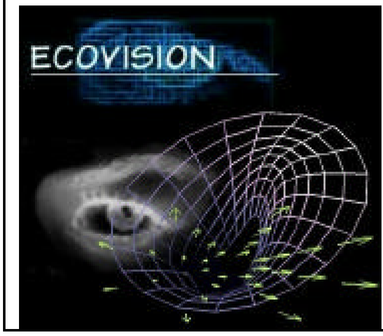

ECOVISION

IST – 2001 - 32114

Artificial vision systems based on early-cognitive cortical processing

Periodic Progress Report N°:1

Covering period 1.1.2002-31.12.2002

Report Version: 1.0

Report Preparation Date: 12.11. 2002

Classification: public

Contract Start Date: 1.1. 2002 Duration: 3y

Project Coordinator: (**Sco**) University of Stirling, Prof. Dr. F. Wörgötter, Stirling, UK

Partners: (**Bel**) KU Leuven, Prof. M. v. Hulle

(**Eng**) Univ. College London, Prof. A. Johnston

(**Ger/**Mun) Ruhr-Univ. Bochum & Univ. Münster, Prof. M. Lappe

(**Ita**) Universita' degli Studi di Genova, Prof. S. Sabatini & Prof. G. Bisio

(**Spa**) Univ. de Granada, Prof. E. Ros

(**Ind**) HELLA Hueck KG, Lippstad, Ger, Mr. M. Mühlenberg

Project funded by the European Community under the "Information Society Technologies" Program (1998-2002)

Revised PPR as of Jan. 03

This version of the PPR represents the state of progress as of 01/01/03.

The work planned for Dec. 2002 has proceeded as planned, thus no changes had to be made to the PPR as submitted for the project review early Dec. 2002.

Table of Contents:

Executive Summary

Goals of ECOVISION: The goal of this co-operative project is to investigate and implement advanced mechanisms which allow for designing high-performance machine vision systems. These mechanisms are motivated by early cognitive cortical processing in the visual system of vertebrates and part of this project is concerned with models of such systems (WP6). The central aspect which distinguishes ECOVISION from other similar approaches is that we are seeking to employ highly co-operative multi-modal mechanisms in order to achieve improved image analysis (WP3-5). This becomes possible because, at the same time, we are addressing the problem of early visual pre-processing by means of dedicated hardware (FPGAs, WP1,2). These front-end algorithms provide in a fast and efficient way the input data for the early-cognitive post-processing. The application goal of ECOVISION is to employ these techniques in driver assistant systems.

Summary of Achievements: We are at month 10 of the ECOVISION project at the time of compiling this report. We are very happy to say that the consortium has achieved in almost all WPs more than what was planned for the first 12 months. Two deliverables have finished ahead of time. At total of 7 journal papers are already published or in press by now, furthermore there are 17 conference contributions. In addition ECOVISION has sponsored two workshops, one in Granada (**jcra'2002**) and another one that will take place mid November in Bochum (*Dynamic Perception*). These activities are listed in the Progress Report Sheets at the end of the report (pg. 51-63). The consortium operates truly in a highly co-operative way as specifically documented by several work-visits (i.e., visits were actual project work was performed, not just discussions, see below). The co-operation with our industrial partner (called Ind) is also very successful and HELLA Hueck KG has set up the application framework for the ECOVISION system and has actively provided us with a test site for recording the necessary camera sequences, which was done in co-operation between Sco and Ind. This is remarkable taking into account that ECOVISION is strongly centered onto basic research which is not necessarily in the focus of application oriented industry.

Two amendments of the contract are currently being processed:

1) Partner Ger has moved from Bochum to Münster (Mun). This has not affected his work and the associated deliverable will finish in time, but more experiments will now be performed in Münster to finally consolidate the results.
2) Partner Bel has faced some unforeseen internal administrative problems which require reallocating his budget fully into personnel covering all other costs from different existing sources. In addition he is now also involved in WP5.2 which requires allocating more man-months for him. (This in itself would not have required an amendment). The reasons for this revision will be described at WP5 below. Furthermore, there is a smaller technical change of WP6.2 because less experimental data can be used as originally planned for this WP (see below).

The diagram below (Fig. 1) summarizes the state of the project. In its center lies the ECOVISION software. This is a huge software package which was in part pre-developed by Norbert Krüger before the start of the project, who is employed in ECOVISION by Sco (http://www.cn.stir.ac.uk/ImageAnalysis/MoInS/index.html). This software is now continued to be developed and adapted to the tasks of the ECOVISION project. Its purpose is to extract multi-modal visual features from image sequences. In every single analysis step there are color, orientation, flow, stereo-disparity, contrast, and contrast-transition extracted from a frame-pair. In addition, ECOVISION has set up a high-resolution camera pair for acquiring the image sequences in video real-time. This has been done in co-operation with Hella Hueck KG, where real driving scenes have been recorded this summer on site (WP7). At the front-end Spa has together with Eng started to implement a specific flow/motion algorithm on an FPGA (WP1, WP2). This part of the work is far advanced and will (probably) lead to a real implementation at the end of the project which is not planned in the initial proposal plan (Only a "feasibility analysis" and a VHDL simulation is part of the Technical Annex (TA) so far!). If we should reach this point, we would be able to ultimately replace the low-level motion part of the ECOVISION package with this hardware solution.

The central goal of ECOVISION is to develop adaptable filters which change their properties in a task dependent way. Eng has developed such a set of filters for low-level feature extraction (orientation, motion, WP4). Sco and Ita have successfully developed higher level filters (which use the extracted low-level information) which perform advanced flow-field and stereo analysis by means of generalized grouping mechanisms, which are not based on any image heuristics (as usual), but instead use measurable statistical image properties (WP3). The filters are meant to change in a scene- and (more importantly) in a task-dependent way. Partner Ger/Mun has by now provided eye-movement data showing where humans look when asked to perform different tasks in a driving simulation (WP5).These results can be used by us to control the adaptation mechanisms in WP3. Partner Bel has performed theoretical neuroscience studies of how different forms of visual attention (spatial attention as compared to featural attention) can be modeled with a map of filters (receptive fields, WP6). Featural attention can be regarded as similar (or maybe even identical) to task-dependent attention. Thus, we expect that the models of Bel will give us ideas of how to actually implement such mechanisms in the ECOVISION software.

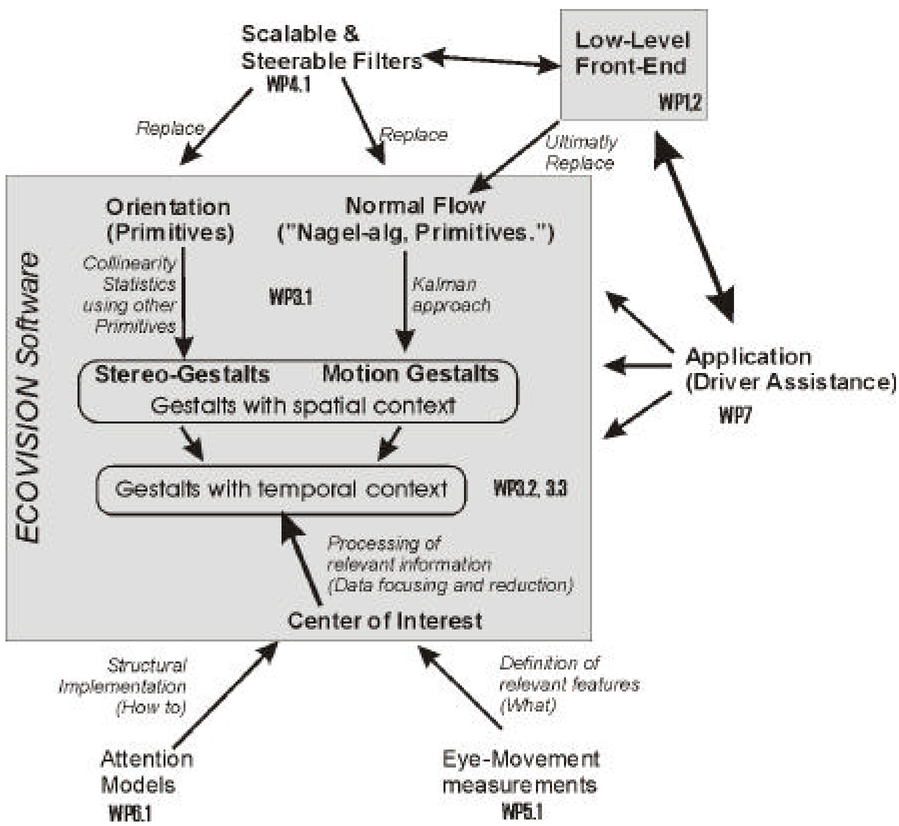

Fig. 1) Cross-link diagram of the different project parts reflecting the state of the work end of October 02

In general this report will be presented in a one-to-one fashion exactly reflecting the steps defined in the TA. This should allow reading the documents side-by-side. We will focus on facts and try to present them in a concise, brisk and concatenated way using diagrams wherever possible. We would ask the reviewers to assess the details of the performed work through the attached complementary materials (deliverables, reports, etc.). This will remove redundancy and save time.

<u>Co-operative activities:</u>

(only members of the consortium are listed as participants)

1) ECOVISION Subgroup Discussion meeting, Granada, Dec, 17,18, (2001) Participants: Spa, Eng, Ind (Eduardo Ros, Alan Johnston, Martin Mühlenberg), Topics: Real time optic flow implementation and its potential application fields.
2) ECOVISION Subgroup Discussion meeting, Jan 5,6, Bochum, Participants: Ita, Ger/Mun, Sco, (Silvio Sabatini, Markus Lappe, Norbert Krüger, Florentin Wörgötter), Topics:
3) ECOVISION Subgroup meeting, Jan 8, Lippstadt, Participants: Ind, Sco (Martin Mühlenberg, Tillman Seubert, Florentin Wörgötter)

Work-visits:

1) Jan 10 to Feb 28, Participants: Eng, Spa (Alan Johnston, Jason Dale and Guillermo Botella). London. Work in performed (in progress): Configuration and parameter study of McGM (optic flow algorithm proposed by Eng).
2) Work-visit: June 10-14, Participants: Sco, Ind (Norbert Krüger, Martin Mühlenberg) Work Performed: Recording of high-resolution color stereo sequences (front-view stereo cameras mounted in a car)
3) Work-Visit: June 16-18, Participants: Sco, Ger/Mun (Norbert Krüger, Markus Lappe), Münster, Preparation of the internal report on: "Using Statistical Properties of Optic Flow for Flow Correction"

First General ECOVISION Meeting: June 29-July 01, Stirling, Scotland: Participants: All. Topics: Interim Reporting and general project planning.

Second General ECOVISION Meeting: Nov. 16,17, Münster, Germany, Participants: All. Topics: First Year Reporting and general project planning.

Project Part 1: The VLSI Front-End: VLSI implementation of motion analysis (WP1, WP2)

Note: Own publications are referred to as code.n, where “code” is the partner code (e.g., Sco, Bel, etc.). The corresponding references are listed in the progress report sheet at the end of this report.

WP1: Motion algorithms

This WP addressed the problem of how to find the best suited low-level motion (i.e., flow-field) algorithm for FPGA implementation.

Scheduling: ahead of planned schedule by 1-3 months.
Central achievements: D1 (due month 13) is to a large degree finished by now. A working integer version of the chosen motion algorithm exists on a Matrox Board. Performance examples can be viewed at
http://www.pspc.dibe.unige.it/~ECOVISION/news/
(see also http://electra.psychol.ucl.ac.uk/jason/).

Planned and performed steps:
Step a1) Assessment of Eng’s motion algorithm (called Multi-channel Gradient Model) if it will in principle be feasible for an FPGA implementation.

Scheduling: Month 0-13
Performed actions: visit of a contracted personnel (of Spa) to stay at the UCL (Eng) for 16 weeks and co-operative work on McGM minimum version.
Results: The Multi-channel Gradient Model (McGM) was found to be feasible and well-suited for FPGA implementation.
Documentation: Internal Technical Report: *The Multi-channel Gradient Model and its Real-Time Implementation* (51 pg) of a Matrox Board integer implementation of the McGM algorithm co-operatively compiled by Eng and Spa, used by Spa for the FPGA-implementation assessment. This document, which was compiled in addition to the original workplan is highly detailed and provided an excellent basis for the assessment of Eng’s algorithm.
Publications: Spa.1-Spa.3

Step 1b) Assessment of other motion algorithms
Performed actions: Ita and Spa performed an assessment of the Simoncelli and Heeger’s motion algorithm.
Results: Rather early on this algorithm was found to be less-well suited for FPGA implementation in terms resource requirements. Two main problems arise: (i) the Simoncelli and Heeger’s algorithm performs on average with respect to other “classical” algorithms, but paying a higher computational load, (ii) the algorithm is contrast-dependent, which is a crucial draw-back in real-world, open-air (automotive) sequences. This, however, needs to be finally confirmed and for the remaining time of WP1 we are planning to make an in-depth comparison with the McGM algorithm.
Documentation: Short report: *Some notes on the Simoncelli and Heeger algorithm* (6 pg, compiled by Ita in co-operation with Spa).
Publications: ---

Steps a2,a3) actual mapping of the integer implementation onto the FPGA architecture (WP1 & WP2, first steps) and accuracy estimation.
Scheduling: Planned for month 12-15 (i.e. after D1)

Performed actions: n/a
Revised Planning: These steps have started just now because D1 is almost and will be finished ahead of schedule.

