to the designing and building of a new generation of flexible, self-adapting artificial visual scene analysis systems. *Gestalts* are thereby simply the explicit realization (at a neuronal level) of the existing statistical regularities in the world.

(2) to give these systems the capacity to operate *in a real-world environment in video real-time* carrying out a VLSI feasibility assessment of the front-end time-consuming processing steps, adapting neuronal algorithmic schemes to VLSI computational primitives.

Most of the existing artificial vision approaches include very high time-consuming processing steps that make them not suitable for VLSI implementation and therefore with little interest for solving real world problems with severe time constraints, like driving assistance, with the current technology. Bio-inspired vision schemes have special interest from the point of view of their suitability for implementation, since they are based on reverse engineering of existing working devices. The search of bio-inspired computational primitives suitable to be implemented in silicon is not trivial. Technologists must be opportunistic, abstracting of biological systems computational meaningful features while discarding other characteristics related to the intimate nature of biological tissues. On the other hand, these computational primitives have to be adapted to be compatible with some outstanding characteristics and schemes provided by the current technology as shorter time delays, multiplexing techniques, etc. All this, constitutes the main grounds of the neuromorphic engineering research field. The ECOVISION scheme is based mainly on characteristics observed in biological vision systems. Therefore it is of interest to study also how can the current technology take advantage of the proposed processing schemes. Furthermore, it seems also worth, to study how can these vision algorithms be modified to make them suitable for VLSI implementation. This work can only be properly carried out through the cooperative work of technologists and vision researchers. ECOVISION consortium represents an optimum opportunity for facing these topics from a multidisciplinary point of view.

Some high level tasks addressed by ECOVISION rest on concepts that will only be clearly defined by the end of the project, but on the other hand, earlier processing stages like motion based structure extraction are of interest for a VLSI feasibility assessment. Bio-inspired motion processing schemes will be studied and characterized in terms of their suitability for FPGA implementation. This stage is considered one of the front-ends of the whole ECOVISION system and therefore must provide a motion representation compatible with the higher tasks foreseen throughout the proposal.

Conventional active vision systems employ a focus of attention, based on a general analysis of visual saliency in the incoming data, to direct the gaze and reduce the area to be analyzed. This is a rather mechanistic process and leads to stereotypical reactions which are not representative of the flexible behavior of higher vertebrates. Recent advances in the psychophysics of human visual perception, on the other hand, demonstrate that specific goal oriented cognitive routines can guide scene analysis at very early stages going beyond the limitations of stereotypical visual saliency. In addition, cortical neurophysiology (particular in higher visual areas but also in V1) has demonstrated that at the level of single cells visual responses are influenced by the visual context and also by intention and planning.

As a consequence ECOVISION will contain context sensitive receptive fields of its processing units ("Gestalt" sensitive receptive fields) similar to context sensitive visual cortical neurons. In order to perform early cognitive analysis we will introduce the concepts of visual relevance and that of a mental fovea (called Center of Interest, CoI) to the ECOVISION system. Through recursive dynamic feedback processes, contextual information shall be weighted with respect to its relevance. This defines a self-organizing process similar to the self-adaptive mechanisms in higher vertebrates. Visual relevance will be enhanced in the Center of Interest, which can be regarded as a mental fovea, where the resolution of the scene analysis shall be magnified. Thus, these novel concepts extend the known ideas of a *focus of attention* and that of *space variant mappings* to a more cognitive, task-driven domain of scene analysis. Visual relevance and the CoI are the central early cognitive structures of ECOVISION, which go beyond the so far existing approaches of active computer vision.

The use of *visual Gestalts*, the *relevance representation* and the definition of a C*enter of interest* are the three entirely novel early cognitive concepts in the system outlined in the next paragraphs. Steps to be undertaken are:
1) Define contextual entities (Gestalts) and thereby create a coarse grain representation of the image employing „context filters“, whereby context is defined by the dynamically changing target application task.
2) Create a dynamic distributed map-representation of the momentarily existing visual relevance.
3) Define a *center of interest* from the task and by analyzing the structure and relationships between the cognitive entities in the scene.

The Goal: A system such as ECOVISION is intended to constitute a vital sensorial control subcomponent in any self-navigating, cognitive artifact. However, the tremendous complexity of cognitive visual scene analysis can initially only be tackled by restriction to a well-defined visual goal. We have chosen a difficult real-world problem in the specific requirements of a driver assistant system, because this is in the center of interest of many industrial applications and provides also the opportunity for benchmarking with other systems. To this end an industrial partner (Hella Hück KG, Lippstadt) will provide the necessary industrial feedback. The main objective of driver assistant systems is to support the human driver by ways of comfort in most traffic situations e.g. automatic distance and cruise control (ACC) on motorways, but also to enhance safety by freeing his capacity for other tasks. Today radar and lidar-based ACC systems are in the market, but more sophisticated applications like automatic Stop&Go in traffic jam situations and advanced collision avoidance will be required in the near future. In particular, real-time computer vision is a key to increase the performance of such systems. An important issue is to obtain 3D and motion information of the driving situation from computer vision. Surely the determination of dynamic and stationary obstacles by binocular stereo vision is an established method, but motion is a well known scheme to get 3D information of an moving object and also to differ between ego-motion and other moving objects. So in very complex environments like urban situations also the motion/flow-field feature map is needed computed as fast as possible to come to a reliable and robust system.

The structure of the consortium provides optimal synergy between the partners for addressing the different work-packages and for integrating the whole system. The layout of the project, however, is modular such that each step can initially be started to a large degree independent of the others. The individual steps will lead to two major scientific challenges. (A) Exploring and implementing an appropriate algorithm, suitable to be mapped onto specific VLSI hardware for the analysis of the motion structure (specifically the "optic-flow") in the visual scene and (B) developing strategies for task driven context analysis as well as implementation and control of the center of interest which constitutes a novel concept of ECOVISION.

ECOVISION faces a multitude of challenges, several of which are of a pre-competitive explorative character, which makes it well suited for an FET-open project. It consists of sub-components which by themselves require special innovative approaches, but the major challenge lies in the co-operativity of the envisioned system. This will rely heavily on the understanding of the integrative character of the processing of visual context neural systems. Therefore, this project requires a strong inter-disciplinary cooperation between experimental and computational neuroscientists, psychophysisists, chip designers and applied researchers from industry to solve the key problems.

3. Participant list

Partici-pant role	Partici-pant no.	Participant name	Participant short name	Country	Date enter project	Date exit project
C	1	University of Stirling, Prof. F. Wörgötter	Sco	UK	Start of project	End of project
P	2	Universiteit Leuven, Prof. M. vanHulle	Bel	B	Start of project	End of project
P	3	University College London, Prof. A. Johnston	Eng	UK	Start of project	End of project
P	4	Ruhr-Universität Bochum, PD Dr. M. Lappe	Ger	D	Start of project	End of project
P	5	University of Genoa, Prof. G.M. Bisio & Dr. S.P. Sabatini	Ita	I	Start of project	End of project
P	6	Universidad de Granada, Prof. E. Ros	Spa	E	Start of project	End of project
P	7	HELLA KG Hueck&Co, Mr. M. Mühlenberg	Ind	D	Start of project	End of project

4. Contribution to Program Objectives

ECOVISION brings together early cognitive and ecological vision. This double meaning very directly captures the intention behind this emerging system. Next generation vision systems must be able to incorporate cognitive aspects of scene analysis in an ecological way. Thus, visual context needs to be taken into account and this information needs to enter the analysis at an early level. These systems will perform a really purposive scene analysis based on intentions and planning (through task adaptation). They will go beyond notions of "saliency" and "attention", which so far have led to the rather reflexive behavioral patterns of current computer vision systems.

Attempts to build artifacts with cognitive properties are a truly challenging undertaking and currently only a few laboratories (located mainly in the US and Japan) have taken this step (A review issue of IEEE Intelligent Systems, July-Aug. 2000, is devoted to these aspects). These groups have successfully invented artifacts, which adapt to their environment. Using these techniques complete robots with sensor-motor interfaces that operate under real-world conditions have been built, with the aim of generating complexity from reflexive insect-like behavior. However, over the last years it has become clear, that even primary sensory processes in higher vertebrates do not share the simple passive reflex-like properties found in invertebrates and attempts have been started to achieve this level of complexity also in computer vision systems or robots (see again IEEE Intell. Syst., 2000). Already early vision is an active process. Novel, more autonomous systems must, therefore, be able to self-adapt and evolve beyond pure programming. Cognition is even one step beyond self-adaptation! This is due to the fact that it reflects part of the strategies humans (and other higher vertebrates) employ in order to directly control their environment. As a consequence we believe that it is now of great importance to introduce these aspects in artificial vision systems, particularly, because such a synthesis of technology with (early) cognition based on a neurophysiological background has not been attempted so far.

The ECOVISION system, which we are building, reaches a new level of behavioral complexity, because response control by early cognitive scene analysis goes far beyond the often rather stereotyped actions of conventional computer vision systems.

5. Innovation:

5.1 State of the art:

This project is grounded in the fields of computer vision and neuroscience/psychophysics of the visual system. A brief overview over these two fields is given covering, however, only the aspects which concern our project.

Computer Vision: Commonly CV systems consist of input devices (cameras), a computational front-end which performs feature analysis (like edge or color detection, motion analysis, etc.) and an integrating device which tries to combine individual features into a consistent representation of the visual scene. This architecture has been successfully employed since its first concepts were introduced by the work of Marr. Currently concepts in the focus of investigation are the following:

1) A central problem yet to be solved is the design of the "integrating device". To this end different data-fusion techniques have been employed, all of which attempt to disambiguate the results from the individual early vision algorithms. The term *data fusion* thereby describes techniques by which results from the algorithmic analysis using different front-end modules are combined and whole conferences were devoted to this aspect.
2) Over the years some more advanced aspects have been added to CV systems subsumed under the term "*active vision*". Commonly this describes a system which actively reacts to changes in the visual scene, usually through moving the cameras and redirecting the gaze. This involves a multitude of different approaches, most notably those who employ a so called "*focus of attention*" in order to control gaze direction. The focus of attention is a concept which derives from visual psychophysics (psychology) and neurophysiology and describes an area in the visual field where an increased saliency of a feature in the momentarily viewed scene elicits enhanced attentiveness. Salient features are, for example, fast moving colorful objects. A normal reaction is then to look at the salient object. Such active vision systems, however, react in a rather stereotyped, reflex-like way, a behavior which is not observed in humans (because we can voluntarily decide not to look at a salient object).
3) In order not to freeze on this particular object its assumed saliency value must then gradually decay such that other (new) objects can become a new focus of attention ("inhibition of return"). This lead to the introduction of the concept of "*visual adaptation*" into CV systems. Neuronal responses in the visual system decay strongly if the stimulus isn't changing. Along the same line adaptation in CV systems helps to neglect a constant scenery and instead concentrate on changes termed "*novelty detection*". *Habituation* is another concept (somewhat similar to adaptation) borrowed from biology which entered CV: organisms *learn* to ignore repeatedly presented identical (or similar) stimuli. This process goes beyond adaptation because it also includes the aspect of learning.
4) *Processing speed* is another central problem for CV systems, despite of the increasing computer power. Therefore, several groups have attempted to accelerate their systems by employing special hardware, which can range from high level hardware (e.g. transputers, dedicated multi-processor boards, etc.) to special micro-chips for early vision processing. Here exists still a high demand to explore new algorithmic avenues and to find efficient ways for high density – high speed VLSI integration.
5) Neuromorphic engineering has been usually related with *analog* VLSI circuits, however some models (like "rank order neural coding") may be efficiently implemented by digital reconfigurable hardware as FPGAs. Indeed, as digital programmable technology has evolved providing, high performance rates, research and design efforts have to be done in order to explore the chances of these technologies to implement more complex and robust neuromorphic approaches. One of the original ideas of Neuromorphic research paradigm states that the complex processing tasks carried out in living systems possible due to architectural topologies that facilitate massively distributed computation. To some extent biological and technological computational primitives share intrinsic characteristics, this motivates the circuit design community to be keen of knowing specific details about the way biological perception and other complex processing tasks are carried out by computational units (neurons or groups of neurons) with severe timing constraints. Ideas taken from reverse engineering of natural systems have been extremely profitable for implementing circuits with new bio-inspired processing schemes.

Some of the above show that several aspects from human vision have successfully influenced computer vision. CV, however, is still far away from human perception capabilities and is in general still not able to implement visual cognitive aspects in a successful way.

(Artificial) Visual Cognition: Commonly visual cognition is thought to be a high level process taking place in higher visual cortical areas and especially in the association cortex. This so-called higher cognition is based on the interpretation of more general sub-components in the visual scene and the analysis of relations between them. While this is certainly of great importance, novel results demonstrate that significant early-cognitive influences already arise at the level of the lower visual cortical areas. Most notably here one finds substantial contextual influences on the cell responses which can change as the consequence of arousal, motivation and attention CV approaches have acknowledged the existence of high level cognitive influences and in order to implement them they have adopted strategies from the field of *artificial intelligence* (like iconic representations). In general this leads to systems in which visual feature (like edges, depth, motion, etc.) are computed from early-vision algorithms and those features are then subjected to a relational analysis. Thus, the important intermediate step of employing contextual influences is left out and these systems have, therefore, a tendency to "jump to conclusions" by drawing contextually-inadequate conjectures. Furthermore, visual features computed from early vision algorithms are usually error-ridden, such that it is rather complicated to subject them directly to a relational analysis. Visual features still embody the aspect of a distributed representation, much less so higher cognitive entities, which have a distinct integrative character. Thus, in summary, it seems that current CV systems leave the distributed stage too fast and integrate information too early. Strangely enough, PDP[1] approaches have had rather small influence on CV so far such that currently the greatest challenge seems to design a distributed computing architecture similar to that in the cortex which embeds the aspects of context analysis *before* integrating the results, for example, employing data fusion mechanisms.

Visual cortical physiology and psychophysics: In the last 5 years visual attention has played a major role in physiological and psychophysical investigations and very recent high-level conferences were devoted to this aspect. However, from this vantage point a continuous shift was observed in two different directions: Several research groups researchers went "upstream" to concentrate on the effects of higher cognitive reasoning and intentional planning (for example of movements) on the cell characteristics in (for example) posterior parietal cortex. Whereas others started to focus on the effects of visual context on the cell responses in the low-level visual areas, realizing that early-cognitive processing obviously takes place by means of these mechanisms already at such an early stage. In ECOVISION, particularly the results on context sensitive receptive fields of the lower visual areas are of central relevance.

5.2 Key innovations in ECOVISION

ECOVISION concentrates on a computer vision system with early cognitive properties. Summarizing the state of the art one finds that a central problem which impedes the design of such CV-artifacts with higher vertebrate (human) properties is currently the lack of a

DISTRIBUTED ARCHITECTURE	which embeds
EARLY-COGNITIVE SCENE ANALYSIS	such as the examination of the visual context by
GESTALT PRINCIPLES.	To this end
SPATIAL CONTEXT	is important to get a spatially relational perspective, while
TEMPORAL CONTEXT	allows to make predictions and to project the gained knowledge about the scene into the future. Contextual influences need to be introduced through
ADAPTIVE PROCESSES	generated by
DYNAMIC FEEDBACK.	The system will attain early cognitive properties by
RELEVANCE REPRESENTATION	which will allow to define the momentarily existing
CENTER OF INTEREST	where the most critical level of scene analysis resides In addition, such a system should still be able to operate at
HIGH FRAME RATES	by assessing the feasability to use dedicated VLSI as a front-end.

[1] PDP = parallel distributed processing

5.3 ECOVISION will lead to significant advances in the state of the art:

The project will lead to significant advances in five aspects superseding the above described state of the art:

1) *Feasibility assessment of a VLSI implementation of depth and motion extraction algorithms*. Adaptation of the front-end processing steps in order to suit them into a FPGA hardware structure. The study of how can the computational primitives that support the early stages of ECOVISON, be modified will be carried out through the cooperative work of partners with complementary backgrounds in the fields of computational vision neuroscience and neuromorphic engineering, all of them with expertise in neuronal processing schemes.
2) *First step of the final stage:* Most importantly the envisioned system will for the first time use the influence of early-cognitive context in order to control scene analysis. This will be based on the Gestalt principles known from psychology. The system itself will utilize a cellular (distributed) architecture at its front end with a topographical representation of the visual field. Integration will take place only rather late in the architecture. The "cellular" computing entities, however, will not be spiking neurons but rather receptive field filters, with a spatial and temporal response characteristic. A hierarchy of such filters will be constructed such that higher visual levels of the system will contain more complex receptive field characteristics. Contextual influences, arising from higher levels and influencing the lower levels, will be introduced via recurrent feedback loops. Two central aspects of early cognitive context will be in the focus of this project: A) Spatial context from the vicinity of the regarded location (which f.e., could lead to adaptation) and B) the temporal stimulation history which will introduce the aspect of early-cognitive predictions in the whole scene. Of importance is it here to emphasize that the system will not be a hierarchy of stacked grandmother cells of increasing complexity, instead more complex behavior is expected to arise from the co-operativity at the different network levels arising through dynamic feedback. In summary: At that stage the system will contain a set of dynamically changing receptive fields of units which represent the *locally* detected Gestalts in the *momentarily* existing context of the visual scenery.
3) *Second step of the final stage:* In the second step we will introduce an entity which we call "Relevance Map". Visual relevance is defined by the momentarily existing task and by the actual structure of the visual scene. Thus, the relevance map is a dynamic structure.

   The relevance map can be interpreted as a map of weighted Gestalts, because it rests on the Gestalt concept from which the receptive fields have been defined in the above paragraphs. Conventional systems use space-variant visual feature maps. In these maps normally the visual resolution at an artificial fovea is enhanced. In our system, the resolution will be enhanced at the Center of Interest (CoI, see next paragraph). Thus, the CoI can be likened to a *"mental fovea"* which shifts according to the demands of the scene analysis on the relevance map. Note, that we do not oppose classical active vision ideas which include eye/camera movements. Instead we wish to investigate this new avenue of a mental fovea as a complementary "cognitive" mechanism that could have far reaching prospects on the capabilities of classical active vision systems. From what was said it is clear that relevance map and CoI are fast-changing adaptive structures.
4) *Third step of the final stage:* Another innovative aspect is the introduction of the Center-of-Interest (CoI) in our visual processing scheme. In humans visual saliency is the main determining factor of attention. The CoI in our system, on the other hand, is strictly task defined and represents a spatial location on the relevance map. Spurious attentional influences cannot immediately change the CoI. However, attention is important. Thus, the CoI might change also as a consequence of deflected attention but only after scrutinizing the reason behind it. This might result in a changed task and as a consequence a change in the CoI. Thus, the stereotypical attention driven reactions are replaced by an early level of simple "reasoning" about the visual situation, which is a more common process in humans that pure reflexive behavior.
5) Furthermore, all these theoretical steps are very strongly related to important questions in computational neuroscience. We will investigate these questions (WP6) in direct conjunction with the process of solving the computer vision problems in ECOVISION in order to benefit from this cross-fertilization between the two fields.

6. Community added value and contribution to EC policies

6.1 Outline of expertise of partners

The participating research groups have a set of expertise with the potential for stimulating inter-disciplinary interactions, which can hardly be found anywhere else in Europe or elsewhere (see diagram). Thus, they complement each other optimally.

	Psychol./ Psychophys.	Exper. Neurosci.	Computat. Neurosci.	Image Processing	Neuromorph. Engineering	VLSI design	Indust. work
Eng	XXXXXXX		XXXXXXX		XXXXXXXX		
Ger	XXXXXXX		XXXXXXX				
Sco		XXXXX	XXXXXXX	XXXXXX	XXXXXXXX		XXXXX
Bel		XXXXX	XXXXXXX				
Ita			XXXXXXX		XXXXXXXX	XXXX	
Spa					XXXXXXXX	XXXX	
Ind				XXXXXX			XXXXX

Table 1: Mutual expertise diagram[2]

The central work-load which is in the field of computational neuroscience and neuromorphic engineering is supported by several groups.

The other fields have strength through support from two partners.

A direct link between industry and academia exists through Sco, who has expertise in industrial work from the development of the stereo-chip/board.

6.2 Requirement for inter-disciplinarity

The table above very clearly lays out why interdisciplinary cooperation is necessary. This cross-linked structure of different expertises cannot be found in just one country.

6.3 Potential strength of this particular collaboration

The ECOVISION working group has been discussing cooperation since October 1999 in order to pursue this particular project of Early COgnitive VISION. During this process a WEB forum had been set up (http://www.pspc.dibe.unige.it/~ecovision) where our ideas were discussed and have now crystallized into this final project layout. Therefore, despite that this is the first time that all the partners have been brought together in a single project, we are confident that the consortium will work well. Another strong reason for this is that successful bilateral and trilateral co-operations exist (have existed) between several of the partners.

6.4 EU-Policies in work-programme 2001

The commission has launched several calls with subaspects which are directed towards cognitive vision systems. Specifically: action line IV.2.1 (2001) about "Real-Time Distributed System" which contains as subaspect (ii) "robust cognitive vision systems".
Furthermore there will be the proactive initiative on “life-like perception systems” (call expected for June 2001).
The topics of the ECOVISION proposal falls also in these domains. The consortium has already started to investigate the scientific and applied problems of ECOVISION.

[2] Note: The distinction between Sco and Eng (both UK) as well as Ger and Ind (both Ger) is for technical/descriptive reasons only.

7 Contribution to Community social objectives

7.1 Competitive international position in research in this area

This proposal opens a path to a group of emerging technology concepts, which so far have been pursued almost only in the US and Japan.

Cognitive context in a computer vision system has basically only been considered in some of the most visionary computer vision systems in a few of the worlds top level laboratories like MIT or Carnegie Mellon. Normally these laboratories are leader of those scientific-technological developments, which will have an immediate impact in industrial application driven ventures only a few years later.

High level sensors and complex sensor-motor interaction schemes are currently at the fore-front of research and development in computer vision and robotics. ECOVISION contributes strongly to this line of research.

Therefore, we expect that our research will contribute to the competitiveness of Europe in a field, which is currently so strongly dominated by US and Japanese industries.

7.2 Value of early industrial inputs

The involvement of an industrial partner at this early stage provides us with the excellent opportunity to immediately adapt the complex concepts contained in the planned system to the long-term needs of industry. The onward research will have to be continued after the funding period. Thus, proven industrial involvement at this early stage will help to attract ongoing funding from both research councils and industry.

7.3. Employment prospects

The high-level technological development which will form within this company (Hella) creating a small number of jobs and growing continuously with the long-term prospects of this project. The project will create further jobs in the medium to long term within companies building systems based on ECOVISION. The wider employment implications in industry will depend on the rate of uptake of this emerging technology. ECOVISION has the potential to contribute to maintaining the competitiveness of Europe in the field of industrial robotics.

The pilot industrial application is in the automotive industry, which is one of the single largest employers in the EU and it’s Associated States. European manufacturers are in a constant battle to maintain competitiveness in this industry and even small cost gains are valuable. We hope that ECOVISION based systems would make quite significant productivity gains particularly in the assembly of complex automotive components, which existing systems cannot deal with.

The ECOVISION technology as it is conceived would be suitable for use in the building of complicated components not only in the car industry but also in other industrial sectors. So it has the potential for wide industrial application and to improve European competitiveness in each of them.

7.4 Training Impact

Within the project researchers from a very different group of disciplines will be brought together in novel and highly interesting way. This will have an impact on both the university and industrial researchers involved and will provide interdisciplinary new insights for all involved. This will be reflected in the future undergraduate and postgraduate training programs of the participating universities, in the way in which companies approach technical problem using neurosciences and psychology concepts to address engineering issues.

The interdisciplinary environment in ECOVISION will hopefully be particularly stimulating for the young researchers employed on the project who will gain experience of the inter-actions at the leading edge of each of their own disciplines and at the boundaries between them and the others. We hope that this will be particularly fruitful between Psychology/Psychophysics, Neuroscience and Image Processing and between Theoretical Physics, Neuromorphic Engineering and VLSI design. There will be other as yet unforeseen interactions, which we hope will contribute to developing their potential for either university or industrial research.

7.5 Social impact of the ECOVISION concept

An important wider social aspect of this research is that we believe that it will help to lessen the controversy in society of trying to build "humanoid machines" through open discussions (see "podiums discussion" in section 8). Public Science Awareness in the UK is an excellent example how controversial scientific concepts which might have a possible impact on society can be successfully discussed between scientists and "the rest of the world".

In section 8 we describe that such a public discussion shall be one aspect of results dissemination. The importance of resolving this controversy cannot be emphasized strongly enough. The raving discussions about genetically manipulated food demonstrates that it is too late to start such a process when the final product has reached market standards. From the viewpoint of our scientific and technologically driven consortium this seems just a sideline to our research, but indeed it carries significant impact for the European society and failure to do so might thwart the whole effort of creating truly autonomous machines for which the ECOVISION system would be a vital component.

8. Economic development and scientific and technological prospects

8.1 World Machine Vision Market - Background

The machine vision market can be segmented into two broad categories: original equipment manufacturers (OEMs) and the factory floor. The former are companies that build standard products sold as capital equipment for the factory floor. The latter are companies that have technical expertise to build programmable, board-level machine vision systems directly into their products, which are then sold to end-users. The ultimate goal of machine vision research, is to replace human vision in a variety of industrial operations as well as world-wide consumer applications/products. In fact, while machine vision systems were first embraced by industrial manufacturers for product inspection, reading (OCR, OCV or barcode), packaging, and for robot guidance in assembly processing, advances in technology and easy-to-use interfaces have made machine vision available to a broader range of users. The end-user marketplace is a potentially large market for machine vision technology.
In perspective, the development of easy-to-use products will enable to accelerate the growth of the rapidly developing end-user market, which shows great promise for machine vision. In fact, while analysts [e.g., see "The Machine Vision Market: 1998 Results and Forecasts Through 2003, Automated Imaging Association, April 1999:"] estimate the current OEM market to be worth $190 million, and the current factory floor market to be worth $280 million, many believe that the end-user potential market is in excess of $1 billion, which creates many exciting prospects for the technological exploitation of the results of ECOVISION project.