Deliverables: No Deliverables due until month 13. In spite of this Eng and Spa have already a running version of McGM on a Matrox, and they are performing experiments to determine the minimum requirements (orientations, resolution, filter masks, bit precision, etc) of the McGM. This is possible through the efficient co-operative work between Eng and Spa.

WP2: FPGA simulation

As planned, only preliminary steps have been performed so far concerning the VHDL implementation of the McGM algorithm on an FPGA (mainly WP2).

Revisions for WP1,2: Progress in WP1,2 is by far better than expected. Steps a2,a3 should finish about 3 months ahead of schedule. Thus, we are hoping to obtain the VHDL code also about 3 months earlier than planned. If this will be the case we will find ourselves in the position to actually implement and test the McGM algorithm, which is not part of the ECOVISION workplan but will be the next consequent step. A the moment this cannot be clearly foreseen but by end of year 2 we will be able to decide about this. If possible we would at the end of the project hold a working prototype in our hands. This is of vital importance for our industrial partner (HELLA Hueck Kg, “Ind”) and we are already in the process of planning the IPR handling.

Influence on the other WPs: We are planning to use the real-time motion analysis Matrox implementation as the motion front end for WP3-WP5. This is expected to take place in year 3 of the project when the ECOVISION system shall be put together.

Project Part 2 Investigations of early cognitive mechanisms and implementation in a distributed system (WP3-WP5)

This project part is centered on the question of how to define measure and use "Gestalts" for visual scene analysis. In general a Gestalt is a visual entity which takes visual context in space and time into account. Thus, we regard Gestalts as advanced visual filters similar to a receptive field of a neuron in a higher visual area (primary visual cortex and beyond). ECOVISION seeks to define Gestalts (implicitly or explicitly) in an adaptive (scene- and task-dependent) way.

WP3 is concerned with the definition of Gestalts in space (Orientation, Stereo) and time (Motion, Flow)
WP4 is concerned with the aspect of how to adapt these Gestalts to a changing context.
WP5 is concerned with the question what are the relevant aspects of a visual scene (with respect to a given task) to which a Gestalt should adapt.

We have made excellent progress in all three WPs. We have arrived at the description of orientation, stereo and motion Gestalts. All these filters are adaptable in a controlled way. We have designed a common framework (large machine vision computer program-package) where these filters will now be implemented in year 2. Within this program it is already by now possible to combine Gestalts (through specific feature binding mechanisms) in order to arrive a more complex semantic image descriptors which we call "semi-complex feature constellations" (SCFC). We have addressed the question how Gestalts should change in a task dependent way in order to create a task dependent semantic context.

WP3 Adaptive receptive fields as context sensitive visual filters – Gestalts in space-time

In this WP we first extracted from stereo image sequences (see Fig. 2, B,G, C,H; pg 27) so called (local, multi-modal) visual Primitives[1] as precursors of the Gestalts (see Fig. 2, D,I, E,J; pg 27). This was achieved by means of a software package pre-developed by Sco, which allows for one-shot analysis of the modalities: position, orientation, contrast-transition, color, stereo and optic flow for any given image pair. In the next step those Primitives were combined into Gestalts. Unlike commonly done, here this was not achieved by applying heuristic assumptions; instead Gestalts were defined by utilizing the intrinsic properties of the images (their statistics and their inherent structural features). Specifically we defined Stereo-Gestalts and Motion-Gestalts by two different algorithmical procedures which combine the Primitives. Thus, those Gestalt are filters which *spatially* adapt to the underlying (spatial) structure of the images (WP3.1). Both Gestalts thereby realize complementary representations as planned in the TA. As a consequence this part of the ECOVISION system implements hierarchical image analysis: from pixels to Primitives to spatial Gestalts. WP3.2 leads to *temporally* adaptive Gestalts (see below) which change their shape following the changes in an image sequence.

[1] A Primitive is not a "feature" as commonly obtained from visual feature analysis algorithms. Features are normally defined at every pixel. Primitives on the other hand are "features" that apply to larger pixel groups in a shared way. Thus, Primitives lead to a sparse representation and to substantial data and noise reduction.

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

Scheduling: WP3.1 in time
Central achievements: Primitives have been extracted from the image sequences. Stereo Gestalts are defined by the statistical properties of the images. Gestalts in the Motion domain have been defined in relation to the structural gradients of flow-fields. A recursive algorithm is being developed to reliably detect these motion-Gestalts.

Planned and performed steps:
Step 1) Measurement and Definition of Gestalts in Space

Scheduling: in time
Performed actions: Sco and Ita agreed to divide the work: Sco focused on stereo-Gestalts, while Ita worked on motion-Gestalts. Both partners co-operated with Ger/Mun (visit of Sco to Ger/Mun,18.-20.6. (addressing the problem of rigid body motion, necessary to embed WP3.1 into WP3.2, discussion between Ita and Ger/Mun (Oct. 15) on concrete implementation of local *vs* global flow templates.)

Step 1a) Motion

Problems: Typical existing flow-field algorithms (which can be used to define motion Gestalts) measure only the "normal flow" (flow orthogonal to an edge/aperture). This flow is ill-suited to objectively (i.e., statistically) measure and define motion Gestalts. The immediate solution was to define motion Gestalts first at a more abstract level, looking at the geometrical properties characterizing vector fields. At the same time we have in WP5.2 started to address the problem of *how to measure* motion Gestalts in an objective way. WP5.2 addresses the aspect of "Visual Relevance", thus the problem of how (and what) to measure motion Gestalts belongs into this WP (see below), and some progress has been made there too.
Results: Flow fields are less diverse than other visual features and normally consist of large patches of flow-patterns which result from a common cause (e.g.; from ego-motion or from object motion). These flow field patches can be characterized on the basis of their first-order differential properties. This allowed defining basic templates which cover the central properties of mixed (real-world) flow-field patterns. Definition of flow-templates was performed in co-operation with Ger/Mun who contributed to evidence relationships between motion Gestalts and motion templates through the common cause idea. These flow-templates currently serve as our motion-Gestalts. (Motion can be extracted in a straight forward way from flow). Ita has invented a recursive algorithm to automatically identify these patterns. This algorithm is based on a Kalman filter. When confronted with a flow-map from a real image sequence the filter converges after several iterations onto the most likely flow-Gestalt for a given frame. (Here we use the term flow-Gestalt and motion-Gestalt as synonyms.) Currently this filter operates frame-by-frame, but the recursiveness of the Kalman filter is expected to allow for an implementation which will continuously adapt to changing visual context. As such the filter can already be implemented in the ECOVISION system.

Documentation: A Kalman filtering approach to early-cognitive vision: basic ideas and progress report (internal report, Sept. 19, 32 pg.)
Publications: Ita.1, Ita.2
Deliverables: 3.1 A
Revised Planning: none

Step 1b) Stereo

Results:

1) A local multi modal descriptor that constitutes our visual Primitives has been implemented. The used Primitives have been motivated by feature processing in the human visual system and functional aspects of information processing.

2) Stereo-primitives were defined on the basis of corresponding Primitives in the left and right image making use of the modalities orientation, contrast transition and color. Optimal weights for the relative importance of these sub-modalities for stereo processing have been computed.

3) Next we have performed extensive statistical measurements on real images and found that *only* collinear (and parallel) line segments are statistical significant primitives in images. The central result of this study is that: If line-segments are collinear then they are with great likelihood also sharing other visual characteristics (such as color, contrast, motion and disparity!). This way an objective statistical measurement called the "Gestalt coefficient" has been defined that allows for defining and actually finding primitives in images without making arbitrary heuristic assumptions.

4) Building on this simple and robust measure we were able to define a grouping process that leads to 2D-Gestalts and Stereo-Gestalts.

5) The Primitives are adaptable according to spatial context once they are part of a 2D or Stereo-Gestalt.

Documentation: Computer Programs & Documentation (MoIns, see: http://www.cn.stir.ac.uk/ImageAnalysis/MoInS/index.html)

Publications: The following papers correspond to the above listed 5 steps in the same order: (Sco.1), (Sco.2), (Sco.1), (Sco.3), (Sco.4). Additional Publications are: Sco.5 – Sco.7.

Deliverables: D3.1 B

Task WP3.2) Visual templates with spatio-temporal context - Gestalts in space-time

Scheduling: about 2-4 months ahead of planning

Performed actions: work by Sco

Planned and performed Steps:

Step 1) Track the development of these receptive fields (Gestalts) over time in order to find a mathematical (algorithmical) way to describe their temporal development.

Scheduling: ahead of time by 2-4 months

Results: The use of the Gestalt-coefficient allowed us to define stereo-Gestalts while the Kalman approach renders motion-Gestalts. The development of both can be tracked in image sequences. In order to do this the rigid body motion (RBM) constraint has been applied to these Gestalts. This provides a robust algorithm, because Gestalts are basically stable. A tracking algorithm has been implemented based on RBM.

Documentation: ---

Publications: Sco.4

Revised Planning: This task is ahead of time, therefore we could already start to address step 2 of this task (see next).

Step 2) Use the tracking to alter the Gestalts in a temporally adaptive way.

Scheduling: started about 2 month ago because of the good progress of step 1 above, originally there was no work planned for year 1

Performed actions: work by Sco

Results: The RBM-based tracking mechanism is structured such that it is not a pure feed-forward analysis (i.e. tracking). The algorithm has already in its first instances been implemented such that a recurrent process is defined which on the one hand does the tracking of the Gestalts and on the other hand continuously adapts their shape to the changing geometry of their projections through feature correction mechanisms, which improve the structure and the inner stability of the Gestalts and by also eliminating outliers. First results have been obtained here.
Documentation: none
Publications: Sco.4

Deliverables: not due

General revised planning: Partner Ita needs now to be involved in order to employ the motion primitives in the right way in the adaptive tracking algorithm. So far this has been done rather ad hoc.

Task WP3.3) Cross-modal interactions between the Gestalts define the first level of recursive scene analysis

Scheduling: no work planned for year 1
Revised Planning: WP3.2 and WP3.3 were found to blend into each other, because some results from 3.2 where RBM principles are used to correct image features show that the stereo and motion domain directly and recursively interact with each other. Thus, the treatment of temporally adapting Gestalts has already by now led to a cross-model interaction between Stereo and Motion. Therefore, we expect that the separation between 3.2 and 3.3 will in year 2 have to be dropped and both WPs will be treated in a more unified Ansatz.
Publications: Sco.4

Task WP3.4) Higher level visual segments

Scheduling: no work planned for year 1

WP4: Space variant representation of weighted filters in a map

This WP is concerned with the design of low-level adaptable filters (mainly for motion analysis) which can be implemented in a map-like fashion into the ECOVISION system. At the same time these filters are supposed to be hardware compatible.

Scheduling: about 2-3 months ahead of planning (Deliverable 4.1, due month 12, is completed apart from revisions which will be performed in the remaining 3 months.
Central achievements: A set of scalable/steerable filters has been designed as a DLL software package which can now (as planned) be implemented into the ECOVISION software package in year 2.

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

Scheduling: about 3 months ahead of planning
Performed actions: A meeting was held in Granada between Eng, Spa and Ger/Mun to discuss the optic flow algorithms. Guillermo from Spa is currently working at UCL with Eng evaluating the optic flow algorithm in the context of WP1 and WP4.

Planned and performed Steps:

Step 1) Design of adaptive Gaussian derivative filters.

Results: Convolution routines for accurate separable Gaussian derivatives generated using Hermite polynomials and also faster numerically differentiated Gaussian filters.

Step 2) Dynamically adapt the scale and preferred orientation of a filter at a given location (general method)

Results: Eng has developed a set of DLLs where low-level visual filters can be controlled by visual context. Filters were designed using a "basis" set of orthogonal X-Y separable Gaussian derivative filter responses. The spatial orientation of the individual filters in the basis set can be modified via a linear recombination of the responses. The scale can also be altered using a weighted recombination, although the range is relatively small. This allows combining such basic filters in a robust, computationally efficient and task-controllable way.

Step 3) Dynamically adapt the local scale of the filtering operations to the demands of the task (task dependency)

Results: An adaptive weighting system in the Taylor expansion representation enables us to combine the responses from filters at different scales in the motion algorithm. This part of the work is planned for the remaining two months of deliverable 4.1.

Documentation: Program (DLL) Description and Documentation
Publications: Eng.1, Eng.2
Deliverables: Software as DLLs for the implementation into the ECOVISION System & Documentation
Revised Planning: finished, but will be revised in the remainder of time.

Task WP4.2) General space variant mapping methods
Scheduling: no work planned in year 1.

WP5: The representation of visual relevance (early cognition)

WP5 is concerned with the question: What is visually relevant during driving? And how can this be used to define a Center of Interest (CoI) for a machine vision system (in the sense that visual analysis should be most dense at the CoI)? Note: visual relevance is task dependent and so is the CoI. Visual attention is not necessarily task dependent, thus the focus of attention is often different from the CoI!

Scheduling: slightly delayed because of the move of Ger to Mun. Basic results have been obtained in Bochum and can be used as planned in the other WPs, such that no delay to the project as such will occur. However, WP5.1 will be extended until July 2003 (overlapping with 5.2) in order to gather more data and analyse it in more depth.

Central achievements: The analysis of real and simulated driving scenes has lead to three novel and surprising findings (see below).

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

Scheduling: slightly delayed because of the move of Ger to Mun. Basic results are available and will be described in deliverable 5.1 by Dec. 2002. WP5.1 will be extended until July 2003 (overlapping with 5.2) in order to gather more data and analyse it in more depth.