Given the current attempts of industry to create smart sensors and more flexible actuators we believe that the ECOVISION system or its subcomponents have the potential to contribute significantly to a new emerging market on "intelligent adaptive systems". The modular concept of ECOVISION might allow to uncouple subcomponents and make them marketable even before a high industrial level of applicability for the whole system is reached.

8.2 Prospects for further basic research using results from ECOVISION

The project addresses quite novel issues in the field of experimental and theoretical neuroscience as well as for computer vision. The new idea to employ a distributed architecture for early cognitive processing is not the end-point of this line of research. ECOVISION finishes at an intermediate level of cognitive complexity. Our system does not do real "reasoning". This is the central field of basic research which open beyond ECOVISION: How do "neuronal" systems (real and artificial) reason about their world and how can this be implemented in technological systems. This seems such a general question without ground. However, one of the advantages of ECOVISION is that it offers a solid entry point into this domain.

8.3 Industrial exploitation plans:
HELLA's central interest in this project rests on two legs:

1) Anticipation of a fully functional low-level, chip-based system for the evaluation of stereo- and motion data. Such a system is of major importance for several companies in the automotive supply area, because currently existing ACC system still have several significant disadvantages. The goal is to supplement these radar or lidar based ACC systems by "human-like" systems; i.e. by systems that use vision. This makes a better man-machine interfacing possible.

2) To this end fast low-level hardware is necessary to decrease the cost and make a mass-product accessible. The existing stereo-board is currently tested by HELLA, but additional motion information would significantly improve this product.

8.4 IPR planning:
The flow-field/motion chip shall be patented at least at the European level. In particular HELLA and the Spanish partner pursue IPRs on the chip design. Sharing of intellectual property and dissemination is regulated independently by the partners in this consortium.

8.5 Dissemination
In general, scientific dissemination (and exploitation) is in the fore-front of this project.

Symposium:
We are planning to hold one larger international symposium at month 36 of the project. The scientific goal of this symposium is to disseminate our own results and to be able to discuss them in a broader forum.

The symposium shall be a bigger event taking place at University of Stirling for 3 full days involving approx. 16 external speakers. In addition we will directly invite another approx. 40 scientists with similar fields of interest. Approx. 150-200 participants are to be expected in total. A podiums discussion about the problems of emulating higher level brain functions into technical systems and the impact that this will have for society shall be central to this meeting. This podiums discussion shall be open to the public and embedded in the public science awareness efforts in the UK. The press will be informed.

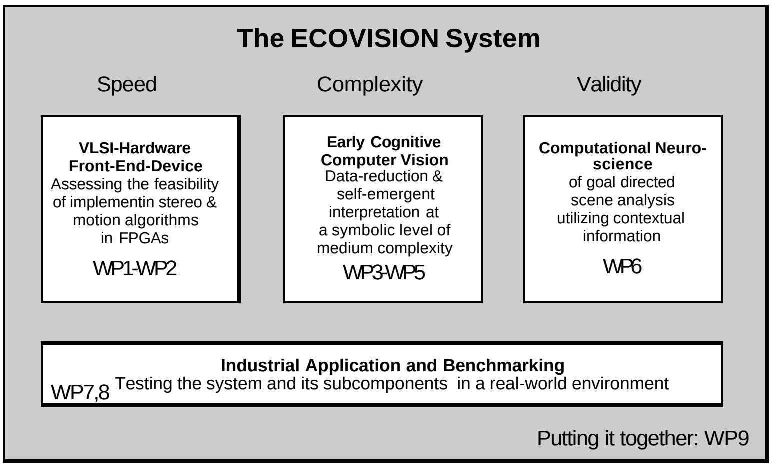
8.7 Publication planning:
We anticipate several contributions to journals and conferences.

9 Workplan

9.1. General Description

The work is divided into 5 sections: 1) VLSI feasibility assessment, 2) Early cognitive mechanisms, 3) Computational neuroscience, 4) Technical workparts and 5) Administrative workparts.

Partner Spa is mainly responsible for part 1, in order to do this Spa requires to set up several PC-based hardware systems for FPGA testing which will allow to work in parallel with 2-3 people on this project. These hardware systems are custom designed to the task of FPGA testing and cannot be easily used for other purposes, therefore Spa will set up one more general purpose PC.
Partners Ita, Ger, Eng are responsible for part 2, partner Bel for part 3, partner Sco for parts 4 and 5. We define a total of 21 deliverables (which is essential one per group per year, only). This way the work is subdivided into rather small chunks in order to assure success and a fine interweaving between the different workparts.
All workparts basically run in parallel and the structure of the deliverables is such that they will be exchanged at certain time-points.



9.2 Workpackage List

Workpackage list

Work-Package	Workpackage title	Lead Contractor	Person Months	Start month	End month	Phase	Deliv-erable
1	Motion algorithms	Sco	40+10	0	13	--	D1
2	FPGA Simulation	Spa	40+7	6	25	--	D2
3	Adaptive receptive fields as context sensitive visual filters – Gestalts in space-time	Ita	72+1	0	33	--	D3.1 D3.2 D3.3 D3.4
4	Space variant representation of weighted filters in a map	Eng	54+7.5	0	36	--	D4.1 D4.2
5	The representation of visual relevance (early cognition)	Ger	72+12.5	0	36	--	D5.1 D5.2 D5.3
6	Contextual effects in human and non human primate visual system.	Bel	35+6.5	0	36	--	D6.1 D6.2
7	Input/Output Specification	Ind	22+2	0	12	--	D7.1 D7.2
8	Benchmarking and Testing	Ind	19.8+2	18	36	--	D8
9	Joining the components	Sco	35+5	18	36	--	D9
10	Presentation, Dissemination & Use	Sco	12+3	0	36	--	D10.1 D10.2 D10.3 D10.4
11	Management & Evaluation	Sco	14+3	0	36	--	none
Coord.	Coordination (not a WP)	Sco	18+4	-	-	-	-
	TOTAL		**433.8+63.5**				

Notes: (MM = Person Months)

1) Additional cost partners have allocated the following MM for permanent staff: Sco: 16MM, Bel: 6.5MM, Eng: 3.5MM, Ger: 19.5MM and Spa: 18MM adding to a total of: 63.5 MM.
2) The "Person Month" column in the table is devided into: A+B, where A is "Person Months of personnel to be hired specifically for ECOVISION" and B is "Person Months of permanent staff".
3) In addition, there are 18MM incurring for Co-ordination activities (cost for a part time secretary) which are not associated to any workpackage (see bottom of table).

9.3 Workpackage Description

9.3.1 Project Part 1: The VLSI Front-End: VLSI implementation of motion analysis

WP1: Motion algorithms:
Lead Contractor: Spa
Partners: Eng, Sco, (all)

The goal of this WP will be to find that particular motion/flow-field algorithm which is most optimally suited for a VLSI implementation. We will not try to generate de-novo solutions to the motion/flow-field problem, because several software solutions exist which emulate possible neuronal mechanisms. Four members of the consortium have designed such algorithms which are "neuronal" (local computations) and rather well understood. First we will focus only on the algorithm of Eng because preliminary results already suggest that this algorithm is excellent for a VLSI implementation. Direct consultations between Spa and Eng will investigate this in the first 4 months of the project. Pre-specifications (and technical constraints will be defined) will be made during this time.

Step 1)

a) If Eng's algorithm passes this stage then see step 2
b) In the unlikely event that Eng's algorithms fails to pass this stage the pre-specifications will be mailed to all other partners. Partners Sco, Ita and now also Bel will *briefly* check their own algorithms against these pre-specifications, Ger, who has a broad background on flow-field analysis, will in addition check algorithms from the literature. This will be done in two weeks only and immediately afterwards a meeting will be made by all partners to discuss the outcome of this. This will lead to the decision which other algorithm can be taken instead of Eng's.

Step 2)

The chosen algorithm will be adapted to a low level software simulation in order to assess the required accuracy levels (bit cutting procedure) and to be optimally mapped onto an FPGA structure.

Step 3)

A quantification of the accuracy of the different cutting depth allowed by the algorithm will be investigated in cooperation with Ind who will provide camera data of driving scenes for benchmarking.

Steps 2 and 3 will take place in the remaining 12-15 month of this WP.

D1: "Best Motion Algorithm": This deliverable will be a low level integer software version of the *best adapted* motion/flow-field algorithm with a well defined bit-depth for every operation. (Deliverable Format: Computer program and short Technical Report/Documentation)

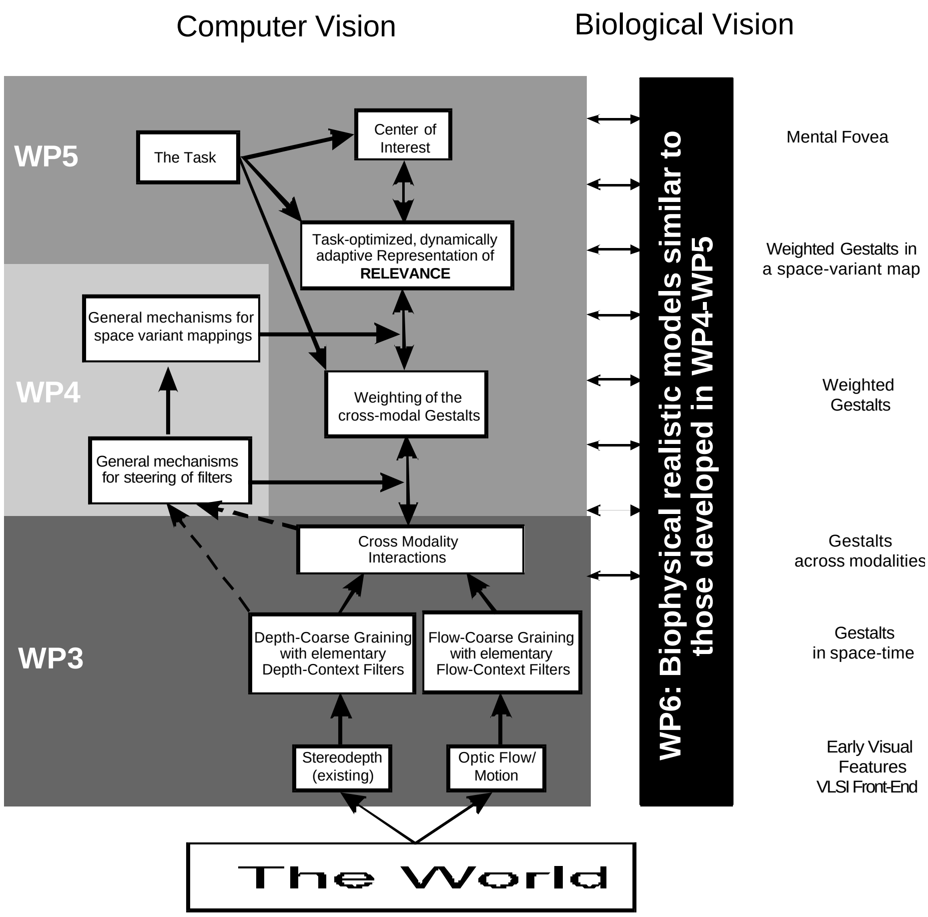
WP2: FPGA simulation
Lead Contractor: Spa
Partners: Ita, Spa

Spa will use the results from the analysis of motion/flow-field algorithms (D1) in order to define concrete technical specifications (bit cutting procedure and others). These technical parameters are essential to face the VHDL programming stage of the algorithm. The simulation of the FPGA functionality in VHDL programming language will provide information about the implementation and computation costs of different parts of the algorithm. Ita's expertise in neuromorphic engineering will contribute to this. This WP will overlap with WP1 because some technical constraints arose in this stage will have to be taken under consideration in terms of its FPGA feasibility. Finally, different possible projections and their simulations will be carried out in order to optimize the FPGA functionality and to investigate possible modifications of the algorithm aroused from the FPGA architecture in which it is intended to be projected will also be taken into consideration, evaluating the speeding up and the time costs of different parts of the algorithm. Therefore the whole process of WP2 will have the following scheme:

1. Define technical specifications, that may be crucial for the algorithm operability, to be taken into account in the first VHDL approach.
2. Simulate the FPGA compatible version of the algorithm in order to allocate possible critical steps and evaluate their relevance in terms of silicon costs and working speed.
3. Study the different possible projections on an FPGA board and their implications in the final efficiency.

D2 "Fully Functional VHDL Simulation": This deliverable will be a fully functional VHDL simulation of the selected algorithm.

MILESTONE 1: The feasibility assessment of a motion real-time processing unit to be used as a front-end of the system together with the existing stereo-unit.