Performed actions: Eye-movement measurement based on artificial images (Ger/Mun) and on driving scene obtained in co-operation with Ind (and Sco).

Planned and performed Steps:

Step 1) Find out which of image parts will attract the viewer's interest.

Results: Real and artificial driving scenes were psychophysically analyzed during two tasks: heading and obstacle avoidance. We found that (a) subjects will find the focus of expansion only after several saccades. (b) Obstacles were, on the other hand, detected immediately even (c) during changing heading direction and, during obstacle avoidance; the focus of expansion was never fixated.

Step 2) How can this viewing behavior be parameterized

Results: As the results show, several parameters influence gaze behavior during driving. Most important is the momentary task. Gaze behavior is very different for heading tracking than for obstacle detection. Since the obstacle detection task also involved the estimation of the changes in heading, it is probably safe to assume that eye movements towards the FOE are not of high relevance and that the obstacle avoidance task is a better reflection of the normal driving situation. The parameters that affect eye movements in the obstacle avoidance task are the distance of an obstacle from the direction of heading and from the observer. If obstacles close to heading appear, eye movements to them are imperative and very quick. If no obstacles are near the heading direction eye movements scan the scene more widely and return to future path from time to time. A further parameter in this situation is the "default" distance at which the eyes look when returning to the path. This differs depending on the subject.

Documentation: preliminary report

Publications: Ger.1 – Ger.3

Deliverables: due 31.12.02, will be completed in time

Revised Planning: The work package will be extended by 6 months in order to gather additional data and prepare it for scientific publication.

Effects on other WPs: Essentially the other WPs will not be affected because the basic data has already been obtained by Ger/Mun which can be used to define the CoI and the task dependence. This will be done in deliverable 5.1. The additional 6 months allocated to WP5.1 are required to perform more experiments in order to consolidate the existing results.

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

Scheduling: only first step planned for year 1

Documentation: none

Publications: none

Deliverables: not due

Revised Planning: An important new component in this WP is concerned with the *data driven segmentation of independently moving objects*. We have decided to implement this, because it is otherwise not possible to define motion-Gestalts in an objective way. This affects also WP3 above, where we currently use

structural assumptions to define the motion Gestalt. The goal of the revision of WP5.2 is to devised an *objective method* to define such motion Gestalts (for WP3) and from them also the CoI (for WP5). As a first step, based on previous research by partner Ger [Lappe, 1993], we will develop a robust and novel method for the computation of ego-heading [Heeger, 1990]. This algorithm will enable the determination of the focus of expansion in a way that is insensitive to distortions in the flow fields caused by flow estimation algorithms and independently moving objects. The method will serve as a Gestalt-steering mechanism in heading tasks (WP 5.2).
In a second step, we will use these robust ego-heading parameters to evaluate, at each flow vector location, the null hypothesis that the scene is stationary, i.e., that there are no independently moving objects. By integrating information over time, the robustness of this procedure will be ensured. The identification of moving objects is important in WP 5.2 in the location of the center of interest (CoI) in tasks such as fixating, tracking and time-to-impact estimation. This information also serves as input to attentional mechanisms, developed in WP 6.
The end results of both phases are motion Gestalts, which are an additional mode that will be used in the multimodal Gestalt merging process.

Partner Bel will be mainly involved in this. To this end some more man-month will be allocated to this new sub-aspect of WP5.2

Heeger, D.J. & Jepson, A. (1990). Visual Perception of Three-Dimensional Motion, *Neural Computation*, **2**, 129-137.
Lappe, M. & Rauschecker, J.P. (1993). A Neural Network for the Processing of Optic Flow from Ego-Motion in Man and Higher Mammals, *Neural Computation*, **5**, 374-391.

Task WP5.3) Implementing early cognitive control in a technical environment:
Scheduling: no work planned for year 1

Project Part 3: Computational Neuroscience (WP6)

This project part is concerned with modeling of attention effects in layer IVc of monkey striate cortex. In those neurons attention leads to a restructuring of their receptive fields, probably as a top-down (feedback) effect arising from higher cortical areas (e.g., area MT). This study is motivated in the context of ECOVISION because it provides novel ideas about of how to restructure (Gestalt-)filters by means of recurrent interactions. WP5 in ECOVISION is concerned with “visual relevance” and the (task dependent) “center of interest” which are supposed to be implemented in WP3 (3.2). WP6 addresses the more basic aspect of visual attention. However, WP6 specifically investigates the differences between spatial and featural attention. Featural attention can be interpreted as attention arising from a task. Thus, this is conceptually very close if not identical to defining a center of interest. WP6 has by now already produced first results which indicate that similar feedback mechanisms and structures can be employed for altering receptive fields (i.e. filters) as a consequence of featural visual attention. Bel has started to talk to Sco in the context of WP3.2 in order to convey these ideas.

WP6: Contextual effects in human and non human primate visual system.

Scheduling: in time
Central achievements: This WP has by now successfully addressed how a temporary changing lateral inhibitory mechanism can be generated that restructures receptive fields in an attention dependent way. Spatial attention has been compared to featural (e.g. task dependent) attention. Most details of the algorithm are still under investigation. It is planned to try using this algorithm for implementing the Center of Interest defined in WP5 in order to define higher order Gestalts in WP3.

Task WP6.1) ***Task-optimized spatial representations – spatial attention case:***

Planned and performed Steps:
Step 1) Algorithmical model of cortical suppression

> Scheduling: in time
> Performed actions: Modeling work by Bel.
> Results: The algorithm is based on the evaluation of 2-DG data taken form a paradigm with spatial and comparing it to one with featural attention. A suppressive mechanism has been found. This has led to the assumption that the suppression is achieved by a centre-surround-like modification of the intralayer connection pattern, which is now modeled in step 2.

Step 2) Implementation of a fuzzy membership function for distinguishing between spatial and featural attention

> Scheduling: in time
> Performed actions: Modeling work by Bel.
> Results: Membership to the two discussed cases is defined through a temporary modification of the centre-surround ranges. This is done with the help of a novel learning rule for recurrent neural networks. The interesting aspect of this implementation is that it will dynamically move from a spatial attention representation to a featural one.

Step 3) Matching the membership mechanism (i.e. function) to visual relevance

> Scheduling: no work planned for year 1

Task WP6.2) Task-optimized representations - general case
No work planned for year 1

Documentation: none
Publications: Bel.1-Bel.7
Deliverables: none
Revised Planning: Further 2DG experiments and thus data relating to early-cognitive attentional mechanisms are not to be expected during the lifespan of ECOVISION. Thus, a more engineering-oriented approach will be adopted in WP 6.2. Specifically, the interactions between attention and context will be further investigated. Classically, both mechanisms are viewed separately, in an attentive and a pre-attentive stage for context. However, advances made since the start of ECOVISION [Raizada and Grossberg, 2001] show that attention and context operate simultaneously, complementing one another [see also Chun, 2000]. We will also adopt the view that attention and context interact simultaneously but we will involve lateral connectivity to be in line with the suppression observed in the 2DG results considered in WP6.1.

Chun, M. M. (2000). Contextual Cueing of Visual Attention. *Trends in Cognitive Sciences*, **4**(5), 170-178.
Raizada, R. and Grossberg, S. (2001). Context-Sensitive Bindings by the Laminar Circuits of V1 and V2: A Unified Model of Perceptual Grouping, Attention, and Orientation Contrast. *Visual Cognition*, **8**(3/4/5), 341-466.

Project Part 4: Technical Workparts

This project part is concerned with the application side of ECOVISION. How can the system be implemented in a driver assistant system (WP7)? WP8 and WP9 address the question of how to put the different components together (planned for year 3 mainly).

WP7 Input/Output Specifications

Scheduling: finished in time
Central achievements: In this WP we have addressed the industrial specification necessary for driver assistant systems which would use ECOVISION technology. Two industry-standard product specifications have been compiled. Furthermore, a set of camera sequences has been recorded as a test-bed for the different parts of ECOVISION.

Planned and performed Steps:
Step 1) Recording of camera sequences

Scheduling: finished in month 6
Performed actions: Visit of Sco to Ind and co-operative work in the week of June 10-14.
Results: Ca. 50 GB of calibrated high resolution stereo camera sequences of different driving scenes. Image resolution was 1000x1400 pixels at 25 Hz.
Documentation: Documented data files for the use of all partners.
Deliverables: 7.1
Publications: Not applicable
Revised Planning: We are planning to record some more sequences with specific sceneries in year 2.

Step 2) Industry-standard product specifications

Scheduling: finished in month 9
Performed actions: work by Ind (consultations with Spa and the other partners)
Results: Two specifications were written: One for a pedestrian-recognition-system and one for a lane-change-assistance-system (rear-view mirror, blind spot problem.)
Documentation: Two industrial specification documents.
Publications: Not applicable
Deliverables: 7.2
Revised Planning: Specifications may be revised according to the results of ECOVISION in the course of the project.

WP8 Benchmarking and Testing
Scheduling: no work planned for year 1

WP9 Complete ECOVISION System (Joining the components)
Scheduling: no work planned for year 1

Project Part 5: Administrative Workparts

WP10 Presentation, Dissemination and Use

Task 10.1 Web page
Scheduling: ongoing
The web page is at http://www.pspc.dibe.unige.it/~ECOVISION/ and it is continuously updated. It contains the latest public results and a private area for the members of ECOVISION only. The password which will allow the referees access to this section is:
username: fusion, password: h2oy2k
Deliverable: D10.1 ongoing

Task: 10.2. Dissemination & Use Plan
Scheduling: finished
The DUP has been written in month 6.
Deliverable: 10.2 DUP. Can be downloaded from our web-pages
(http://www.pspc.dibe.unige.it/ECOVISION/private/index.html).

Task D10.3 Scientific Publications
Scheduling: ongoing
Publications are listed at the individual Progress Report Sheet of each partner.
We have by now published 17 conference publications and 7 journal publications.
Deliverable: 10.3, ongoing

Task D10.4 Workshop at Stirling.
Scheduling: in time
We are in the process of planning a workshop for May 2004 in Scotland. A conference site at the Isle of Skye has been contacted. A preliminary program has been compiled. Speakers will be invited in spring 2003.
Deliverable 10.4. not due

WP11 Management and Evaluation

Scheduling: ongoing
Administrative aspects: Several smaller budget shifts have been made and approved by the coordinator. Two major contract amendments have been performed. Partner Ger/Mun has moved from Bochum University to Münster taking the position of a full professor there. The University of Münster is now our new contract partner since June. Partner Bel has found other sources for durable equipment and for the budget planned for auditing. As a consequence, Bel has shifted this budget into Travel and Personnel. In all other respects the administration of the ECOVISION project proceeds without problems.

Co-operations and Meetings: See pg. 5 of this report.
Publications: See Progress report sheets of each partner (pg. 51-63).

Dissemination and Use: This project is centered on the scientific aspect. Thus, dissemination will take place mainly at conferences and by means of publications.
The following conferences will be addressed (those marked with an * have already received a submission from the ECOVISION project). Conference venues are listed only for 2002 and only for those conferences where papers have already been submitted.
*DAGM (Zurich 2002, Workshop on “Cognitive Vision”)

*ECVP (Glasgow 2002)
*NIPS (Neural information processing systems, USA)
*IEEE Workshop: Neural Networks for Signal Processing (Martigny, Switzerland, 2002).
*JCRA'2002 (Granada, Spain)
*Workshop on Dynamic Perception, Bochum 2002
*2nd Workshop on Biologically Motivated Computer Vision, Tübingen,
Annual Neuroscience Meeting (USA, for the neuro-aspects of ECOVISION)
IEEE Workshop on Neural Networks for Signal Processing
Human Brain Mapping
Vision Sciences Society meeting
WS on “Early Cognitive Vision”, Stirling, 2004 (Organized by ECOVISION)
Europ. Sci. Conf. “Three-dimensional Sensory and Motor Space”, Italy

The following journals will be addressed (those marked with an * have already received a submission from the ECOVISION project).
*Neural Networks
*J. VLSI Signal Processing Systems for Signal, Image, and Video Technology
*Neural Computation
*Scientometrics,
*IEEE Trans. Neural Networks
*Vision Research
*Int. Journal of Computer Vision
*Journal of Physiology (Paris)
Biological Cybernetics
Neuroimage
PhysicaD
IEEE Trans. on Biomedical Image Processing,
Journal of Neuroscience
Proceedings Royal Society London B

Cost Breakdown Summary

As a consequence of the pending amendments and in agreement with the Scientific Project Officer of the Commission (Mr. A. Bakalakos), we will provide the Cost Breakdown after completion of the full first year in January 2003.Therefore no cost calculations and no man-month tables are provided in this PPR.

World-wide State of the Art Update

WP1,2 Motion estimation and FPGA

Motion estimation focal-plane solutions have been proposed by different authors using analog VLSI to combine the photo-sensors and the motion detection processing circuits within the same chip (Kramer et al 1997, Etienne-Cummings et al 1997, Higgins et al 1999). These motion sensors are useful for robot vision and similar applications. On one hand, they have the weakness of any analog circuit (circuit parameter variations due to the fabrication process, aging, temperature variations, etc), but on the other hand they are approaches that work with low-power and in real-time. Because of having the processing elements and photo-sensors built together, it is difficult to estimate the accuracy of these vision chips with textures and in different scenarios and compare them with more "classical" optic flow algorithms (Barron et al, 1994).

A more flexible approach consisting of a front-end sensor performing no pre-processing task and a digital circuit implemented in a FPGA makes it possible to implement different optic flow algorithms taking into account the accuracy vs efficiency trade-off (Liu et al 1998). The main advantage of digital computing is that the precision in the different stages is controllable and different devices will exhibit exactly the same behavior (which is not the case when computing with analog circuits). Digital implementation of motion estimation algorithms has different application fields, such as video compression (Ramachandran and Srinivasan 2001), artificial vision, etc. Different authors have addressed this implementation issue in the last years (Ramachandran and Srinivasan 2001, Boluda and Domingo 2001, Zuloaga et al 1998, Cobos and Monasterio-Huelin 1998, 2001). But the complexity required by the "classical" optic flow algorithms (Liu et al. 1998) has not encouraged the direct implementation of these models onto specific hardware.

WP3-5 Gestalt in Space and Time and the Problem of Visual Relevance

Statistics of Images:
Decades ago, Brunswick and Kamiya (1953) first had stated that Gestalt principles should be related to the statistics of the natural world. Unfortunately the limited computational power at this time made it difficult to quantitatively support this statement. The strong prevalence of collinearity (and parallelism) in natural images has been investigated first by Krüger (1998) and Elder & Goldberg (1998). These results have been confirmed and extended by Sigman et al (2001) and Geisler et al (2001). In addition to those first observations, in Sigman et al (2001) a co-circularity rule has been established as a generalization of collinearity, which says that after straight lines, circular structures form the most common second order relations of line segments in natural images. While Krüger (1998) and Sigman et al (2001) have investigated the second order relation of line segment responses without considering whether the two line segments belong to the same 3D contour or not (called 'absolute co--occurrence' in Geisler et al 2001). Geisler et al (2001) and Elder & Goldberg (1998) have investigated conditional densities which take this contour coincidence into account (so called 'Bayesian co--occurence'). This distinction is especially important since only in the case of a collinearity event caused by a 3D contour the grouping of entities is justified. However, Geisler et al (2001) have shown that in both cases the second order statistics are similar which indicates that collinearity is mostly determined by projections of 'real' 3D contours. In addition to collinearity and parallelism, Elder and Goldberg (1998) show

that other Gestalt cues (proximity and luminance similarity) can also be related to the statistics of natural images.

Motion and Stereo:
Our approach has some interesting properties compared to other systems. Firstly, different from the classical structure of motion approaches (see, e.g., Klette et al (1998), Faugeras (1993) and Hartley& Zisserman (2000) we do not intend to acquire 3D—information only, but we are interested in attributes that are relevant for perceptive tasks. Our representations therefore consist of multiple modalities. Note that our stereo algorithm does not make use of the ordering constraint (such as, e.g., in Pollefeys et al, 2000) but that the system can also be used in case of depth continuities. The only restriction is the occurrence of local line segments in the scene. In our representations, semantic properties of features and their reliability are explicitly coded. Both, semantic properties and reliability, are subject to contextual influences. The integration of contextual information and its modeling by recurrent processes that modify reliabilities is the central aim of our current project and our method differs in that respect to classical structure-from-motion algorithms (see, e.g., Hartley & Zisserman 2000). We are using reliability information to improve RBM estimation by picking out certain 'good' feature constellations from the big feature pool. This way handling outliers works complementary to other well established methods such as RANSAC (Fischler & Bolles, 1981). Coding information with its reliability allows keeping hypotheses that are unlikely (when looking at them locally) but may become likely taking the context into account.

Eye-movements:
With respect to eye movements during driving, a study on the gaze behavior of a race-car driver has shown that eye movement strategies not only depend on task demands but also on experience or learning level. Rather than fixating the tangent point, as normal drivers do during curve negotiation, this race car driver directs gaze to points on the curve that are specific (and repeatable) for each curve on the track, showing that he has established a reliable and consistent gaze strategy that is adapted to the track (Land and Tatler 2001). Our own results from the gaze recordings suggest that heading can be estimated even when gaze (and presumably attention) is directed to objects on the scene in front of the driver. Wann et al. (2000) have investigated heading perception under various conditions of attention allocation. They find, consistent with our results, that heading can be estimated reliably even when a secondary task requires attentional resources, provided that extraretinal information about ongoing eye movements is available. For the developments in the project this suggests that we should treat the heading task as a pre-attentive routine.

WP6 Modeling of attentional processing

There is ample experimental evidence for the effect of cognitive context and spatial attention on visual processing (Desimone & Duncan, 1995; Colby & Goldberg, 1999; Sengpiel & Huebner, 1999; Vanduffel *et al.*, 2000; for a recent review, see Chun, 2000). Contextual information embodies important properties of the visual environment and serves to steer attentional mechanisms. Since the start of ECOVISION, a possible mechanism for the interaction between attention and context was demonstrated in a recent paper by Raizada & Grossberg (2001). In this model, an on-center, off-surround attentional input modifies the representation in V1. At the same time, this representation is being altered by contextual mechanisms. A model for the roles of and interactions between the dorsal and ventral streams in relation to attention are discussed in a recent paper by Van der Velde and co-workers (Van der Velde et al., 2001). An alternative model that also addresses the relations between spatial and featural attention has been developed by Deco (2001). The model is closely related to the biased competition paradigm in which attention biases the competition in favor of objects containing the attended features or situated at the attended location. Feedback from extrastriate modules dynamically adapts the strength of the activities in the early visual cortex. These feedback connection weights are determined by Hebbian learning in connection with a shifting focus of attention but this takes far too much computational time to be a realistic model.

References

Barron, J.L.; Fleet, D.J.;Beauchemin, S.S.: “Performance of Optical Flow Techniques”, *International Journal of Computer Vision*, Vol. 12, No. 1, pp. 43-77, 1994.

Boluda, J.A. and Domingo, J.J.; “On the advantages of combining differential algorithms, pipelined architectures and log-polar vision for detection of self-motion form a mobile robot”, *Robotics and Autonomous Systems*, Vol. 37, No. 4, pp. 283-296, December 2001.

Bradley, D.C., & Andersen, R.A. (1998). Center-surround antagonism based on disparity in primate area MT. *J. Neurosci*, 18, 7552-7565.

Brunswik, E. and Kamiya J. (1953). Ecological cue--validity of 'proximity' and of other Gestalt factors. *American Journal of Psychology*. LXVI. p:20--32.

Chun, M. M. (2000). Contextual Cueing of Visual Attention. *Trends in Cognitive Sciences*, 4(5), 170-178.

Cobos Arribas, P.; Monasterio-Huelin Maciá, F.: “FPGA implementation of the Horn & Shunk Optical Flow Algorithm for motion detection in real time images”, DCIS’98 *Proceedings, Design of Circuits and Integrated Systems Conference*, pp. 616-621, Spain, 1998.

Cobos Arribas, P.; Monasterio-Huelin Maciá, F.: “FPGA implementation of Camus correlation optical flow algorithm for real time images”, DCIS’01 *Proceedings, Design of Circuits and Integrated Systems Conference*, Spain, 2001.

Colby C.L, Goldberg M.E. (1999) Space and Attention in Parietal Cortex. *Annu. Rev. Neurosci*. 22:319-349.

Deco, G. (2001). Biased Competition Mechanisms for Visual Attention in a Multimodular Neurodynamical System. *Emergent Neural Computational Architectures Based on Neuroscience*, 114-126.

Desimone R., Duncan J. (1995) Neural Mechanisms of Selective Visual Attention. *Annu. Rev. Neurosci*. 18:193-222.

Elder H. and Goldberg R.M. (1998). Inferential reliability of contour grouping cues in natural images. *Perception Supplement* 27.

Etienne-Cummings, R.; Van der Spiegel, J.; Mueller, P.: “A focal plane visual motion measurement sensor”, *IEEE Transactions on Circuits and Systems I*, Vol. 44, No. 1, pp. 55-66, 1997.

Fischler R. and Bolles M. (1981). Random Sample Consensus: A paradigm for model fitting with applications to Image Analysis and Automated Cartography. *Communications of the ACM*. 24(6), p:619--638

Geisler W.S. and Perry J.S. Super B.J. and Gallogly D.G. (2001). Edge co--occurrence in natural images predicts contour grouping performance. *Vision Research*. 41, p:711--724.

Hartley R.I. and Zisserman A. (2000). Multiple View Geometry in Computer Vision. Cambridge University Press.

Higgins, C.M.; Deutschmann, R.A.; Koch, C.:”Pulse-based 2D motion sensors”, *IEEE Transactions on Circuits and Systems II*. Vol. 46, No. 6, pp. 677-687, June 1999.

Klette R., Schlüns K. and Koschan A. (1998). Computer Vision - Three-Dimensional Data from Images. Springer.

Kramer, J.; Sarpeshkar, R.; Koch, C.: “Pulse-based Analog VLSI Velocity sensors”, IEEE Transactions on Circuits and Systems II. *Analog and Digital Signal Processing*. Vol. 44, No. 2, pp. 86-101, February, 1997.

Krüger N. (1998). Collinearity and Parallelism are Statistically Significant Second Order Relations of Complex Cell Responses. *Neural Processing Letters* 8(22). p:117--129.

Land MF, Tatler BW (2001) Steering with the head: The visual strategy of a racing driver *Current Biology* 11 (15): 1215-1220.

Liu, H.; Hong, T.H.; Herman, M.; Camus, T.; Chellappa, R: “Accuracy vs. Efficiency Trade-offs in Optic-Flow Algorithms”, *Computer Vision and Image Understanding*, No. 72, pp. 271-286, December 1998.

Pollefeys M., Koch R. and van Gool, L. (2000). Automated Reconstruction of 3D scenes from sequences of images. *Journal of Photogrammetry and Remote Sensing*. 55(4), p:251--267.

Raizada, R. and Grossberg, S. (2001). Context-Sensitive Bindings by the Laminar Circuits of V1 and V2: A Unified Model of Perceptual Grouping, Attention, and Orientation Contrast. *Visual Cognition*, 8(3/4/5), 341-466.

Ramachadran, S.; Srimivasan, S.: “FPGA implementation of a novel, fast motion estimation algorithm for real-time video compression”, *Proceedings of the Ninth International Symposium on Field Programmable Gate Arrays* (FPGA-2001), pp. 213-219, Monterey, California, USA, 2001.

Sengpiel F., Huebner M. (1999) Visual Attention: Spotlight on the Primary Visual Cortex. *Current Biology* 9:R318-R321.

Sigman, M. and Cecchi, G.A., Gilbert, C.D. and Magnasco M.O. (2001). On a common circle: Natural Scenes aand Gestalt Rules. *PNAS*. 98(4), p:1935--1949.

Van der Velde, F. & Kamps, M. (2001). From Knowing What to Knowing Where: Modeling Object-Based Attention with Feedback Disinhibition of Activation. *Journal of Cognitive Neuroscience*, 14(4), 479-491.

Vanduffel W., Tootell R.B.H., Orban G.A. (2000) Attention-dependent Suppression of Metabolic Activity in the Early Stages of the Macaque Visual System. *Cereb. Cortex*, 10: 109-126.

Wann JP, Swapp D, Rushton SK (2000) Heading perception and the allocation of attention *Vision Res*. 40 (18): 2533-2543.

Zuloaga, A.; Martin, J.; Ezquerra, J.: “Hardware Architecture for optical flow estimation in real time”, pp. 97-976, *IPIC-C’98*.

Work Progress Overview

This part of the document contains the individual descriptions of the work done for the different workpackages. It contains more technical details than the executive summary above and is meant to provide deeper insight into the progress of the ECOVISION project.

WP1

The goal of this WP is to find the best motion/flow –field algorithm in terms of its VLSI implementability. Several members of the consortium have designed such algorithms and others exist in the literature (e.g.; Barron et al. 1994).

As planned, we have studied the FPGA implementation viability of Eng's algorithm (Johnston et al 1992). A first assessment of its implementability has been done, but WP1 is not yet finished, a complete study with different algorithm configurations (image resolution, filter sizes, bit accuracy, etc.) is to be performed in order to evaluate how all these factors affect the algorithm accuracy.

The Multi channel Gradient Model (McGM) computes optic flow using a gradient technique. As such, it makes use of many spatio-temporal derivatives of the image sequence, obtained through many image filtering operations. The model extends the basic motion constraint equation by expressing the image sequence as a Taylor expansion, thus introducing higher order derivatives into the image description. Use of higher order derivatives in a Taylor expansion framework has been shown to add robustness to optical flow calculations and provides a convenient method for analyzing image structure contained within a small region of an image. The ratio of the spatial to temporal derivative of any one of the terms in the Taylor expansion can be used to compute image speed, although each individual ratio will be ill conditioned by itself. The McGM combines the multiple derivative terms of the Taylor expansion and pools over a space-time volume to arrive at a much more robust measurement of optical flow. In addition, the Taylor expansion is extracted in a number of directions that correspond to the orientation columns found in the primate visual cortex, providing a measurement of speed as a function of filter orientation. This multiple-measurement strategy helps reduce noise and allows the translational component to be separated from differential components of optic flow.

In this way the McGM gives a very robust optic flow estimation that can be used for real world open-air sequences. The computational cost of this estimation robustness is high, therefore we will evaluate later in the project if this robustness level is necessary with the driven sequences provided by the Ind partner.

We address the objective in two stages:

Task 1.1: FPGA implementability assessment of the McGM algorithm.

In order to study the hardware implementability of the model, it has been divided in three stages well differentiated by the computational primitives on whom they are based:

1. Spatial and temporal filter. Computational function: Convolution.
2. Computation of the speed and orthogonal speed. Computational function:

arithmetic operations.