9.3.2 Project Part 2 Investigations of early cognitive mechanisms and implementation in a distributed system

WP3 Adaptive receptive fields as context sensitive visual filters – Gestalts in space-time
Lead Contractor: Ita
Partners: Sco,Bel,Eng,Ger,Ita, Spa

In WP1-2 the output of stereo and motion/flow-field analysis yields so called feature maps. These feature maps are always rather erroneous and/or ambiguous. However, the structure of the world provides intrinsic statistic properties within and across both maps that allow to define Gestalts.
The goal of WP3 is:
(1) to determine the structure of intrinsic correlations for localized pixel groups (*spatial* Gestalts);
(2) to determine their *temporal* interactions (*spatio-temporal* Gestalts);
(3) to investigate how these Gestalts change over time and recursively interact with each other across (stereo and motion) modalities to obtain *adaptive cross-modal context-sensitive filters*
(4) to exploit the outputs of such filters to obtain more reliable feature maps to be used in further processing

This will be done in four steps.

WP3.1) *Visual templates with spatial context - Gestalts in Space:*

Specific work-steps to generate spatial Gestalt filters are:
1) We (lead by Ita who has started attempts into that line of research) will develop or identify in the literature algorithms to individuate pixel groups with
   A) a *shared* motion/luminance pattern and
   B) those with a *shared* disparity/luminance pattern.

In order to define what *shared* means we will

2) assume and investigate several simple null-hypotheses about the possible geometrical transformations (and their combinations) which could underlie the feature-structure of such a pixel-group.
   A) For example at a given moment in time the motion vectors of a localized pixel group could have derived from rotation, translation, shear, etc. transformations and
   B) similarly for stereo by constant or linearly changing disparities (more complex changes shall be approximated by step-wise linear changes).

Similar to the visual system of primates, where – for example – rotation sensitive cells have been found in area MST, we will, thus, in this first step define *receptive fields of units (templates)* which are sensitive to these particular flow or disparity patterns. Finally we will:

3) Test these receptive fields against real (driving) scenes.

   For every analyzed point on the image these receptive fields can be understood as first order *spatial* context filters This results in a context-dependent spatially coarse-grained description of the image structure through wide-field depth and motion sensitive receptive fields that capture coherent properties of large (overlapping) pixel groups.

The context filters, thought as deformable templates in a high dimensional domain, will be defined in a topological space to allow a direct extension of the space-variant mapping strategies investigated in WP4.

D3.1 "Visual templates with Spatial Context - Gestalts in Space": This deliverable will be several sets of receptive field filters which are sensitive to either stereo- or motion and change their filter characteristic with respect to the spatial context (Deliverable Format: Mathematical (filter-) equations stage 1 in a technical report and computer implementation thereof).

WP3.2) *Visual templates with spatio-temporal context - Gestalts in space-time:*
The concept of Gestalts described there will now be extended into the temporal domain.
Specifically, we will:

1) Track the development of these receptive fields (Gestalts) over time in order to find a mathematical (algorithmical) way to describe these temporal developments.
   A) Here we will mainly rely on the paradigm of coherent motion.
   B) In addition, we will utilize the prediction of motion parameters from optic flow and rigid body motion models (RBM), which is basically a solved problem, but not in the context of receptive fields, which will require to reformulate the know mathematical relations in this context.
   C) In a very similar way, we can also rely upon spatio-temporal constraints imposed on the *deformation* of retinal image, during rigid relative motion between the observer and the environment (RBM), to evidence spatio-temporal correlations in motion and depth parameters

2) Then we will use these algorithmical descriptions to alter the Gestalts (the receptive fields) from time step to time step and thereby adapt them to the changing structure of the visual scene. Continuous time adaptation is one of the resources most widely used in biological systems to overcome hard constraints imposed by the individual cells. This approach is also related to the idea of adaptive filters in electronics only that we will re-define it in the context of visual receptive fields.

Accordingly we have now arrived at Gestalts in space-time.

D3.2) "Visual templates with Spatio-temporal Context - Gestalts in Space-time": This deliverable will be several sets of receptive field filters which are sensitive to either stereo- or motion and change their filter characteristic with respect to the spatio-temporal context. (Deliverable Format: Mathematical (filter-) equations stage 2 in a technical report and computer implementation thereof).

WP3.3) *Cross-modal interactions between the Gestalts define the first level of recursive scene analysis:*
Here we define cross-modality interactions between stereo and motion and this is the first step towards a recursive way of scene analysis within ECOVISION. So far the receptive fields are restricted to either

motion or disparity analysis. Chaired by Sco in this step receptive fields shall be defined which integrate stereo and motion. This requires the identification of the underlying processes by which stereo and motion information are related. Two mechanisms can be immediately used to define the first work-steps:

1) Static inference: The structure of one set of receptive fields allows to immediately infer parts of the structure of the other receptive fields. For example, the stereo Gestalt-RF-map allows to make direct predictions of the relative motion/flow field vectors. Specifically we will algorithmically implement the following properties:
   A) objects close by will lead to larger flow field vectors than those far away.
   B) Objects close to each other in depth will have similar flow vectors.
   C) Tilted objects at a given location in the map will produce only a distinct set of possible flow-field patterns.

   And similarly for stereo:
   D) Objects at a flow-field vortex will turn such that there stereo-map changes in a predictable way
   E) Objects at the focus of expansion will gradually increase their disparity

More such static inference relations can be defined but shall not be mentioned here for reasons of brevity. Given the state of the art implementation of static inference principles is straight-forward.

2) Dynamic recursion**:** This is a more complex algorithmical principle and can be best understood directly from an example: Disregarding a lot of the actually existing complexity, let us consider the behavior of a certain group of joined pixels which have at one point in time all the same stereoscopic disparity[3]. The system will track the behavior of these joined&equal disparity pixels (and all the other pixels, too). If those pixels move in a joined (viz. correlated) way, predictable from the motion map, then their disparity should also change jointly. After such a joint disparity change, however, one will now be able to also predict with some certainty a joint motion-vector change for all these pixels, and so on. Thus, the intrinsic correlations between those pixels define a recursive process between the changes in disparity and the changes in their motion vectors. Accordingly we need to
   A) first rigorously analyze the statistical properties of large image sequence data sets in order to measure the structure of the cross-correlations between motion and stereo.
   B) Derive possible dynamic cross-modality interaction from this analysis.
   C) Implement these dynamic cross-modality interactions initially in artificial scenes with known ground-truth.
   D) Assess and quantify the robustness of these interaction mechanisms.
   E) And finally implement these mechanisms for the analysis of real scenes in ECOVISION.

Note, dynamic recursion[4] defines a process which rests entirely on a few basic assumptions about the structure of the world (like rigidity, etc.) and on the inspection if pixel groups obey the Gestalt *laws*. By this we arrive at an *self-emergent* image segmentation at the Gestalt level. We believe that dynamic recursion is one of the challenging new ideas of ECOVISION and here we enter uncharted terrain.

D3.3 "Cross-modal interactions between the Gestalts define the first level of recursive scene analysis": This deliverable will be a set of adaptive receptive field filters which are sensitive to stereo *and* motion and change their response characteristics with respect to the spatial *and* temporal context. (Deliverable Format: Mathematical (filter-) equations stage 3 in a technical report and computer implementation thereof).

---

[3] Naively it could be assumed that these pixels belong to the same object. However, we strongly emphasize that our approach does not make such a priori assumptions, which would essentially reduce the approach to a data-driven image segmentation step.

[4] At this point we extrapolate from the existing knowledge in neurophysiology. It is known that receptive fields change their shape in a time-dependent manner. While it would be an over-interpretation to directly associate these change with higher level image processing, it is, on the other hand, attractive to utilize a mechanism similar to a temporally changing receptive field structure to encode higher order spatio-temporal correlations between image parts. In general we will seek the solution to this particular distributed-data correlation-problem by employing mutual feedback between the stereo- and the motion-specific receptive fields defined in step 1. This strategy finds in justification again in neurophysiology, where results demonstrate that the generation of higher order receptive field structures rests heavily on the employment of feedback loops.

WP3.4) ***Higher level visual segments:***

We will focus on how the results obtained from context filtering operations can be used to obtain more reliable feature maps to be used in further processing. Here a specific interaction is needed with WP4 and WP5.
The spatio-temporal Gestalts defined in the parametric space and the resulting "Gestalt segmentation" back in the image domain, will lead to more reliable stereo and motion feature maps to be used as input for WP5. Specifically, Gestalts will provide labels on the image by grouping pixel that share the same stereo-motion properties. Such information will be straightforwardly used as confidence measure of the original disparity and motion maps and also organized in a proper vector data structure representing complex segments, such as surface curvature and so on, towards a higher-level description of the dynamic scene.

D3.4 "Context-based confidence maps": This deliverable will be a set of confidence maps to make the original disparity and motion maps more reliable (Deliverable Format: Data Sheet Confidence measures and Data Format definition).

WP4: Space variant representation of weighted filters in a map
Lead Contractor: Eng
Partners: Sco, Eng, Ger

In WP3 we have arrived at *adaptive receptive fields that represent Gestalts in space-time across modalities*. In WP4.1 we define *weighting* ("steering") mechanisms to produce filters of different scales and orientations in computationally efficient ways. In WP4.2 techniques shall be developed to embed filters into a space-variant map.
The two deliverables (D4.1 and D4.2) focus on the technical aspects of how to construct adaptive scale space-variant mappings given the constraints of available technology.

WP4.1) ***Adaptive receptive fields as steerable-scalable filters:***

The goal of this work package is to design algorithms, which will finally allow for the adaptive behaviour of Gestalts (in WP5). To this end the following steps are necessary:

1) To make the problem tractable we will initially disregard the specific computational structure of Gestalts outlined in WP3 and concentrate on space variant image filtering with Gaussian derivatives.
2) Algorithms will be developed to dynamically adapt the scale and preferred orientation of a such a filter at a given location by weighting the filter outputs such that the final calculation is approximately equivalent to explicitly calculating Gaussian derivatives at a different scale. This is achieved by "steering" and "scaling" techniques, known from the literature, that can be adapted for use here. The aim is to calculate the value of a Gaussian derivative operator for any direction or scale by linear addition of the outputs of a small basis set of filters (tuned to a limited set of orientations and scales) with appropriate weights. This can be thought of as an efficient way to produce orientation selective fields. Given a fixed number of filters, which span all directions about a point of application, we can adaptively manipulate orientation selective filters, as well as the spatial scale of operation, to adjust the properties of the filter to the input.
3) The procedure developed here will then be used to dynamically adapt the local scale of the filtering operations to the demands of the task (for "task" see WP5) without the explicit implementation of space variant filtering at the initial stage leading to a significant reduction in computational load.

These steps will be performed in a way that will allow adaptation of Gestalts in WP5.

D4.1 "Adaptive Receptive Fields as Steerable-Scalable Filters": This deliverable shall be a software implementation of an algorithm to steer and scale filter kernels in order to obtain different resolutions with minimal computational cost at the initial filtering stage (Deliverable Format: Computer program and documentation).

WP4.2) ***General space variant mapping methods:***
Our goal is to derive mapping techniques which can be generalized to Gestalt filters and this requires the following work-steps:

1) Implementing convolutions using space-variant filters is a very computationally expensive process since filter sizes and shapes need to be specified for each location in the image. Thus, we will, establish a local Taylor series representation of the image by filtering with a range of Gaussian derivative operators at each location. This will allow us
   A) to substitute the convolutions with a simple global linear filtering operation and thereby reduce the computational load tremendously. Only by this, an extensive investigation of different mapping strategies becomes possible. The global filter sheet can have differential operators of different scales at different spatial locations.
   B) Computation of results at intermediate locations (between filter centers) can be achieved using the approximation and interpolation properties of the Taylor representation. This way adaptable scales and, thus, different space-variant mapping structures can be obtained.

2) Different mapping strategies can now be tested against real image sequences.
   A) conformal mappings like the complex logarithmic mappings which are invariant with respect to scaling and rotation
   B) more general (non-conformal) polar-mapping procedures combined which affine transformations of the map which allow to locally enhance/decrease the resolution in specific map regions (like at the CoI, see below).
   C) specific mappings which allow to transform the map-representation of flow-patterns into that of stereo-patterns.

Again we note as an important point that all these techniques need to generalize to Gestalts.

D4.2 "Space Variant Mapping Methods": This deliverable shall be a software implementation of an algorithm for optimal efficient implementation of space-variant filtering in a generalizable way (Deliverable Format: Computer program and documentation).

MILESTONE M2: This milestone covers the "sub-cognitive" side of the system. The final outcome of this research will provide a visual scene analysis system which captures cross-modal interactions between motion and stereo filters and which incorporates contextual (spatial and temporal) influences thereby defining Gestalt-sensitive receptive fields.

WP5: The representation of visual relevance (early cognition)
Lead Contractor: Ger
Partners: Sco, Bel, Ger

Summary:
WP5.1 will analyze (psychophysically): What is "visually relevant" during driving?
We know that visual relevance directly relates to the actual, momentarily performed *task* (e.g., heading, fixating, tracking, etc.). But how?