3. Final estimation of the speed. Computational function: matrix determinants.

We have evaluated the three main factors that need to be taken into account when studying its hardware implementability: (1) computational speed, (2) memory requirements and (3) bandwidth of buses between the processing circuits and memory blocks.

The computation resources required by the different stages have been evaluated making several assumptions based on the previously acquired experience with the model, such as image size, minimum number of orientations, number of spatial filters, etc. Specific purpose resources present in current FPGA devices that will be required are: (1) embedded memory blocks, (2) embedded multipliers and (3) embedded processors. Some high performance FPGA devises have been pointed out as possible final candidates for this implementation because of their internal specific purpose blocks.
Finally, a first block processing scheme has been proposed that uses different levels of parallelisms along the whole process.

Task 1.2: McGM characterization: configuration study.

The Multi-Channel Gradient Model has been implemented in a high level language (MATLAB) in order to easily study the different accuracy levels, required resolution, size of the oriented filters, and other configuration details. One of the ECOVISION contracted personnel of the group of the University of Granada is currently visiting the UCL (16 weeks duration: 01/10/2002 till 28/02/2003) to gain experience with the motion model in order to study how different simplifications or modifications of the model that facilitate its implementation may affect its functional performance.

As contingency plan a second motion algorithm is also being studied to asses its VLSI implementation feasibility. This is the Simoncelli and Heeger model (Simoncelli and Heeger 1998).

The Simoncelli and Heeger's algorithm is a very bio-inspired processing scheme that consists in two stages of computation corresponding to neurons in cortical areas V1 and MT. Both V1 and MT neurons compute a linear combination of their inputs. The linear weighting of a V1 neuron determines the selectivity for stimulus position and orientation and also for stimulus direction. An MT neuron computes a linear sum of V1 responses, such that it becomes selective for stimulus velocity in a particular direction and at a particular speed. This algorithm has been proposed as "A Model of Neuronal Responses in Visual Area MT".

This very regular computational structure can be implemented as a highly parallel data-path, but there are two main drawbacks:

1. Related with the algorithm performance and hardware implementation: S&H algorithm performance is on average in line with respect to other "classical" algorithms, but it is much more computational demanding, due to its choice of adopting a population coding. More concretely, the Simoncelli and Heeger's algorithm does not explicitly compute single velocities; rather, they can be inferred from the population activity of a set of cortical-like cells tuned to different velocities.

The precision of the "measured" velocity would strictly depend on the size of the population set that span the velocity space. To achieve a high precision it would imply for each pixel a large number of filters. Detailed considerations on the required precision would be needed to asses the implementation viability (Liu et al 1998). Some simulations are being done to evaluate the performance/hardware-requirements trade-off, compared to McGM algorithm.

2. Related with real world application: This algorithm is contrast-dependent, which is an important drawback for real world open-air sequences.

Other authors have addressed the implementation on FPGAs of other optic flow algortihms (Zuloaga et al. 1998, Cobos and Monasterio-Huelin 1998, 2001) but they have used very simple devices with limited resources and they have not reported any performance results obtained from these implementations.

WP2

This workpackage depends on the results and conclusions obtained in WP1, which is not yet finished.

Currently the group of the University of Granada (Spa) is testing different Hardware description languages (Handel-C, VHDL, etc) and tools such as the System Generator (of Mathworks and Xilinx). These promising higher level description tools may change the way complex systems are designed in the next years. These higher level description methodologies can reduce the final implementation time.

Now, at the month 10 of the project, WP2 is in a very preliminary stage (it began at month 6 of the project), mainly focussed on evaluating different description tools and environments, and setting up the prototyping platforms.

References:

[1] Johnston, A., McOwan, P. W. & Buxton, H.: "A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells", Proc. R. Soc. Lond. B., 250, pp. 297-306, 1992.
[2] Simoncelli, E. P. Heeger, D. J.: "A Model of Neuronal Responses in Visual Area MT". Vision Reseach, Vol. 38, No. 5, pp. 743-761, 1998.
[3] Barron J. L., Fleet D. J., Beauchemin S. S., "Performance of Optical Flow Techniques", *International Journal of Computer Vision* **12**(1), pp.43-77 (1994)
[4] Liu, H.; Hong, T.H.; Herman, M.; Camus, T.; Chellappa, R.: "Accuracy vs. Efficiency Trade-offs in Optic-Flow Algorithms", Computer Vision and Image Understanding, No. 72, pp. 271-286, December 1998.
[5] Zuloaga, A.; Martin, J.; Ezquerra, J.: "Hardware Architecture for optical flow estimation in real time", pp. 97-976, IPIC-C'98.
[6] Cobos A.P.; Monasterio-Huelin M. F.: "FPGA implementation of the Horn & Shunk Optical Flow Algorithm for motion detection in real time images", DCIS'98 Proceedings, Design of Circuits and Integrated Systems Conference, pp. 616-621, Spain, 1998.
[7] Cobos A.P.; Monasterio-Huelin M.F.: "FPGA implementation of Camus correlation optical flow algorithm for real time images", DCIS'01 Proceedings, Design of Circuits and Integrated Systems Conference, Spain, 2001.

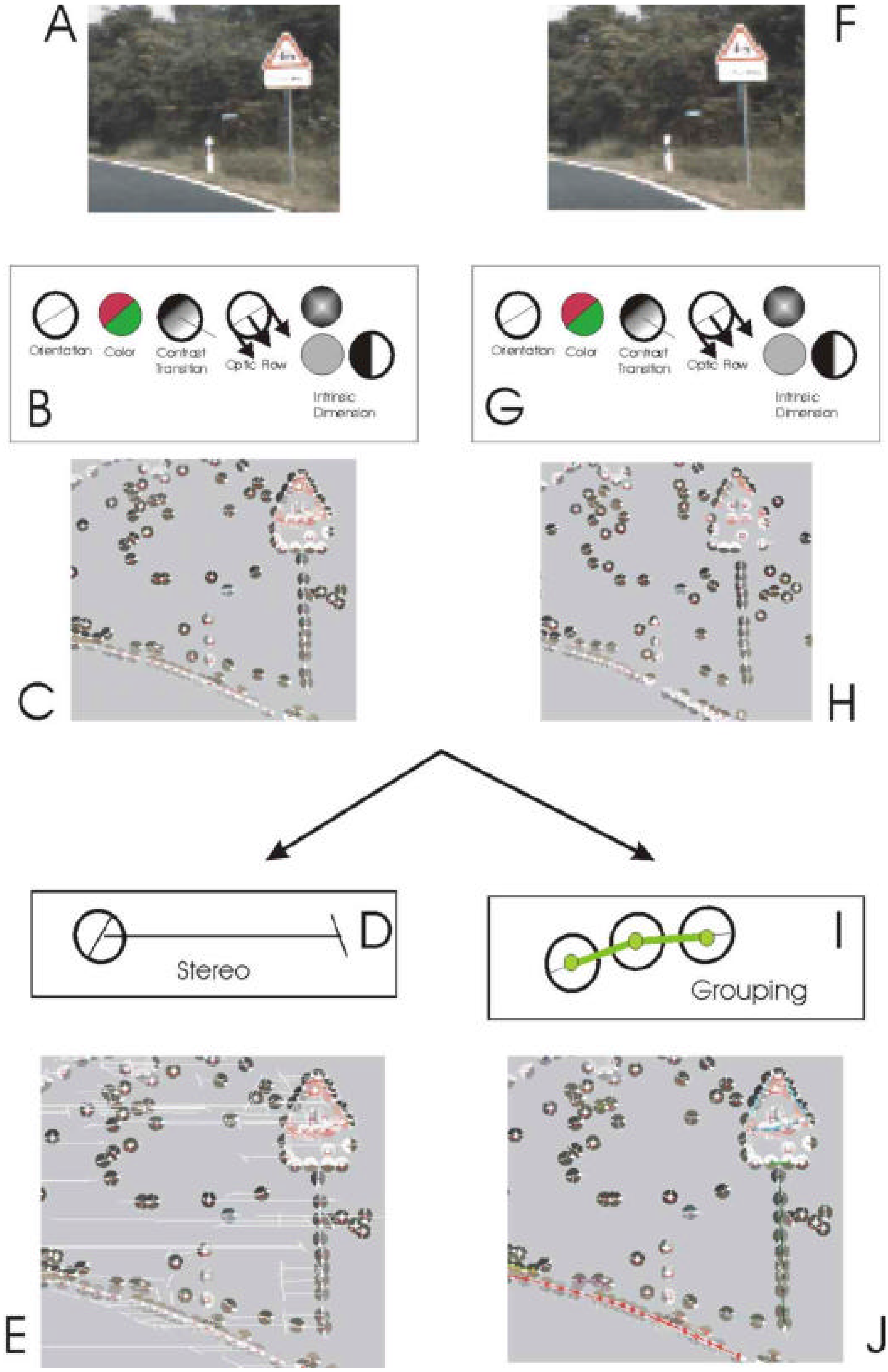

Fig. 2) Feature Processing. A: left image. F: right image. B, G: Visual Primitives representing the modalities orientation, Color, Contrast transition and optic flow. C: Extracted Primitives for the left image. D: Extracted Primitives for the right image. D: The modality stereo is associated to the Primitives. The position and orientation of the found corresponding primitive in the right image is attached to the Primitive in the left image. E: Stereo correspondences extracted. I: Grouping is symbolized by links. G: Realized grouping for the left image.

WP3

We addressed the problem of defining a computational framework to perform an adaptive analysis of dynamic scenes. The adaptive processes have a data-driven character, relying on contextual information, especially with respect to motion and stereo features. In this first stage of the project, the activity has been conducted by joint collaboration between Ita, Sco, and Ger/Mun who attempt to define complementary solutions for context-sensitive filters and specific operative ways and metrics to define the context itself. Specifically, sighting the entire workpart as a whole, two approaches emerged, led by Sco and Ita, respectively. The former defines the context as the statistical interdependences among local multi-modal (stereo) image descriptors to modulate the confidences associated to the features (deliverable D3.1a). The latter, defines context-sensitive motion filters (CSMFs) exploiting Kalman filter theory. In this case, Gestalt detection occurs through a spatially recurrent filter that checks the consistency between the spatial structural properties of the feature map and a structural rule expressed by the process equation of the Kalman filter.

In general we found, that Task 1,2 and 3 started to blend into each other at a rather early stage. Therefore, we decided to describe the progress in WP3 here in this more technical section as a whole referring to the task only "in-line".

a) Stereo Gestalts: A local visual multi-modal descriptor and its contextual adaptation according to statistical and deterministic image properties

1. Generation of Primitives: (WP3.1) We address the problem by defining a *local multi-modal visual descriptor* that generates our basic *Primitives* that represent (stereo) image information in a condensed and sparse way. Figure 2 shows the Primitive processing on stereo sequences (see Fig. 8) recorded in cooperation with Ind. The Primitives comprise 3D information in the modalities 2D-position and 2D-orientation, contrast transition (phase), color and optic flow (Fig. 2 primitives: B,G, results: C,H) and 3D information (3D-position and 3D-orientation) extracted by stereo processing (Fig. 2 primitives: D, results E). We use standard algorithms as well a new algorithm to extract Primitives in these modalities. However, we *also condense, stabilize and sparsify pixel groups* to achieve a more efficient and less redundant representation. Our primitives are adaptable in the sense that they adapt their parameters to the current image content (e.g., they move towards a pixel position with maximal energy.) Furthermore, confidences are associated to all aspects of the descriptor that are modifiable according to the context (see 5, below).
2. Biological and functional Motivation: (WP3.1) Our processing of Primitives has been motivated by the human visual system in the sense that they represent modalities that are processed in V1 and other specialized areas. Sparse representations play an important role for information processing in the brain. They essentially allow coding of information efficiently according to memory requirements and most importantly they permit improved data-exchange processes across spatial and temporal visual entities and across modalities, because of their sparseness.
3. Statistical Interdependencies: (WP3.1) We have investigated the spatial statistical interdependencies of our Primitives. We introduced a statistical measure that codes the interdependency between events, called 'Gestalt Coefficient'. Using this measure we could show that classical Gestalt laws like collinearity and parallelism can be linked to the statistics of natural images: Going beyond the 1/f law formulated by Field (1987) we could show that collinearity is the most dominant spatial second order context for edge

like features in natural images. Moreover, we could show that these interdependencies increase significantly when we take also other modalities that are coded in our Primitives (e.g.; color, contrast transition) into account. This corresponds to improved multi-modal grouping effects.

4. Grouping: (WP3.2) Based on the statistical measurements (see 3) we have defined a grouping process. By linking Primitives with high statistical interdependencies and applying the transitivity relation, groups of Primitives are formed which finally constitute our spatial Gestalts (Fig. 2 I, J). Once entities are grouped they can influence each other (see 5).

5. Modification according to spatial context: (WP3.1, WP3.3) The Primitives are subject to spatial contextual modification. In this context we have drawn a distinction between two principally different kinds on context that, in our view, have to be treated in different ways: We distinguish between deterministic and statistical interdependencies and we make use of both kinds of interdependencies. Deterministic interdependencies allow for mandatory predictions and are mostly based on geometric laws (such as the change of a rigid body under a motion). Statistical interdependencies only allow statements about the likelihood of other events.

We have implemented a modulation according to the spatial context in two ways:

A) Confidences associated to Primitives can be enhanced when they are part of a larger Gestalt (Fig. 3 A-C).
B) Confidences can be reduced when primitives are not part of a Gestalt and can be ignored for certain processing steps. In this way we could eliminate outliers according to the 3D-context applied to the task rigid body motion (RBM) estimation (Fig. 3 D-I).

These activities have led to the deliverable D3.1a and several publications.

b) Kalman-based context sensitive filters based on deterministic (i.e., geometric) spatial Gestalts

We defined a general framework to specify context sensitive motion filters (CSMFs) through Kalman filter theory. Due to its recurrent formalization, Kalman filter appears particularly promising to design CSMFs based on recurrent cortical-like interconnection architectures. To validate the approach, we defined a set of CSMFs based on deterministic (i.e., geometric) spatial motion Gestalts. In particular, the geometric properties of the optic flow field have been described through a specific set of elementary gradient-type patterns (e.g. see Fig. 4c), as cardinal components of a linear deformation space (Koenderink 1986, Raiguel, 1995).

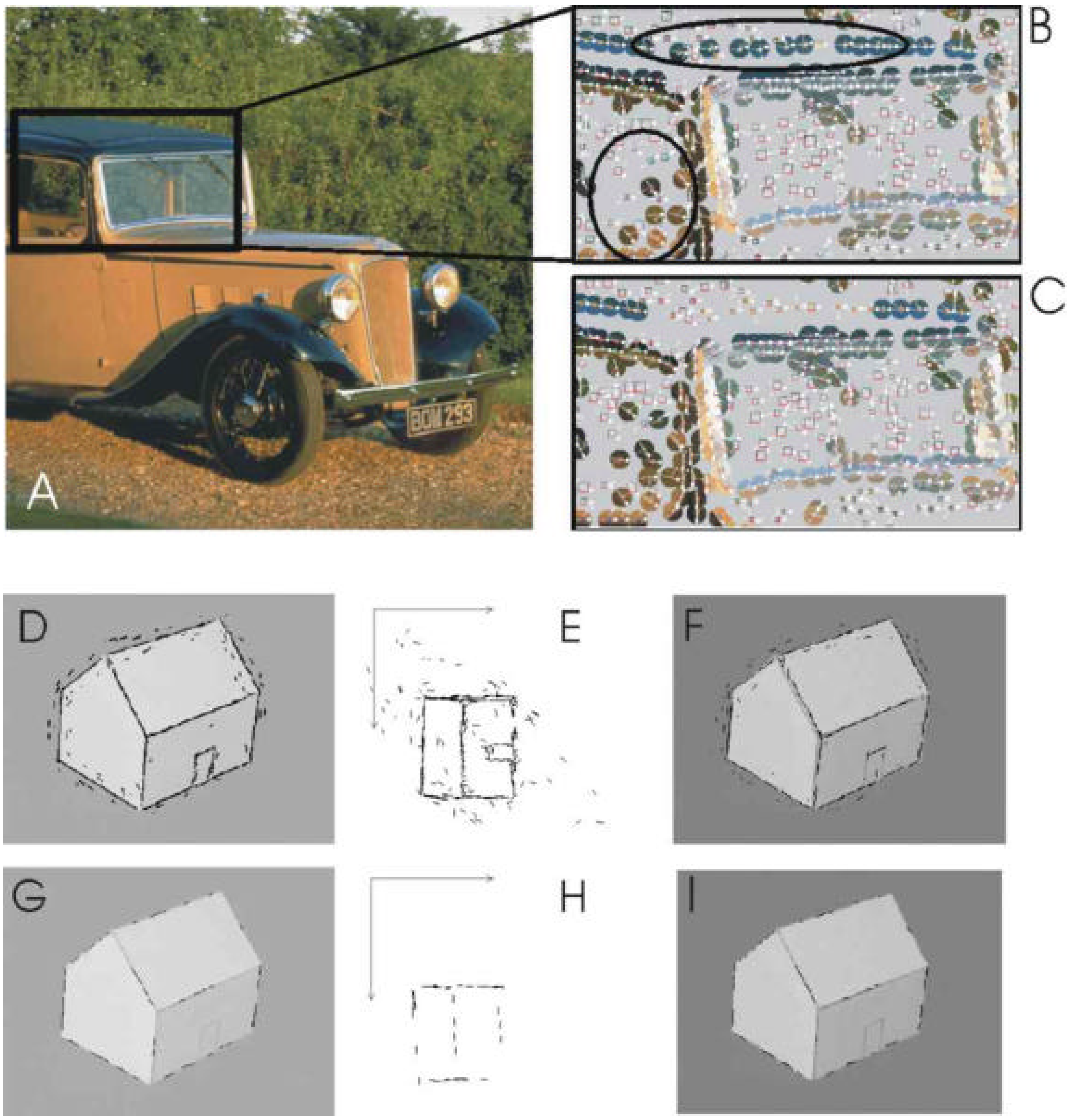

Fig. 3) A-C: Example for feature enhancement based on spatial Gestalts. A: Image of a car. B: Extraction of features utilizing feature enhancement with grouping based on the Gestalt coefficient. C: Feature extraction without grouping. In B and C all features above the same threshold are displayed with a filled circle while features below this threshold are displayed without circles. Note the successful detection of the low contrast edge in B (horizontal ellipse) and the reduction of non--meaningful features (vertical ellipse) compared to C. D-I: Utilizing Stereo Gestalts for stable RBM estimation. D: Projection of Stereo Representation with outliers corresponding to wrong correspondences. E: Projection of the representation to the x-z plane. F: Result of RBM estimation with representation with outliers. Note the imprecise RBM estimation. G: Representation that only consists of stereo primitives that are part of a group. H: Projection of the representation to the x-z plane. I: Result of RBM estimation with the representation in G,H. The RBM estimation has improved significantly.

By checking the presence of such Gestalts in optic flow fields, we make the interpretation of visual motion more confident. Given motion information represented by an optic flow field, we recognize if a group of velocity vectors belong to a specific pattern, on the basis of their relationships in a spatial neighborhood. Casting the problem as a Kalman filter,

the detection occurs through a spatial recurrent filter that checks the consistency between the spatial structural properties of the input flow field pattern and a structural rule expressed by the process equation of the Kalman filter. A motion-Gestalt emerges from a noisy flow as a solution of an iterative process of spatially interacting nodes that correlates the properties of the visual context with that of a structural model of the Gestalt. An example of motion-Gestalt detection in real-world driving sequence is shown in Fig. 4. The resulting CSMF behaves as a template model (Perrone and Stone 1994). Yet, its specificity lies in the fact that the template character is not built by highly specific feed-forward connections, but emerges by stereotyped recurrent interactions. Furthermore, the approach can be straightforwardly extended to consider adaptive cross-modal templates (e.g., motion and stereo). By proper specification of the transition matrix in the process equation of the Kalman filter, it can, indeed, potentially model any type of multimodal spatio-temporal relationships (i.e., multimodal spatio-temporal context).
These activities have led to the deliverable D3.1b and has been partially published in Ita.1

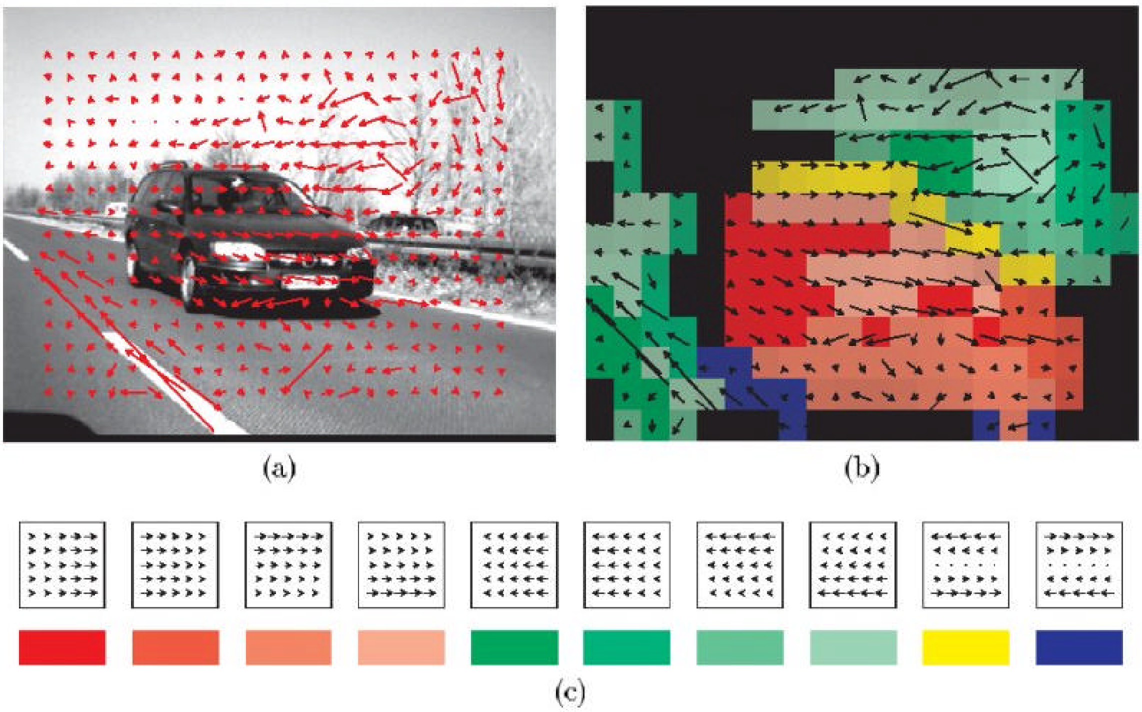

Fig. 4) a) The computed optic flow superimposed on a frame of the sequence provided by Hella. b) The resulting motion segmentation superimposed to the optic flow evidenced in a). c) The color code of the most probable elementary flow components (EFCs) encountered. We use the HSV color space: the hue identifies the direction of motion, the saturation discriminates the kind of EFC for a given hue and the value represents the probability of the given EFC. The two opponent EFCs are identified only by the hue, while the saturation and the value are fixed. The color black indicates that, for the considered region, the reliability of the segmentation is below a given threshold. It is worthy to note that the CSMFs detected the motion edges (yellow and blue colors) between the areas of coherent motions (green and red regions for leftward and rightward motions, respectively).

c) Receptive field structure of flow detectors for heading perception

Considering motion Gestalts as signals that originate by a common cause, and assuming “heading” as that cause, a comparative analysis of physiologically-based heading models has been performed by Ger to evidence common constraints at the level of local motion detectors (WP5). We were able to show that three of these models predict a somewhat counterintuitive circular receptive field (RF) structure with local or global motion opponency. These similarities in the RF structure across models suggest common principles that underlie human visual heading perception and that may be useful for technical applications. Parts of the results has been published Ger.1.

References:

J.J. Koenderink. Optic flow. Vision Research, 26(1): 161-179, 1986

J.A. Perrone and L.S. Stone. A model of self-motion estimation within primate extrastriate visual cortex. Vision Research, 34(21):2917-2938, 1994.

S. Raiguel, M.M. Van Hulle, D.K. Xiao, V.L. Marcar, and G.A. Orban. Shape and spatial distribution of receptive fields and antagonist motion surrounds in the middle temporal area (V5) of the macaque. Eur. J. of Neurosci., 7: 2064-2082, 1995

D. Field (1987) Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America, 4, 2379-2394.

WP4

Most of our activity has been directed towards producing a general low-level filtering framework that can be used for steerable-scalable filtering to be implemented as front-end of ECOVISION.

In WP4.1 we have developed routines using a bank of orthogonal Gaussian derivatives that can be efficiently steered to arbitrary orientation. In addition, we have developed an algorithm that can use this bank of Gaussian derivatives to predict local scale-space behaviour of image structure. This permits us to dynamically adapt the scale of filtering operations within a finite range.

We have also been continuing to develop the real-time implementation of our optic flow algorithm with the group in Spain. We currently have a real-time version operating using a low-cost framegrabber and a high-end PC that is can be used to investigate performance in real-world visual environments and develop efficient computational architectures. This system has been used to evaluate image sequences of driving situations from the industrial partner.

Task 4.1 Adaptive receptive fields as steerable-scalable filters

The aim of this work package is to deliver a software module for adaptable low level spatial filtering using Gaussian derivatives. These steerable-scalable filters can then be used in an adaptive algorithm in WP5.

Techniques have been developed (Freeman and Adelson, 1991) that use a “basis” set of orthogonal X-Y separable Gaussian derivative filter responses (Fig. 5). The spatial orientation of the individual filters in the basis set can be modified via a linear recombination of the responses. Since the basis set is separable the implementation is computationally very efficient.

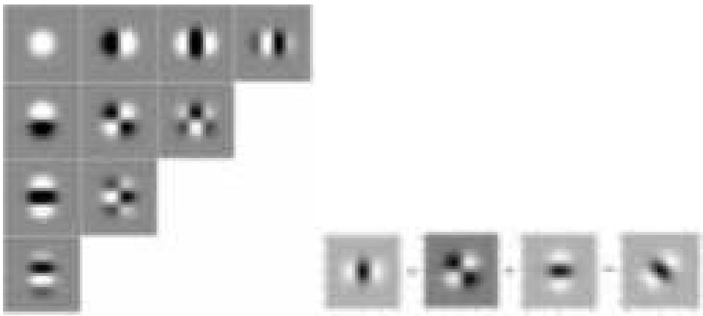

Fig.) 5 The impulse responses of the basis set of separable Gaussian derivative filters (left). X-derivative order increases from left to right, Y-derivative order increases from top to bottom. Diagonal sets of filters are of the same total derivative order. Any filter in the basis set can be `steered` to a new orientation using a weighted combination of filters of the same total derivative order (right).