WP5.2 will bring the results from WP5.1 and WP5 together by using the psychophysical results to actually set the weights (from WP5.1) *within* the map (from WP4) and thereby define the early cognitive aspect of "visual relevance" (the *actual Gestalt-weights*) in an algorithmical way for ECOVISION. By this it will become possible to define a *Center of Interest* (CoI) within the Gestalt-map of the ECOVISION system. The CoI is a *mental fovea* which adapts in size and location according to the task.

WP5.3 Is the implementation of early cognitive properties in a robot environment

WP5.1) ***Psychophysical investigations of driving behavior and CoI in humans:***

Psychophysical experiments by will provide deeper insight into the viewing behavior of drivers. To this end the pre-recorded driving scenes will be analyzed using an eye-tracking system. These experiments rest on the existing data from studies of eye movements in car drivers. For some situations, such as negotiating a curve on intersections, it is known that drivers perform characteristic eye movements that are likely intended to aid the momentary task. This behavior is quite reliable and predictable, and, because it is very dependent upon the task, it seems to be a good indicator for the CoI. We will use a similar strategy to learn more about the CoI in driving related tasks.

To this end experiments will be made:

1) to find out which of the actually segmented image parts will attract the viewers interest and
2) how his temporal viewing behavior can be parameterized.

In particular, we are interested in possible conflicts of interest. For example it is known that highly salient image features (f.e., red-colored objects) can very strongly divert the attention of a person (and in particular of a traditional saliency driven active vision system). However, despite of this, a colorful advertisement sign should not be in the CoI during driving! Little is known about eye-movement control – which is currently the best measurable for visual interest – during such conflicts of interest/attention. Results along these lines will, thus, be potentially influential beyond the core-questions of ECOVISION.

The experiments will record eye movements of human subjects that watch a driving simulation on a computer screen. Direction of gaze will be recorded and analyzed in relation to the momentary flow field and the driving parameters. Different tasks will be given to the subjects in order to study the relationship of eye movements to the task. Possible tasks include the localization of the direction of heading, the localization of obstacles along the path, or the steering of a given course.

These experiments will be performed with healthy human subjects in accordance to the guidelines laid out in the declaration of Helsinki (see http://www.wma.net/e/policy/17-c_e.html).

D5.1 "Psychophysical Investigations of Gaze and CoI during Simulated Driving in Humans" This deliverable will be a quantitative description of the eye movements during simulated driving tasks in humans (Deliverable Format: Scientific Report)

WP5.2) ***Task optimized representations and the CoI:***

A computational procedure to take advantage of space-variant mapping and steerable filter techniques (WP4.1 and WP4.2) to the specification of context Gestalt-filters will be developed. Choice of the actual Gestalt-steering process (*how to actually set the weights*) will be driven by the task and the knowledge about the kind of optic flow/stereo information relevant to it (WP5.1); e.g., heading vs. time to contact, etc. **Task** is a short-living goal of the scene analysis process and is not to be confused with the general long-term goal of (in this case) a driving situation.

We will

1) investigate what representations are specifically beneficial for a specific task. To this end
   A) we will, first, define different *specific (concrete) tasks* from the psychophysical analysis of WP5.1. These tasks will be named in a figurative way (like "heading", "fixating", "tracking", estimating the "time to contact", etc.). However, they will indeed be formalized in a mathematical framework of how to treat a scene during such a task.
   B) Next we will formalize the demands of "how and where" the weights need to be increased/decreased for a given task. For instance, for the task of "heading", the global structure of the representation needs to have large weights for those filters with a preference for low speed near the FoE and smaller weights else. For the task of "fixating", the weights need to be larger around the fixation point which is normally different from the FoE ("Focus of Expansion").
   C) In the next step we will improve these formalizations by utilizing the known (from WP3.3) cross-modal (stereo/motion) interactions. For example, in part B we left it unspecified which filters at the fixation point shall receive the large weights. This depends on the distance to the fixation point. If the fixation point is close by then filters which prefer high velocities need the largest weights and vice versa. Thus, the stereo depth map indicates which parts of the flow will be alike in terms of

speed and direction, thereby providing a cross-modal information to choose an optimally adapted representation.

2) The next major work-step is to assure that the parameters of the representation are adaptable. To this end algorithmic extensions need to be found which are compatible with work-step 1 and which allow to:
   A) flexibly alter the shape of the representation as soon as the task changes. This alteration must be fast, i.e. following the time course of the task change, but it must also keep the already present information continuously available. We will therefore specify algorithms that ensure a gradual, smooth transition of the representation rather than an abrupt change. The course of change will hence depend on the new as well as the current form of the representation.
   B) Specifically we will test mechanisms of *lateral inhibition* which spreads from the Center of Interest (CoI) to the periphery.
   C) *Mutual inhibition* between competing possible CoIs.
   D) *Inhibition of return* which prevents to immediately switch back to an older CoI if a new one has taken over.

These mechanisms come from the theory of attention control and should with ease be adaptable to the concepts of an early cognitive representation such as in ECOVISION.

3) In the next step the such weighted Gestalts will be arranged in the space-variant mapping structure. This step is rather simple because it is identical to WP4, only now, we have task relevant weights, whereas in WP4 weights were still arbitrary.

4) Finally we will
   A) Scrutinize the employed mechanisms from 1 and 2 if they are robust against noise, and
   B) Test their results against artificial scenes with known ground truth and also against real driving scene in order to find out to what degree the obtained representation are correct and in how far they correspond to the actually existing (psychophysically measured, WP5.1) visual relevance.

D5.2 “Task Optimized Representations and the CoI”: This deliverable will be the design of algorithmic strategies to determine the actual weights of the Gestalt in relation to a given task (Deliverable Format: Technical Report)

<u>WP5.3)</u> ***<u>Implementing early cogitive control in a technical environment:</u>***

Recent results of Sco make it possible to attempt to implement the results from WP3-4 and WP5.1-5.2 in a robot environment. During the time this proposal was written and evaluated Sco has developed a small robot which can integrate different sensor inputs and is controlled by a PC.

ECOVISION faces the challenge that it shall operate on real image sequences taken from a driving car. This is a hard problem. For this reason the project will start first with artificial images (as laid out above). The newly developed robot now offers the unique opportunity to test the ECOVISION algorithms in a *controlled* real-world environment, using real sensor data but within a robot lab (where, for example, the lighting conditions and the scenery can be well controlled and measured) before testing it with complex real-world scenes.

This workpackage is introduced, because during the last months Sco was able to establish the necessary robot environment which can serve as a testbed allowing for a straightforward implementation. In view of the complexity of the complete ECOVISION system, we regard it is as a very sensible step to test the results of WP3-5.2 in such an environment, which is still better controlled than the real-world.

Steps to undertake are:
1) Record the vision sensor output from the robot and use it as input to the ECOVISION algorithms from WP3-5.

2) Use the output of the ECOVISION algorithms to control robot motion. The main goal here is to learn to avoid obstacles by means of ECOVISION based scene analysis.
3) Analyze the robot behavior in order to adapt and change the ECOVISON algorithms.
4) Use the results from this analysis to refine the ECOVISON algorithms from WP3-5 in order to make them better adapted to true real-world problems.

The goal is to arrive at an implementation which makes use of the above results and thereby reaches some complexity in its behavior. We expect, that the properties of WP3-5.2 will lead to a certain predictive power in the behavior of this machine.

This provides us with an excellent additional test- and control situation before the complete ECOVISION system shall be put together in WP9.

D5.3 "Implementing early cogitive control in a technical environment" : This deliverable will be a robot implementation of early cognitive properties with predictive power as to scene analysis. (Deliverable Format: Demo-Robot and Scientific Report)

MILESTONE M3: WP5 covers the "early-cognitive" side of the system. Visual relevance and the Center of Interest are defined here and the ECOVISION system will thereby attain the first levels of a "mindful" scene analysis.

9.3.3 Project Part 3: Computational Neuroscience

WP6 Contextual effects in human and non human primate visual system.
Lead Contractor: Bel
Partners: Sco, Bel, Ita

Summary:
In WP6.1 we will develop an algorithmic procedure derived from computational neuroscience models for specifying the weights in a space-variant map of neurons, given a spatial attention task. We will use the experimental evidence as a guidance to put constraints on the nature of the algorithmic procedure.

Then, in WP6.2, we will extend the previous algorithmic procedure so as to include tasks that require the weights to depend on contextual influences other than spatial ones. Together with WP6.1, this will lead to an algorithmic procedure for generating adaptive representations of relevance by enhanced/decreased cell activity, which is probed psychophysically in WP5.1.

Strategies of how to introduce contextual and attentional feedback into our map of Gestalts, in order to arrive at a relevance map, will be mainly derived from (computational) neuroscience studies. Indeed, there is ample experimental evidence[5] for the effect of *cognitive context* and *spatial attention* on visual processing.

[5] In the laboratory of Bel, experiments with the double label 2-Desoxyglucose labeling technique have shown that, when a monkey is using a central grating task, compared to one in which the grating is irrelevant, the representation in the lateral geniculate nucleus (LGN) and primary visual cortex (area V1) is surrounded by suppressed metabolic activity. In a further series of experiments, in which two peripheral gratings were used, one in the left and another in the right visual field, the monkey had to identify the orientation (at threshold level) of the left grating in one condition (labeled with C14) and that of the right grating in a second condition (labeled with H3). The results show that the suppression is deepened in area V1 and that the activation by the grating in extrastriate cortical areas is increased in the hemisphere contralateral to the side attended [unpublished results]. These experiments are indicative of a two-step mechanism: 1) a selection/filtering component of spatial attention planning, and 2) a re-allocation of resources, which are by default biased towards the central visual field, as part of contextual processing. The selection/filtering is also supported by the findings of others who recorded from macaque visual cortical area MT (middle temporal), the neurons of which are known to be sensitive to both motion and stereo. They found that spatial attention yields an enhanced response to stimuli at an attended location in the visual field and a reduced response to stimuli elsewhere. This was also observed in the efferent area MST (medial superior temporal). Both these and the 2DG findings place important constraints on the neural mechanisms of attention. They can be shown to be compatible with the properties of the surround that envelopes the MT neuron's receptive field, and that plays a crucial role in the segmentation of moving objects from motion information.

The biological data for the development of these theories will be taken from Prof. Guy Orban, who will co-operate in this project (see Appendix A).

WP6.1) *Task-optimized spatial representations – spatial attention case:*
We define a space-variant map of neurons (e.g. by the standard complex-logarithmical mapping) and investigate how to assign weights to each cell's activity, given a task that involves spatial attention.

1) We first model algorithmically the cortical suppression observed in the 2DG and electrophysiological experiments (see footnote 5). The obvious way to do this is by introducing a center/surround type of connection kernel in the map.
2) Next, in order to make the suppression task-dependent, we equip the neuronal representation with fuzzy membership functions of which the degree of fuzzification depends on the presence of (nearby) spatial clusters of similar activities (fuzzy membership in activity contrary to fuzzy membership in clusters). The idea is to allow, within a task-dependent spatial range, any combination of neuronal activities. In the context of WP3-5 these group-activities can be associated with the filter-like Gestalts (at a distributed multi-cellular level).
3) The *relevance* of these group-activities is derived. By the type of membership function used, continuous graded representations are created, which are indicative of their degree of relevance for the task.

D6.1 A report on the "CNS[6]-algorithm for generating task-optimized spatial representations of neural activities" This deliverable is needed for generating general task-optimized representations in WP6.2. It takes the available 2DG and electrophysiological data as a guide for constraining the design of the algorithm (Deliverable Format: Scientific Report).

WP6.2) *Task-optimized representations - general case:*
We extend WP6.1 by allowing for contextual influences other than spatial ones to specify the weights that are assigned to the cell activities in the space-variant map. We will address this issue from an information-theoretic standpoint.

1) The weights are assigned in basically the same way as in WP6.1 but, in order to keep the problem manageable, our starting point will be that only specific combinations of group-activities (Gestalts) need to be further analyzed (to save resources in the current step). These combinations will be found by an unsupervised, information-theoretic learning procedure, called *infomax* learning, which will detect co-occurring sets of group-activities. More specifically, we intend to represent the "Gestalt" combinations as points in a data space and derive the (not necessarily orthogonal) axes so as to make the representations projected along these axes statistically independent (namely derive the Independent Components, cf. ICA). This is akin to a data compression procedure.
2) The next step is to address the fact that, especially when using experimental data, which is the intention of this project, the combinations of group-activities ("Gestalt-combinations") are likely to be noisy and non-linearly transformed. Hence, a more robust algorithmic procedure than infomax will be developed. Most likely we need to take into account the skewness as well as the kurtosis of the group-activity distributions.
3) Once this robust ICA procedure is developed, a binding process (e.g., neuronal synchronization) will be defined to bind all those different group-activities which belong to the same task. This way we arrive at "labeled" activity patterns, which represent a meaning (namely the momentarily existing task). The desired binding can be easily established by selecting the most likely co-occurrences between task and combinations of group-activities ("Gestalt-combinations").