Using a Taylor expansion with an additional scale excursion parameter allows us to predict the response of filters of different scales (Florack et al., 1996). Derivatives of image structure with respect to scale can be equated to spatial derivatives that we have already computed in the basis set using equation 4.1. Thus using the basis set we can construct an image representation at a different scale from that of the filters.

$$\left(\nabla^2 - \frac{\partial}{\partial s}\right) I(\bar{x}, s) = 0 \qquad \text{Eqn 4.1}$$

The Gaussian derivative measures can also be used to interpolate in regions of interest or between sparsely sampled regions using a Taylor expansion – this should lead us into a general space-variant sampling strategy during WP4.2.

One drawback with the current methodology is that scale-space prediction is limited to a finite range. We intend to try and extend this range by using a multi-scaled basis set.

Deliverables: (due month 12, finished month 10)
The software deliverable currently consists of a set of C++ object classes and DLLs that encapsulate separable Gaussian filtering, the Taylor expansion representation, and orientation/scale steering functionality. At present this software relies on Intel`s Performance Primitives for fast convolution. An optimized stand-alone version is under construction to free us from this restriction.

Task 4.2: General space variant mapping methods

The goal of this work package is to develop techniques that enable us to perform space variant image operations.

The research in this area is at a preliminary stage since it relies somewhat on the technology from WP4.1. However, two strategies are currently being investigated; the interpolation properties of the Taylor expansion representation and an adaptable recursive filtering strategy.

The research is not yet at the stage where any specific space variant mapping strategies have been implemented and tested.

References:
Freeman W. T., Adelson E. H., "The Design and Use of Steerable Filters", *IEEE Transactions PAMI* **13**(9), pp.891-906 (1991)
Florack L., ter Harr Romeny B., Viergever M., Koenderink J., "The Gaussian Scale-Space Paradigm and the Multiscale Local Jet", *International Journal of Computer Vision* **18**, pp.61-75 (1996)

WP5

Task 5.1: Psychophysical investigations of driving behavior and CoI in humans

We have conducted psychophysical experiments to gain deeper insight into the viewing behavior of drivers. Pre-recorded and computer-generated driving scenes were presented to human observers while their gaze behavior was recorded. We have analyzed the patterns of eye movements in order to find out which parts of the scene attract the viewer's interest an how his temporal viewing behavior can be parameterized. We focused particularly on possible conflicts of interest because then the pattern of eye movements is most informative about the attentional processes of the viewer. We were interested in how the direction of gaze related to the momentary flow field, the driving parameters, and the momentary task. Different tasks (heading, obstacle avoidance, etc.) were given to the subjects in order to study the relationship of eye movements to the task.

One set of experiments used driving scenes recorded with a video camera from inside a car driving in urban or motorway settings. These recordings contained rich scenery and several simultaneous tasks of the driver. The data showed many eye movements that were directed to objects in the scene. The analysis of these data proved difficult, however, because neither the scenery nor the momentary task could be defined rigorously in these sequences as both were depending on elements that could not be controlled by the experimenter (e.g. other cars, the placement of objects in the scene, the momentary movement parameters of the car, etc.). Therefore, in another set of experiments we recorded eye movements on computer-generated driving scenes that were under full experimental control. In these driving scenes the direction of motion was chosen by the computer, the scenery contained only few visible elements that were precisely defined in shape and location, and the task of the observer was strictly specified with respect to the driving parameters and the scene objects.

We compared two tasks that were modeled after typical subtasks of driving. The first was a heading tracking task in which the observer was required to track the direction of motion of the car (i.e., the visible focus of expansion) with the eyes as the car moved in a changing course. The second task was an obstacle avoidance and time-to-collision task in which the subject was required to quickly identify obstacles on the path and take appropriate action as the course of the vehicle changed over time. In this task, objects in the scene could be either target obstacles or irrelevant distractors, both being identified by their shape. Because we were interested in the role of optic flow in heading detection and obstacle avoidance in these situations, the scene did not contain road markings but rather depicted travel across an open field terrain. Images of road markings would have provided extra-flow cues to travel direction that would have interfered with the use of visual motion in these tasks.

The results of these experiments revealed several novel and surprising findings about the interplay between optic flow and eye movements.

First, in the heading tracking task, it was apparently very difficult to align gaze with the direction of heading. Subjects typically required a series of several saccadic eye movements until they reached the focus of expansion. The first saccade typically covered only 60% of the required distance to reach the FOE.

Second, in the obstacle avoidance task, subjects were quick to identify obstacles on the path even when the direction of heading changed. Typically the first saccade after a direction change was already on target. Since the obstacle avoidance task also required the estimation of the heading of the car in the absence of road markings the quickness with which the task was completed suggests that heading estimation was much quicker in these tasks than in the heading-tracking tasks described above.

Third, in the obstacle avoidance task observers virtually never looked at the FOE. Yet they were able to quickly and accurately estimate heading as demonstrated by the obstacle avoidance performance. This suggests that fixation of the FOE is not required for heading estimation.

Fourth, when there was no immediate obstacle on the path subjects adopted a scene scanning behavior in which gaze was directed to irrelevant elements of the scene (either potential obstacles that were not on the path or distractors) and from time to time switched back to the ground in front of the car in the direction of heading.

The results are consistent with a model of attentional resource distribution in which attention is (a) prevalently and immediately directed to obstacles on the course, (b) when idle directed to scanning of the environment irrespective of direct relation to the driving task, (c) from time to time directed back to the course for routine checking, and (d) seldom directed to the FOE.

The first set of experiments in WP 5.1 has thus been concluded as planned. In the following, more experiments will be required to consolidate these results. These experiments have to be quantitatively analyzed and prepared for scientific publication in the additional 6 months allocated to this WP.

Task 5.2: Task optimized representations and the CoI:

(Not much work was planned for year one, in spite of this we have started to address some of the questions.)
The goal of WP5.2 is to develop computational procedures to take advantage of space-variant mapping and steerable filter techniques (WP4.1 and WP4.2) to the specification of context Gestalt-filters. Choice of the actual Gestalt-steering process shall 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. Early work in this WP was intended to, first, define different specific tasks from the psychophysical analysis of WP5.1 and than formalize the demands of "how and where" the weights need to be increased/decreased for a given task. This work has begun as scheduled as tasks can be defined from the qualitative results of experiments in WP5.1 (see above). The concept which underlies these investigations is that of the "*common cause*": A given task will lead to a specific driving and viewing behavior. This leads to a specific type of global flow-field which is determined by the ego-motion and the center of gaze. Thus, this specific global flow-

pattern can be associated with the common cause of the behavioral pattern (driving and viewing) which has led to it. The global flow-pattern will normally be locally disturbed by other common causes (moving objects). These will lead to local deviations of the flow-pattern. The goal of this WP is to discover those different "flow-segments" by means of data-clustering techniques which start with simple common-cause Null-hypothesis flow-fields. From there on the clustering algorithm should be able to detect multiple (few) specific common causes. Partner Bel, who has deep knowledge about data driven clustering algorithms from other projects, is currently deepening their algorithmic knowledge of optic flow, and heading in particular, which will enable them to develop the phase space in which they can process driving scenes, in such a way that similar driving scenes, including the objects moving in them, belong to the same clusters which we then consider as Gestalts.

Revision of WP5.2: The revision concerns the plan to devise an objective method for the segmentation of motion Gestalts. It is described in detail in the Executive Summary above (see pg. 13-14).

Task 5.3: Implementing early cognitive control in a technical environment:
No work planned for year 1

WP6

Task 6.1: Task-optimized spatial representations – spatial attention case

The workpackage consists of three steps, for which the following progress was made:

1) We started with the data that were previously obtained in the laboratory of Bel by means of 2DG[2] experiments in monkey striate cortex (Vanduffel et al., 2002). The data are available under the form of digitized radiographic images showing the distribution of metabolic activity in layer IVC. We first pre-processed the radiographic images in order to cope with the noise due to artifacts inherent to 2DG experiments. For this, we developed an input transformation method. There are two types of images, one obtained during a task involving spatial attention, and another featural attention (left and right panels in Fig. 6). A model architecture was developed consisting of a two-dimensional array of neurons with a centre-surround intralayer connection pattern. The parameters controlling the ranges of the center and the surround parts of this connection pattern can be individually and for each neuron separately adapted. By virtue of our center-surround *intralayer connection* pattern, which in addition can be modulated, our model is essentially different from the (fixed) center-surround *input* pattern of Grossberg and co-workers (Raizada & Grossberg, 2001), and the biased competition model of Deco (2001).

2) When subtracting the spatial attention image from the featural attention image, we clearly observe a suppression of the metabolic activity in the cortical tissue that surrounds the representation of the stimulus (left and right panels in Fig. 6). We decided to model this suppression by a temporary modification of the centre-surround ranges (middle panel in Fig. 6). A novel learning rule for recurrent neural networks, based on a numerical determination of the error surface gradient, was tested and its robustness

[2] The 2DG or double-label desoxyglucose technique allows the registration of neural activity during two tasks in the same animal with high spatial but low temporal resolution.

verified. When initialised with the activations observed during a task involving spatial attention, the network's activation dynamically evolves to that observed during an orientation discrimination task which requires featural attention (i.e., from left to right panel in Fig. 6). Current investigations are concerned with the precise shape of the intralayer connection pattern. Difference of Gaussians and spline-based connection kernels have been explored, but still contain weaknesses (mainly relating to border effects) that need to be addressed.

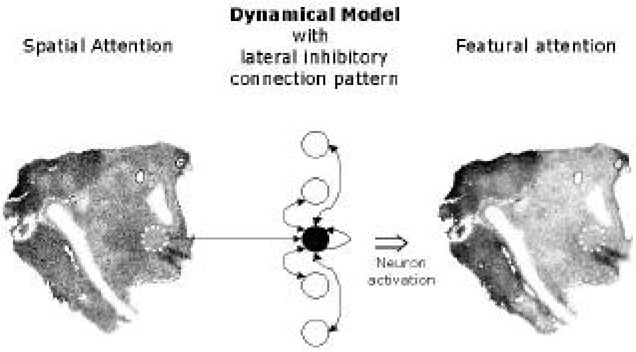

Fig. 6) Schematic overview of model components and operation

3) Once the algorithm converges, the connection patterns can be studied. Since the pattern is represented by two parameter values for each neuron, clustering procedures can be used to detect regularities. These results will enable the formulation of hypotheses concerning the nature of the modulatory attentional mechanism. We expect to find an increase in orientation selectivity in the regions corresponding to the stimulus representation and a decrease in the regions surrounding the stimulus. Simulations have shown that this can be achieved by a narrowing, respectively a widening of the inhibitory regions (see Fig. 7).

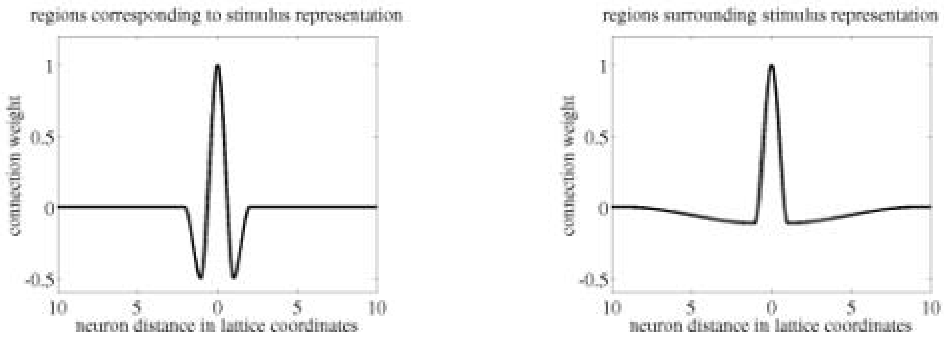

Fig. 7) Expected intralayer connection patterns

Task 6.2) Task-optimized representations - general case
No work planned for year 1.

<u>Revisions</u>: Since further 2DG experiments are not supported by the project, and since this class of experiments are extremely demanding in terms of material and personnel costs, additional data are <u>not</u> to be expected in the near future. As a consequence this WP is revised towards a more technical approach. This revision is described in detail in the Executive Summary above (see pg. 16).

References:

Deco, G. (2001). Biased Competition Mechanisms for Visual Attention in a Multimodular Neurodynamical System. *Emergent Neural Computational Architectures Based on Neuroscience*, 114-126.

Raizada, R. and Grossberg, S. (2001). Context-Sensitive Bindings by the Laminar Circuits of V1 and V2: A Unified Model of Perceptual Grouping, Attention, and Orientation Contrast. *Visual Cognition*, 8(3/4/5), 341-466.

Vanduffel W., Tootell R.B.H., Orban G.A. (2000) Attention-dependent Suppression of Metabolic Activity in the Early Stages of the Macaque Visual System. *Cereb. Cortex*, 10: 109-126.