D6.2 "CNS-solution for relating tasks to combinations of group-activities". This deliverable consists of: 1) a report on the robust ICA-based algorithm, for finding the combinations of group-activities ("Gestalt-combinations") that belong to a given task, and 2) a report which describes the strategy for binding Gestalt combinations to tasks (Deliverable Format: Scientific Report).

[6] CNS=Computational Neuroscience

9.3.4 Project Part 4: Technical Workparts

WP7 Input/Output Specifications
Lead Contractor: Ind
Partners: Spa, Ind

Ind will provide camera sequences from different driving scenes. These will be used to test the performance of the motion/flow-field chip and of the (existing) stereo-board. At the same time these sequences will be used to test the results of Project Part 2 (WP3-WP5). System simulations in laboratory are suitable methods to realize first efficiency tests of the FPGA based motion/flow-field algorithm and the complete ECOVISION system, but only real-world data of driving situations reveal the performance of the chip in an embedded car application. For this, numerous video sequences are required in the chip-performance test phase. Due to the fact that ground truth video data can vary in a unforeseeable manner a systematic categorization of driving situations is evidently needed, before. Video sequences will be taken from a forward looking camera mounted in the cars roof module similar to the planned position of future embedded automotive vision systems. The camera is part of a stereo setting. Some other car data like ego-velocity, yaw rate, e.g. should be recorded synchronously.

D7 "Input/Output Specifications": This deliverable will be
D7.1 Well adjusted driving sequences (Movies) with defined scene parameters (Deliverable Format: Movie files)
D7.2 Set of agreed-on I/O parameters (Deliverable Format: Data sheet).

WP8 Benchmarking and Testing
Lead Contractor: Ind
Partners: Spa, Ind (and to a small degree most other partners)

Ind will perform a quality control particularly for the low-level chip-based part. Ind will provide expert help joining the components in WP8 and testing them. This WP is essential for the complete project, because only through industrial involvement we will be able to test ECOVISION and its sub-components against hard and concrete demands.
This task demands investigations in the generation of quality and robustness describing determination parameters. Ind will also insert the experience concerning data-fusion framework (stereo + motion) in an automotive environment and strategies in case of sensor /chip signals with a poor reliability. Experiments with images taken on public roads will be done to test especially the FPGA efficiency. The wide spectrum of real-world data leads directly to a categorisation of test-driving scenes.
General driving areas:
1.) motorways 2.) federal roads 3.) country roads 4.) city streets
Area characteristics and geometries:
1.) horizontal curves 2.) vertical curves (e.g. road inclines)
Illumination and weather conditions:
1.) day/night 3.) rain/snow 4.) low contrast (fog) 5.) high contrast (glaring)
different moving objects:
1.) pedestrians 2.) motorbikes 4.) cars 5.) trucks (lateral and longitudinal movements)
Ego-motion:
1) low-speed/high speed 2.) pitch-, roll and yaw-angle movements of the car

D8 "Benchmarking and testing": This deliverable will present the rationale that guides the definition/choice of the benchmarks (Deliverable Format: Technical evaluation Report) .

WP9 Complete ECOVISION System (Joining the components)
Lead Contractor: Sco
Partners: Sco, Spa, Ind, (all)

WP1 to WP2 build on top of each other and similarly WP3-WP5. Sco will chair the development of the complete system (Sco owns the already existing stereo FPGA-board). Provided the 2 sub-components from WP1-2 and WP3-5 are functional, this is a straight-forward rather technical step and requires mainly intelligent handling of the input-output interfacing between the two components and the adaptation of the

processing speed of component 1 and of component 2 (expected to be somewhat slower). Spa is needed because they have developed the motion FPGA implementable algorithm.
The system integration of workparts WP3-WP5 will demonstrate the feasibility of the Gestalt approach. Joining the components into the complete ECOVISION system is mainly a software problem, because motor control arises within a simulated environment. This limits the total effort of the WP to 31 MM, with real sensor-motor units it would be much higher.

D9 "Complete ECOVISION System": This deliverable shall be the complete system. It comprises the last Milestone (Deliverable Format: Lab-Model Demonstrator)

MILESTONE M4 is the complete ECOVISION system

9.3.5 Project Part 5: Administative Workparts

WP10 Presention, Dissemination and Use
Lead Contractor: Sco
Partners: all

This workpackage will be concerned with the presentation of the ECOVISION results to the "outside world".

We will:
1) Maintain and up-date the already existing Web page with
    a) Non-confidential Reports
    b) Images and Movies of the analysis results
    c) Sets of computer sub-routines (functions) as a computer vision library of Gestalt-filters
    d) Public meeting announcements
    e) Miscellaneous

2) Publish in international scientific journals.

3) Make one international meeting at the end of the project with external speakers. We are planning to hold a larger meeting with a substantial number of external speakers (also from overseas). The meeting shall have about 100 participants in total. Invited speakers will have their accommodation and travel cost refunded, in addition we need to rent lecture halls from the university. This leads to a total cost of about 30,000 Euro as specified in the CPFs.

4) The aspect of "use" will be addressed by an internal screening done by Ind where the achievement will be juxtaposed with the actual demands from HELLA in the context of car cruise control.

Deliverables: Dissemination and Use according to
D10.1 Web page,
D10.2 Dissemination & Use Plan,
D10.3 Scientific Publications,
D10.4 Workshop at Stirling.

WP11 Management and Evaluation
Lead Contractor: Sco
Partners: all

General management structure: The Project Management Board (PMB) is formed by the Project Manager (F. Wörgötter) supplied by the Prime Contractor (Sco) and by the Workpackage Coordinators as listed below:

WP1: Spa (E. Ros)
WP2: Spa (E. Ros)
WP3: Ita (S.P. Sabatini)
WP4: Eng (A. Johnson)
WP5: Ger (M. Lappe)
WP6: Bel (M. Van Hulle)
WP7: Ind (M. Mühlenberg)
WP8: Ind (M. Mühlenberg)
WP9: Sco (F. Wörgötter)
WP10: Sco (F. Wörgötter)
WP11: Sco (F. Wörgötter)

The PMB will report to the project manager and will resolve possible conflicts between partners. The project manager will be responsible for all contacts with the European Commission and other projects. and for re-planning and rescheduling of resources, objectives and deliverables in order to maintain the project within the original proposed finished date, whilst achieving the objectives as originally declared, based on the feedback obtained from the PMB members responsible for each work-package.

Decision making procedure: The partners in ECOVISION have built up a good relationship while preparing this project and it is anticipated that decisions will normally be by consensus, Should a vote be necessary each member of the PMG will have vote and the Coordinator will have an additional casting vote if required.

Communication structure: At regular intervals (every two months), the workpackage coordinators will be required to communicate briefly by email to the coordinating partner all necessary information relative to the progress of their work-package, each task within that workpackage, resource changes, cost changes, changes in critical activities, and forecasts with regard to the attainment of objectives. On arrival of this information, it will immediately be integrated into the master plan so that the project manager can continuously monitor all project activities. We will use the already existing WEB Forum (password secure members section) at http://www.pspc.dibe.unige.it/~ecovision to facilitate communication between the partners.

Meetings: Regular meetings [as required but at least every six months] will be held between the project manager and the PMB members responsible for each work-package any meetings may be bilateral and specific to the context of the specific workpackage
The PMB will meet once per year. The aims of these meetings will be to resolve any plan or schedule conflicts that may have emerged from the monthly feedback reports and to make technical decisions based on results achieved and to allocate responsibilities for the following periodic progress report when relevant.
Meetings will also be held on a work-package level between partners involved in that work-package where the project manager will also be present on request of the PMB member responsible.
The frequency of meetings for each work-package will be decided between the PMB member and the project manager depending on the nature and number of partners in the workpackage. In particular, additional meetings are necessary between Ita and Spa for the coordination of WP1-WP2 as well as between Sco, Bel, Eng, Ger and Ita for the coordination of WP3-WP5.

Reports: Internal progress reports from the different WP-coordinators are scheduled every 6 months. Formally, progress of the work will be assessed by means of twelve-month Progress Reports prepared by the Project Management Board for the Commission on the basis of the internal reports. These 12-month Progress Reports will contain a Management Section. They will be sent together with cost statements. Given the explorative character of this project it does not seem to be reasonable to send reports more often. A Final Project Report will be prepared at the end of the project.

Progress Monitoring: The progress of the project will be monitored in relation to the Scientific and technical deliverables as described in Section 9.5.

9.4 Deliverables List

Deliverables list

Deliver-able No	Deliverable name	WP no.	Lead Partic i-pant	Esti-mated Person month	Del. Type	Security	Delivery (month)
D1	Best Motion Algorithm	1	Sco	40+10	Tech. Report and Program	Rest.	13
D2	Fully Functional VHDL Simula-tion	2	Spa	40+7	VHDL Program	Rest.	25
D3.1	Visual templates with spatial con-text - Gestalts in Space	3	Ita	19	Tech. Report and Program	Rest.	9
D3.2	Visual templates with spatio-temporal context - Gestalts in Space-time	3	Ita	19	Tech. Report and Program	Rest.	18
D3.3	Cross-modal interactions between the Gestalts define the first level of recursive scene analysis	3	Ita	19+1	Tech. Report and Program	Rest.	27
D3.4	Context-based confidence maps	3	Ita	15	Data sheet	Int.	27
D4.1	Adaptive Receptive Fields as Steerable-Scalable Filters	4	Eng	27+3.5	Program & Docu	Rest.	12
D4.2	Space Variant Mapping Methods	4	Eng	27+4	Program & Docu	Rest.	24
D5.1	Psychophysical investigations of gaze and CoI during simulated driving in humans	5	Ger	27+2	Scientific Report	Pub.	12
D5.2	Task optimized representations and the CoI	5	Ger	27+6	Tech. Report	Rest.	27
D5.3	Implementing early cogitive con-trol in a technical environment	5	Sco	18+4.5	Demo-Robot & Sci. Report	Pub.	36
D6.1	CNS*-Algorithm for generating task-optimized spatial representa-tions of neural activites	6	Bel	20+3.5	Scientific Report	Pub.	18
D6.2	CNS-solution for relating tasks to combinations of group-activities	6	Bel	15+3	Scientific Report	Pub.	36
D7.1	Calibrated Camera Sequences	7	Ind	12+1	Movie files	Int.	6
D7.2	I/O Specifications	7	Ind	10+1	Data Sheet	Int.	9
D8	Benchmarking and Testing	8	Ind	19.8+2	Tech. Eval. Report	Rest.	36
D9	Complete ECOVISION system	9	Sco	35+5	Lab Dem-onstrator	Pub.	36
D10.1	Ecovision Web Page	10	Sco	3	Web Page	Pub.	Monthly updates
D10.2	Dissemination and Use Plan	10	Sco	3+1	Plan Sheet	Rest.	6
D10.3	Scientific Publications	10	Sco	3+1	Journal Paper	Pub.	18,36
D10.4	Scientific Workshop	10	Sco	3+1	Workshop held	Pub.	36
----	Management	11	Sco	14+3	---	---	---
---	Coordination	No WP	Sco	18+4	---	---	---
Total				433.8+ 63.5			

Notes: (CNS = Computational Neuroscience)

1) Person months are subdivided into A+B, where A=hired staff and B=permanent staff.
2) Bottom rows are included in order to account for the complete number of Person Month.
3) Int. = Internal circulation within project (and Commission Project Officer if requested)
4) Rest. = Members of the consortium and Commission Project Officer
5) Pub.=Public document

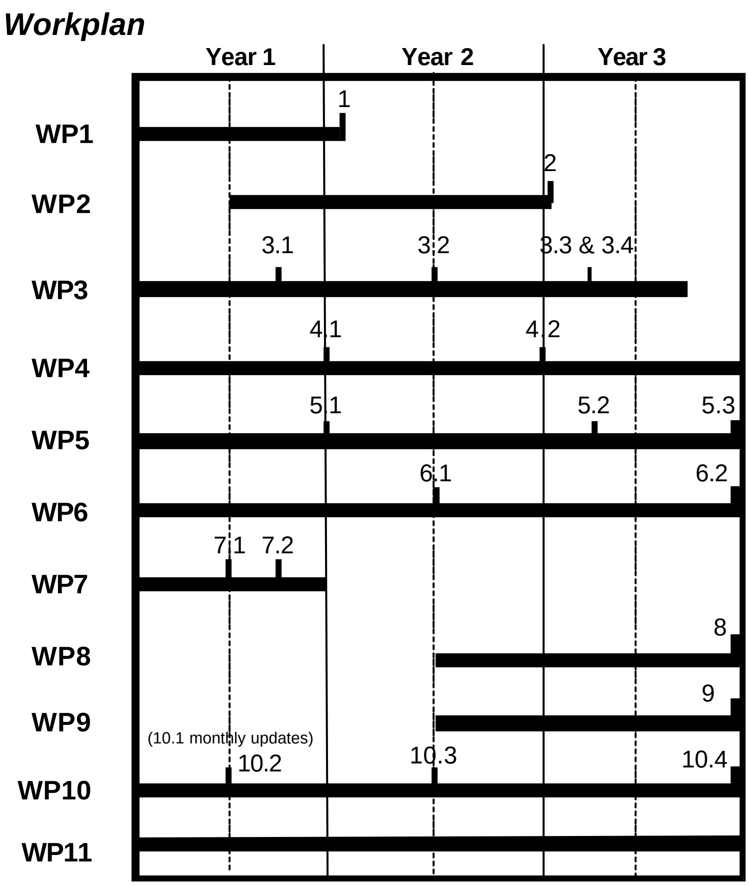
9.5 Project Planning and Timetable

The project planning and timetable is most clearly laid out in the diagram in 9.6. Specific time points are:

There will be a mid-term review at timepoint 18 by which time we will hope to have established:
1. The VHDL simulation of the best motion algorithm
2. Visual Gestalts in space-time as steerable filters
3. A preliminary mapping structure for visual relevance.
4. First strategies for predictive motor control

In the second 18 months the main "progress points" are:
5. Clear-cut evidence how humans define a CoI
6. Definition of the CoI in the ECOVISION system
7. Clear-cut task switching strategies
8. Neurophysiological evidence (or counter-evidence) for the algorithms which were implemented in ECOVISION
9. A fully functional complete ECOVISION system.

9.6 Graphical Presentation of the project components



Workpackage durations and deliverables

The continuation of a WP after its deliverable indicates a refinement process.

9.7 Project Managment

Project management constitutes an individual workpackage (see WP11, above).
The steps for monitoring the management process and its evaluation are described in detail in this WP and will, therefore, not be repeated here.

10. Clustering

Cooperation with NeuroInformatics related projects

A close co-operation with other projects in the Neuroinformatics for Living Artifacts initiative is foreseen as well as with those of the Life-like Perception Systems initiative and with related projects coming from the FET Open Scheme. The intended interaction with other projects will be centered around specific topics and take the form of joint events (at least one meeting per year), exchange of researchers and information. Part of this co-operation may take place within a specific network of excellence

11. Other contractual conditions

a) Sharing of intellectual property and dissemination is regulated independently by the partners in this consortium.
b) Subcontracting undertaken by Ita in the course of the ECOVISION project is regulated according to the legal procedures of this partner. The subcontractor(s) will take care of two distinct assignments: the major one will concern the setting up, the updating and administration of the website of the project; the minor one will concern the refinement of the software modules of Ita. Ita's personnel participating in the project will be fully engaged in the core research matters of the project, and it is thus necessary to entrust such duties to external professionals (not yet identified).

12. Supplementary reports and concertation activity

Not applicable

13. Other considerations

The partners have the clear intention to hire postdocoral students, based on their qualifications, or, when no suitable candidates are found, due to a high market demand, promising doctoral students will be hired instead. Both the postdoctoral and doctoral students will be hired under a postdoctoral and doctoral grant contract, respectively.

APPENDIX A: Consortium description

Sco: University of Stirling

The University of Stirling at the foot of the mountains centered between Edinburgh and Glasgow has one of its major strength in the field of Psychology where it is ranked amongst the top five UK Psychology departments in an external evaluation by the government. Within this field perceptional studies play a major role as well as computational neuroscience of the visual system. Broad expertise exists in psychophysics and a main center of research interest has been the influence of early cognitive aspects on the psychophysical performance in humans. The CCCN (Center for Cognitive and Computational Neuroscience, http://www.stir.ac.uk/cccn/) has successfully pursued these lines of research since the mid eighties. In this center there are 8 senior researchers and approx. 20 PhD students and postdocs.

Key Personnel: Prof. Dr. Florentin Wörgötter

EDUCATION:	1979-1985	Study of Biology and Mathematics at the Univers. Düsseldorf, FRG.
	1985-1988	Ph.D. student at the Universitätsklinikum Essen, FRG
	June 30, 1993	Degree: "Privatdozent" (PD) in theoretical and experimental Neuroscience, official member of the medical faculty of the Univ. Bochum
	Jan. 2000	Title: Professor ("apl.") at the medical faculty in Bochum
	March 2000:	Appointed "Chair" at the Dept. of Psychol., Univ. Stirling, Scotland

SPECIALIZATION

main field: Computer vision, Computational and experimental neuroscience in the visual system.

current research interests:

Visual cortex. Connectivity and the dynamic receptive field organization.

Technical applicability of ``cortex-like'' information representation in parallel computer architectures and special-purpose hardware (VLSI).

INDUSTRIAL COOPERATIONS:

MAZ Hamburg-Harburg in the framework of developing a VLSI chip for stereo-vision, based on cortical processing principles (German patent No.: 19625727).

Role in ECOVISION: Principal investigator, Project Manager, Coordinator of several WPs, Link between Ind and the others, Main characteristics: Broad interdisciplinary background.

SELECTED RELEVANT PUBLICATIONS:

1) Wörgötter, F. and Eysel, U.T. (Aug, 2000) The effects of context and state on receptive fields in the striate cortex. TINS, 23, 497-503. (comment: high-level review journal).

2) Cozzi, A. and Wörgötter, F. (2001) COMVIS, a communication framework for computer vision. Int. J. Comp. Vis. (IJCV) in press. (Comment: Neurally inspired Data fusion in Computer Vision, one of the best Computer Vision journals)

3) v. Ferber, C. and Wörgötter, F. (2000) Clustering and recognition. Phys. Rev. E (rapid comm.) 62, 1461-1464 (Neurally inspired data clustering in real Physics, second best possible hard-core Physics publication).

4) Suder, K., Wörgötter, F. and Wennekers, T. (2000) Neural field model of receptive field restructuring in the visual cortex. Neural Comp. in press. (Comment: The best journal for theoretical neuroscience)

5) Wörgötter, F., Cozzi, A. and Gerdes, V. (1999) A parallel noise-robust algorithm to recover depth information from radial flow fields. Neural Comp. 11, 381-416. (Comment: See above)

6) Wörgötter, F., Suder, K. ,Zhao, Y., Kerscher, N., Eysel, U. and Funke, K (1998). State-dependent receptive field restructuring in the visual cortex. Nature, 396, 165-168. (Comment: Theory and Experiments)

Bel: Laboratorium voor Neuro- en Psychofysiologie, KU Leuven Medical School, Belgium

The laboratorium voor Neuro-en Psychofysiologie is involved in the interdisciplinary study of the neuronal mechanisms of Visual perception, attention and memory. It counts 20 researchers (3 permanent staff and 17 doc or post doc) and 7 technicians. It features an unique range of techniques used in human and non human primates.: single cell recording, lesion and behavior, psychophysics, functional imaging (PET and fMRI) in humans metabolic mapping (double label 2-deoxyglucose) and functional MRI in the awake fixating monkey , neural networks and modeling. It has studied many different regions of human and monkey cortex. Present research interest centers around extraction of 3D surfaces from stereo, motion and monocular cues, motion processing, task-dependence, spatial and featural attention , perceptual learning and self-organization,.

Role in ECOVISION: Coordinator of WP6, central biophysical computational neuroscience partner.

Key Personnel: Prof. Dr. Marc M. Van Hulle is the leader of the computational neuroscience group of the Laboratorium voor Neuro- en Psychofysiologie. He is currently an associate Professor at the K.U.Leuven medical school. He received an M.Sc. degree in Electrotechnical Engineering (Electronics) and a Ph.D. in Applied Sciences from the K.U.Leuven, Leuven in 1985 and 1990, respectively, and also holds B.Sc.Econ. and MBA degrees. In 1992, he has been with the Brain and Cognitive Sciences department of the Massachusetts Institute of Technology (MIT), Boston (USA), as a postdoctoral scientist. Marc M. Van Hulle has authored the monograph Faithful representations and topographic maps: From distortion- to information-based self-organization, John Wiley, 2000, and more than 60 scientific publications. He is also founder and director of Synes N.V., the data mining spin-off of the K.U.Leuven (30 employees).

SELECTED RELEVANT PUBLICATIONS:

1) Van Hulle, M.M. & Gautama, T. (2000). Modeling motion processing in macaque area MT/V5: From single cells to population codes. In: Modulation of Neuronal Signalling: Implications for Active Vision, G.T. Buracas, O. Ruksenas, T.D. Albright (Eds.) (invited paper). NATO-ASI series, in press.
2) Van Hulle, M.M. (1999). Faithful representations with topographic maps. Neural Networks, 12, 803-823.
3) Van Hulle, M.M. (1999). Density-based clustering with topographic maps. IEEE Trans. Neural Networks, 10(1), 204-207.
4) Van Hulle, M.M. (1998). Kernel-based equiprobabilistic topographic map formation. Neural Computation, 10(7), 1847-1871.
5) Liu, L., & Van Hulle, M.M. (1998). Modeling the surround of MT cells and their selectivity for surface orientation in depth specified by motion. Neural Computation, 10(2), 295-312.
6) Van Hulle, M.M. (2000). Faithful representations and topographic maps: From distortion- to information-based self-organization, New York Wiley (also to be published in Japanese by Kaibundo, 2001).

Other personnel: Prof. Dr. G.A. Orban, Laboratorium voor Neuro- en Psychofysiologie, KU Leuven, Medical School, Belgium

Role in ECOVISION: Advisory role in developing the algorithmic procedure from computational neuroscience models in WP7. He will also deliver the experimental data, in addition to the already published and, thus, publicly available data and results, that we will use for constraining our algorithmic procedure.

SELECTED RELEVANT PUBLICATIONS

1) Sáry Gy., Vogels R. and Orban G.A. (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science, 260:995-997.
2) Vandenberghe R., Duncan J., Dupont P., Ward R., Poline J.B., Bormans G., Michiels J., Mortelmans L., Orban G.A. (1997) Attention to one or two features in left or right visual field: a positron emission tomography study. Journal of Neuroscience, 17:3739-3750.
3) Vanduffel W., Payne B.R., Lomber S.G. and Orban G.A. (1997) Functional impact of cerebral connections. Proceedings of the National Academy of Sciences USA, 94: 7617-7620.
4) Orban G.A., Sunaert S., Todd J.T., Van Hecke P. and Marchal G. (1999) Human cortical regions involved in extracting depth-from-motion. Neuron, 24:929-940.
5) Vanduffel W., Tootell R.B.H. and Orban G.A. (2000) Attention-dependent suppression of metabolic activity in the early stages of the macaque visual system. Cerebral Cortex, 10:109-126.

Eng: University College London

University College London is one of the foremost teaching and research institutes in the United Kingdom. It is consistently rated among the top three multi-faculty universities in the UK. The Psychology Department, one of the largest in the UK (>40 teaching staff), received a (best possible) 5-rating in the last UK Research Assessment exercise. Since then UCL has established the Institute in Cognitive Neuroscience and the Gatsby Centre for Computational Neuroscience, which has attracted further large internationally prominent groups to the Psychology Department. In addition UCL has set up CoMPLEX, an interdisciplinary grouping to support work in mathematical biology, and CAIS (Centre for Advanced Instrumentation). Prof. Johnston has links (including PhD supervision) to both these cross-departmental groups.

Key Personnel: Prof. Alan Johnston, Dept. of Psychology

EDUCATION

1972-1976	University of Aberdeen, BSc Psychology (Hons) 2.1
1976-1980	University of York, D.Phil.

APPOINTMENTS

1980-83	Research Fellow, Brunel University
1983-84	Experimental Officer, Brunel University
1984-87	Research Fellow and MRC Grantholder, Brunel University
1987-94	Lecturer in Psychology, University College London.
1994-99	Reader in Psychology, University College London.
1999-	Professor in Psychology, University College London

Role in ECOVISION: Link between Psychology and the other partners, prime psychophysisist, expertize in steerable filters, coordinator of WP4 and links WP3 to WP5 via WP4.

SELECTED RELEVANT PUBLICATIONS.

1. Johnston, A., McOwan, P.W. and Buxton, H. (1992) A computational model of the analysis of first and second order motion by simple and complex cells. Proc. R. Soc. Lond. B., 250, 297-306.

2. Johnston, A, McOwan, P.W. and Benton, C.P. (1999) Robust velocity computation from a biologically motivated model of motion perception. Proc. R. Soc. Lond., B., 266, 509-518.

3. Johnston, A, Benton, C.P. and McOwan, P.W. (1999) Induced motion at texture-defined motion boundaries. Proc. R. Soc. Lond., B., 266, 2441-2450.

4. Johnston, A. and Wright, M.J. (1983) Visual motion and cortical velocity. Nature, 304, 436-438.

5. Nishida, S and Johnston, A.(1999) Influence of motion signals on the perception of spatial pattern. Nature, 397, 610-612.

The first article [1] introduced our gradient model of motion processing in the human visual system, which describes how one can combine multiple spatio-temporal derivative operations to compute motion in a robust way. This work is extended in [2] which describes how a 2Dspace-time model can compute motion direction as well as speed in a way that us robust to the addition static noise. The influence of static noise is removed at late stage in processing. In [3] we show the model can detect and predict perceived speed of motion in second-order or texture-defined motion sequences. In [4] we show that motion thresholds and motion adaptation varies with visual eccentricity in a way that was predictable from the cortical magnification factor. Later work indicated that the variation was such the ground plane would appear to move at a constant speed (while looking ahead) even though image velocities vary from fovea to periphery. The work in [5] showed that the motion after effect induce changes in perceived spatial position indicating that there is an interaction between motion computation and processing of spatial location.

Ger: Ruhr-Universität Bochum
The Ruhr-Universität Bochum with more than 30.000 students is one of the main centers of vision science in Germany and hosts the nationally funded research center on 'neurobiology of vision'. The vision community at the Ruhr-Universität includes members of the Biological, Medical, and Psychological Faculties, and the Institute of Neuroinformatics. The applicant is the head of the work group on 'Computational and Cognitive Neuroscience' which is part of the Department of General Zoology and Neurobiology. The Department consists of 8 staff scientists and associated postdocs and students and a permanent staff of 10 technicians. Within the work group, there are two postdocs, three PhD students, and one technician. The work group has been active on computational and physiological aspects of motion vision for several years. Research included computational modeling of optic flow analysis, psychophysical and electrophysiological experiments on optic flow processing in humans and monkeys, experimental investigations of motion-stereo combinations and modeling thereof, as well as studies of eye movements and low-level motion perception. More information can be found at http://www.ruhr-uni-bochum.de/neurobiol/ml.htm

Key Personnel: PD Dr. Markus Lappe
1982-1989 Study of physics and mathematics in Marburg and Tübingen
1990- 1992 Guest researcher at NIH, Bethesda, USA
1992 PhD in Physics at the Max-Planck-Instute for Biological Cybernetics, Tübingen
1993-1998 Researcher at the Ruhr-University Bochum
1998 Habilitation in Neurobiology
since 1999 Group leader in 'Computational and Cognitive Neuroscience' at the Ruhr-University Bochum

EXPERTISE:
Computational analysis of optic flow
Psychophysics of optic flow perception and heading estimation
Physiology of optic flow processing in primates
Role of eye movements in optic flow analysis
Contribution of stereo vision to optic flow analysis, computational issues and experimental observations
Virtual reality systems for movement simulation

Role in ECOVISION: Prime optic flow partner

SELECTED RELEVANT PUBLICATIONS.

1) Lappe, M. and Rauschecker, J. P. (1995). Motion anisotropies and heading detection. Biological Cybernetics 72:261–277. (Presents a neural implementation of an efficient algorithm for heading estimation from optic flow in the context of anisotropic retinotopic maps in the visual cortex)

2) Lappe, M. (1996). Functional consequences of an integration of motion and stereopsis in area MT of monkey extrastriate visual cortex. Neural Computation 8:1449–1461. (Presents an analysis of stereo-flow interactions in primate area MT and an neural algorithm that uses these interactions to increase the robustness of optic flow representation)

3) Grigo, A. and Lappe, M. (1999). Dynamical use of different sources of information in heading detection from retinal flow. Journal of the Optical Society of America A 16:2079–2091. (Psychophysical investigations of the dynamics of human optic flow perception. It shows how visual motion perception and eye movement signals interact on the timescale of several hundred milliseconds in order to provide estimates of heading)

4) Lappe, M. and Grigo, A. (1999). How stereo vision interacts with optic flow perception: Neural mechanisms. Neural Networks 12:1325–1329. (Compares three proposed types of stereo-flow interactions in primate cortical neurons with recent experimental results form human psychophysics)

5) Lappe, M., Bremmer, F., and van den Berg, A. V. (1999). Perception of self-motion from visual flow. Trends in Cognitive Sciences 3:329–336. (A review of human optic flow perception and associated neurocomputational models)

Lappe, M., editor (2000). Neuronal Processing of Optic Flow. Academic Press.
(A comprehensive description of neural mechanisms of optic flow analysis in animals and humans)

Ita: University of Genoa

At the University of Genoa, the School of Engineering, ranked third among 36 ones in a recent national survey, has one of its major strength in the field of Information and Communication Technology, area of development and reorganization for various industries in the Liguria Region.

The Department of Biophysical and Electronic Engineering (DIBE), with its research groups active in Electronics, Telecommunication, Bioengineering, Circuits and Systems, Electromagnetic Fields, and Mathematical Physics, is a well established centre of collaborations and contracts with the European Commission, national and international Research Agencies and Industries. It counts 34 professors and researchers, 10 lecturers, 29 Ph. D. students, 20 units of technical and administrative staff.

Key Personnel 1: Prof. Giacomo Mario Bisio
1965 Electronic Engineering at the University of Genoa
1971 M. Sc. degree in Electrical Engineering from Stanford University, CA, USA.
1966-1983 Researcher at the CNR, (Italian) National Research Council.
1983 Associate professor at the School of Engineering, University of Genoa.
1990 Full professor of Microelectronics.
1992-98 Director of the Department of Biophysical and Electronic Engineering, and
Co-ordinator the Board of Department Directors (1996-98).
Co-ordinator at local, national, and European level of the following projects: CEC Esprit Parallel Computing Action, 1990; CEC Esprit VLSI Design Action, II phase, 1992-94; CEC Esprit - LTR, 1994-97 Cormorant Project "Cortical Maps of Resistive Anisotropic Networks".

Key Personnel 2: Dr. Silvio P. Sabatini
1992 Laurea in Electronic Engineering (summa cum laude) at the University of Genoa
1996 PhD in Electronic Engineering and Computer Science at DIBE, University of Genoa
since 1995 Research fellow at the ``Physical Structure of Perception and Computation'' (PSPC)
Research Group at DIBE (http://pspc.dibe.unige.it/)
since 1999 Assistant professor in Computer Science.

Role in ECOVISION: Coordinator of WP3, Links chip design with theoretical neuroscience.
Expertize: Physical processes of biological vision to inspire the design of artificial perceptual machines based on neuromorphic computational paradigms.

- Gabor operators and their discrete analogs; Representation properties of cortical maps.
- Stereo depth and motion analysis.
- Signal and image processing; Circuits and system theory; Microelectronics.

SELECTED RELEVANT PUBLICATIONS

1) S.P. Sabatini, and F. Solari. An architectural hypothesis for direction selectivity in the visual cortex: The role of spatially asymmetric intracortical inhibition. Biological Cybernetics, 80(3):171-183, 1999. (A model that evidences average structural/functional principles of intracortical connectivity and relates them to the underlying single-cell motion properties.)

2) L. Raffo, S.P. Sabatini, G.M. Bo and G.M. Bisio. Analog VLSI circuits as physical structures for perception in early visual tasks. IEEE Trans. Neural Net., 9(6):1483-1494, 1998. (Arrays of simple interacting analog units reacting collectively to spatio-temporal input stimuli.)

3) S.P. Sabatini, L. Raffo, and G.M. Bisio. Functional periodic intracortical couplings induced by structured lateral inhibition a linear cortical network. Neural Computation, 9(3):525-531, 1997. (A neural field model of simple cell functionality. The origin of highly-structured Gabor-like receptive fields is related to the propagation properties of cortical activation triggered by the spatial organization of the recurrent lateral inhibition scheme on the orientation map.)

4) L. Raffo, S.P. Sabatini, M. Mantelli, A. De Gloria, and G.M. Bisio. Design of an ASIP architecture for low-level visual elaborations. IEEE Trans. on VLSI Systems, 5(1):145-153, 1997. (A VLSI digital architecture suitable to map cooperative-competition processing schemes in multilayer networks.)

Spa: University of Granada

The University of Granada with more than 60000 students is considered one of the most significant centers of high education and research in Spain. During the last years is growing mainly in the area of technological departments. The research group CASIP (Circuits And Systems for Information Processing) at the Department of Computer Architecture and Technology (http://atc.ugr.es/) is composed by 21 staff members and 10 PhD students actively involved in research tasks. CASIP has long expertise in VLSI implementation, signal processing, smart control, parallel computation, among other fields. One of the main research lines focuses on implementation of neural network circuits as ASICs from different design paradigms: full-custom, semi-custom, and programmable logic devices (PLD's). In particular, we have designed focal plane Cellular Neural Networks, Self-organization maps (Kohonen maps), artificial retinas and general purpose spiking neuron prototypes.

Key Personnel: Prof. Dr. Eduardo Ros
1993 B.Sc. Degree in Physics in 1993, University of Granada, Spain
1996 B.Sc. Electronic Engineering, University of Granada, Spain
1997 PhD, University of Granada, Spain.
1998 Guest researcher at the King's College, London.
Since 2000 appointed Associative Professor at the University of Granada (Department of Technology and Architecture of Computers).

Further staff personnel that will be involved in the project:
Prof. Dr. F.J. Pelayo (Full Professor)
Prof. Dr. B. Pino (Associate Professor)
Prof. D. Palomar (Assistant Professor)

EXPERTISE
- Neuromorphic Engineering and VLSI Implementation of Neural Systems.
- Multiple-chip communication paradigms.
- Signal and image processing; ASICs implementation.
- Adaptive systems.

Role in ECOVISION: Chip and board designer, Coordinator of WP1 and WP2.

SELECTED RELEVANT PUBLICATIONS.

1) F.J. Pelayo; E. Ros; X. Arreguit; A. Prieto: VLSI Implementation of a Neural Model Using Spikes, "Analog Integrated Circuits and Signal Processing", Special Issue on Neuromorphic Engineering, Kluwer Academic Publishers, Vol. 13, pp. 111-121, 1997. (This paper describes a full custom VLSI implementation approach of a neural model whose inherent characteristics are useful for processing tasks such as stimuli correlation and motion detection.)

2) M. Anguita; F.J. Pelayo; E. Ros; D. Palomar; A.Prieto: Single and Multiple Chip Approaches to Cellular Structures for Vision Tasks and Image Processing, "Analog Intagrated Circuits and Signal Processing", Kluwer Academic Publishers, Vol. 15, pp. 263-275, 1998. (Different neural circuit approaches are described in order to be used for vision tasks.)

3) E. Ros; F.J. Pelayo; D. Palomar; I. Rojas; J.L. Bernier; A. Prieto: Stimulus Correlation and Adaptive Motion Detection using Spiking Neurons, "International Journal of Neural Systems", Vol. 9, No. 5, pp. 485-490, 1999. (Specific neural structures perform motion detection and stimulus correlation with adaptive time constants.)

4) E. Ros; F.J. Pelayo; P. Martín-Smith; D. Palomar; A. Prieto: Competitive and Temporal Inhibition Sturctures with Spiking Neurons , "Neural Processing Letters", Vol. 11, 3, 197-208, 2000. (This paper describes how the neural prototypes proposed in the previous papers can be used to carry out competitive processing tasks. Temporal inhibition structures can be useful to emulate attention mechanisms.)

All these papers are related with VLSI approaches of neuron models that can be used for different simple vision tasks.

Ind: Hella KG Corp.
Hella Corp. is a automotive industry supplier with a worlwide staff of 22.000 and a total sales result of about 2.2 Mio €. Main products are automotive lighting and electronics.

Hella - the name of this traditional company has been a byword for innovative lighting technology for 100 years. Founded in 1899 as a factory for streetlights, headlamps and horns, the company today enjoys the confidence of nearly all the world's automobile manufacturers in Hella lighting technology and electronics. With over 30 production locations worldwide, the group is a global partner for the automotive industry. Over 1,000 engineers and technicians work in the field of research and development (R&D). In Germany, Hella employs nearly 12,000 persons. A further 10,000 work at production locations outside Germany, as well as in subsidiary and affiliated companies (in Europe and overseas). Worldwide, Hella Group turnover for business year 1998/99 came to 4.3 thousand million DM.

In addition to the Lighting Technology sectors, the Automobile Electronics Division is a second field of equal importance for the activities of Hella. The main product groups are heating and climate control, body electronics, lighting and electronics, drive train control and systems and components.

Products from the drive train control sector are supplied to all European carmakers. The product range includes speed regulators, E-gas systems, drive train regulators, actuators for switched intake manifolds, turbocharger regulators and intake air throttles.

Range scanning sensors based on Radar or Lidar (Laser) cannot provide all the information that will be needed for ACC (Adaptive Cruise Control) and Collision Avoidance systems in future smart cars. Vision systems can provide the missing information needed and can increase redundancy and safety. Stereo distance measurement and motion estimation in parallel are both tasks to detect stationary obstacles and moving objects in front of a moving vehicle. For all the state-of-the-art processor parallel image data analysis in real-time is still a high-performance job. Dedicated Hardware-based solutions like early vision chips (motion, stereo) can help to overcome this problem and support a lean software-development.

Hella Corp. is a automotive industry supplier with a worldwide staff of 22.000 and a total sales result of about 2.2 Mio €. Main products are automotive lighting and electronics.

Key Personel: Martin Mühlenberg is currently a member of the technical staff at the predevelopment department electronic systems of Hella KG Corp. His fields of work are vehicle guidance and driver assistant systems based on machine vision. He received the diploma degree in electrical engineering at University of Paderborn in 1996.

Role in ECOVISION: Industrial testing and benchmarking. Control of the ECOVISION system design.

APPENDIX B: Contract preparation forms