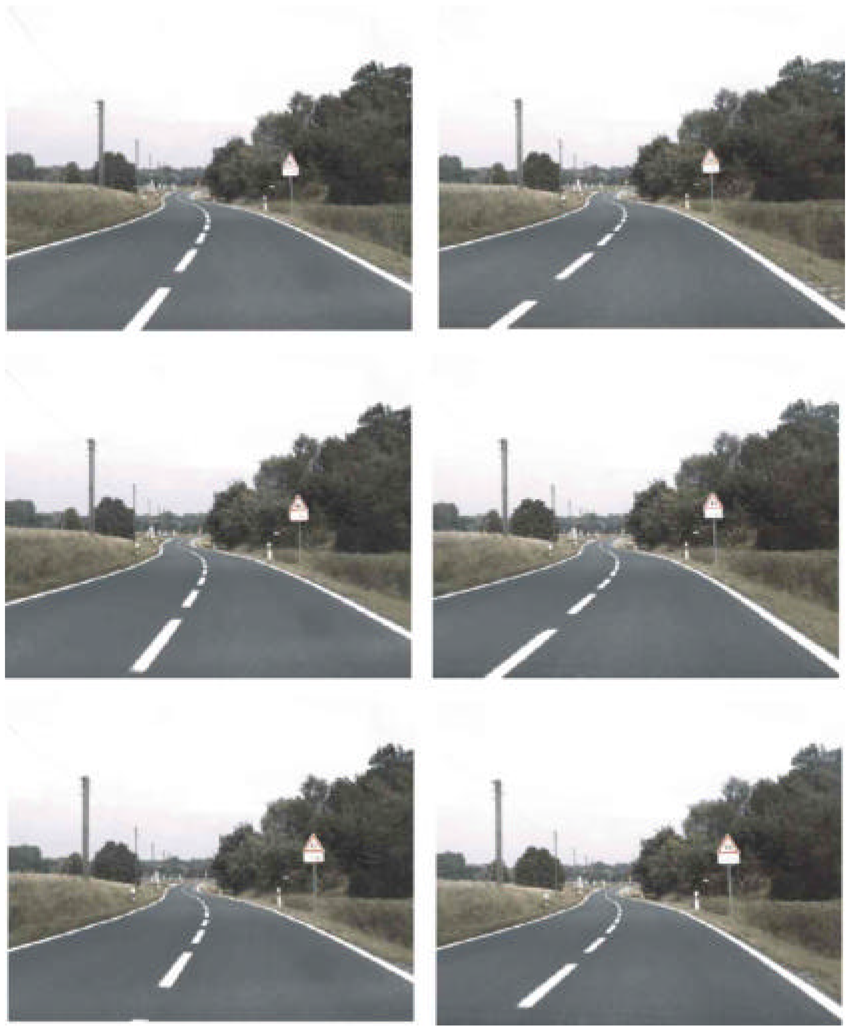

Fig. 8) Samples from a stereo movie recorded while driving (can be free-fused).

WP7

Task 7.1: Well adjusted driving sequences (Movies) with defined scene parameters

The first movies were recorded in April 2002. They show sequences of an overtaking car in different situations with sunny weather without rain or fog. In the beginning of June Norbert Krüger from Stirling visited HELLA for one week. He brought Sco's stereo-camera equipment with him for temporary installation in a testcar. With this stereo setting a large number of movies were recorded (in total 50 GB of data). A collection of this two types of movies was made available for the partners on the internet homepage in the members' area.
(http://www.pspc.dibe.unige.it/ECOVISION/private/index.html).

The recorded movies only show a small part of possible driving and weather situations. This was the reason for the decision to permanently implement a blindspot-/stereo-camera-system at HELLA (as a copy of the stereo-camera rig from Sco). This way it will be possible to record more blindspot and/or stereo movies under different weather conditions.

For the next set of movies a special movie recording program is currently written, which also records the time and speed data from the CAN (Controller Area Network) bus – a typical bus system in cars – in the testcar and show it also in the movie. The program is also running without a camera, so it will be possible to record the time and speed data of an overtaking car, too. This will allow calculating the true speed difference between the two cars during the hole overtaking process.

Task 7.2: Set of agreed-on I/O parameters (Deliverable Format: Data sheet)

For the specification of the I/O parameters two detailed specifications reports were compiled in co-operation with all partners who contributed their notions about feasibility of certain "product" features. One document describes a Pedestrian-Warning-System (PWS), the other one a Lane-Change-Warning-System (LCWS). In the reports it is described in detail how this system will have to work, which requests the sub-systems will have to grant and which I/O parameters will be necessary for it. The two documents have been made public for all partners middle of October.

WP8

no work planned for year 1

WP9

no work planned for year 1

WP10

Task 10.1: Web page
The web page is continuously updated and can be viewed at:
http://www.pspc.dibe.unige.it/~ECOVISION/

A written description of this in this report appears obsolete and is omitted here.

Task 10.2: Dissemination & Use Plan (DUP)
The DUP has been finished by month 6 and can be downloaded at:
http://www.pspc.dibe.unige.it/ECOVISION/private/index.html (from "Official report")

Task 10.3: Scientific Publications
At total of 7 journal papers has been published, furthermore there are 17 conference contributions. In addition ECOVISION has sponsored two workshops, one in Granada (*jcra'2002*) and which will take place mid November in Bochum (*Dynamic Perception*). Please see Progress Report Sheets at pages 51-63 of this report for details.

Task 10.4: Workshop at Stirling
We have started planning an international workshop with <100 participants for May 2004 in Scotland. A conference site at the Isle of Skye has been contacted.
Title of the WS: *Early cognitive vision*
Duration 4 days + 1 day for tutorials.
Topics:

1) Low-level front ends
2) Aspects of intermediate complexity level vision ("semi complex feature constellations")
3) Tasks and Applications

Semi-open format: Students can apply, planned a bit like a summer school format.
Proceedings shall be on the Web or on a CD.
More money shall be sought for this (planned to apply in Feb 2003)
Speakers shall be invited late spring 2003.
During the November ECOVISION meeting we will agree on the list of speakers.

WP11

A first management report has been written by month 6
Smaller budget reallocations are:

1) Ind moves less than 10% from personnel to durable equipment in order to set up their own camera rig.
2) Ind moves approx. 1400 Euro to Sco to cover Norberts travel.

Spa requested a reallocation of dedicated time (number of man-months) without affecting its budget, for the reason that more effort was required for his workpackages as originally planned. Since this shift did not affect Spa's budget at all it was approved by the coordinator.

Two contract amendments are currently being processed by the commission:

a) Markus Lappe (original partner Ger) moved from Bochum to Münster to take up a full professorship there (new partner Mun). All obligations have been without changes transferred to Mun.
b) Partner Bel requested to move his budget almost completely into "personnel". This was justified by additional resources that became available to Bel early during this project which can cover for the other costs (equipment, auditing, etc.). Thus, this major budget reallocation was approved first by the coordinator and then by the commission. The reason for this amendment was the increase work-load of Bel due to the revision of WP5 (see there)

DELIVERABLES TABLE

Project Number:	IST – 2001 - 32114
Project Acronym:	ECOVISION
Title:	Artificial vision systems based on early-cognitive cortical processing

Del. No.	Revision	Title	Type[1]	Classification[2]	Due Date	Issue Date
3.1	1	**Visual templates with spatial context - Gestalts in Space**	**R+S**	**Rest.**	**30.9.02**	**30.9.02**
4.1	1 ahead of time	**Adaptive Receptive Fields as Steerable -Scalable Filters**	**S**	**Rest.**	**31.12.02**	**30.10.02**
5.1	1 Ahead of time	**Psychophysical investigations of gaze and CoI during simulated driving in humans**	**R**	**Publ.**	**31.12.02**	**31.12.02**
7.1	1	**Calibrated Camera Sequences**	**I**	**Int.**	**30.6.02**	**30.6.02**
7.2	1	**I/O Specifications**	**R**	**Int.**	**30.9.02**	**30.9.02**
10.2	1	**Dissemination and Use Plan**	**R**	**Rest.**	**30.6.02**	**30.6.02**

1 *R: Report; D: Demonstrator; S: Software; W: Workshop; I=Image Sequences (movie files)*

2 *Int.: Internal circulation within project (and Commission Project Officer + reviewers if requested)*
Rest.: Rest. = Members of the consortium and project group (i.e. employed under ECOVISION) and Commission SO + reviewers only
IST: Circulation within IST Programme participants
FP5: Circulation within Framework Programme participants
Pub.: Public

document

DELIVERABLE SUMMARY SHEET

Project Number:	IST-2001-32114
Project Acronym:	ECOVISION
Title:	Visual templates with spatial context – Gestalts in space

Deliverable N°:	3.1
Due date:	30-09-2002
Delivery Date:	30-09-2002

Short Description: The problem of defining a computational framework to perform an adaptive analysis of dynamic scenes is addressed. Sighting the entire Workpart as a whole, two complementary approaches emerged, led by Sco and Ita, respectively.

In part A, we define the context as the statistical interdependences among local multimodal image descriptors to modulate the confidences associated to the features and arrive at stereo-Gestalts.

In part B, we define context-sensitive filters (CSFs) exploiting Kalman filter theory. In this case, Gestalt detection occurs through a spatial recurrent filter that checks the consistency between the spatial structural properties of the feature map and a structural rule expressed by the process equation of the Kalman filter

Partners owning: Ger/Mun, Ita, Sco

Partners contributed: Ger/Mun, Ita, Sco

Made available to: Partners, Commission, Reviewers

DELIVERABLE SUMMARY SHEET

Project Number: IST – 2001 - 32114

Project Acronym:ECOVISION

Title: Adaptive Receptive Fields as Steerable-Scalable Filters

Deliverable N°:	4.1
Due date:	31.12.02
Delivery Date:	31.10. 02

Short Description: The software deliverable currently consists of a set of C++ classes that encapsulate separable Gaussian filtering, the Taylor expansion representation, and orientation/scale steering functionality. Two Microsoft Visual Studio workspaces are provided as demonstrations.

At present this software relies on Intel`s Performance Primitives for fast convolution operations.

An optimized stand-alone version is under construction to free us from this restriction.

Partners owning: Eng

Partners contributed: Eng, Spa

Made available to: Partners, Commission.

DELIVERABLE SUMMARY SHEET

Project Number: IST-2001-32114

Project Acronym: ECOVISION

Title: Psychophysical investigations of gaze and CoI during simulated driving in humans

Deliverable N°: 5.1

Due date: 31.12. 2002

Delivery Date: preliminary report exists.

Short Description: Scientific report of results of psychophysical investigations of gaze behavior and center of interest during simulated driving in humans. The report describes eye position recordings of human observers in different simulated driving tasks. In these tasks, participants watched a screen displaying a movement simulation of driving along a course with variable directions and speeds, while they were given tasks like monitoring the driving direction or estimating potential collisions with obstacles. Their directions of gaze were recorded and analyzed over time in relation to momentary driving parameters and scene layout. From this analysis we infer the distribution of interest of the driver.

Partners owning: Ger/Mun

Partners contributed: Ger/Mun, Sco, Ind

Made available to: Preliminary Report to Partners

DELIVERABLE SUMMARY SHEET D7.1

Project Number: IST – 2001 - 32114

Project Acronym: ECOVISION

Title: Well adjusted driving sequences (Movies) with defines scene parameters

Deliverable N°: 7.1

Due date: June 2002

Delivery Date: July 2002

Short Description: Blind spot movies are recorded and placed in the members' area on page http://www.pspc.dibe.unige.it/ECOVISION/private/index.html.

Also with the help of Norbert Krüger stereo sequences are recorded and also placed on the same page.

For the possibility to reproduce the camera setting in different weather situations, a stereo- and blindspot-camera-setting was build up by Hella for the permanent installation in the Test car.

With the new setting more movies are planed to record the next time and to publish for the partners.

Partners owning: Ind

Partners contributed: Ind , Sco

Made available to: Partners, Commission.

DELIVERABLE SUMMARY SHEET D7.2

Project Number: IST – 2001 - 32114

Project Acronym: ECOVISION

Title: Set of agreed-on I/O parameters

Deliverable N°: 7.2

Due date: September 2002

Delivery Date: October 2002

Short Description: For the specification of driver-assistant-systems, there were two reports written:

One for a pedestrian-recognition-system and one for a lane-change-assistance-system.

In these specifications the requirements for the two systems are described in detail.

The reports were published for the Partners by middle of October.

Partners owning: Ind

Partners contributed: Ind,

Made available to: Partners, Commission.

DELIVERABLE SUMMARY SHEET

Project Number: IST – 2001 - 32114

Project Acronym: ECOVISION

Title: Dissemination and Use Plan

Deliverable N°: 10.2

Due date: July 2002

Delivery Date: July 2002

Short Description: This document consists of an overview of the expected results followed by the approach to dissemination and use and a section about market projections. The more detailed part of the DUP focuses on the description of dissemination plan and covers conferences, publications and web presence. The last section (Description of use plan) is concerned with the plans of HELLA (partner Ind) describing their future plan to use such systems.

The DUP can be downloaded at:

http://www.pspc.dibe.unige.it/ECOVISION/private/index.html (from "Official report")

Partners owning: Sco

Partners contributed: all

Made available to: Partners, Commission

Effort in person months for reporting period 1/1/2002 -31/12/2002

<<Due Jan 2003>>

	Partner A				**Partner B**				**Partner C**				**Partner D**				**Total**			
	Period		**Total**		**Period**		**Total**		**Period**		**Total**		**Period**		**Total**		**Period**		**Total**	
WP/Task	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
WP1																				
Task 1.1																				
Task 1.2																				
WP 2																				
Task 2.1																				
Task 2.2																				
Total																				

Period: Est.: estimated effort in contract for period Act.: effort actually spent in period

Total: Est.: estimated cumulative effort to date in contract Act.: cumulative effort to date actually spent

Costs in euro for reporting period 1/1/2002 -31/12/2002

<<Due Jan 2003>>

	Partner A				**Partner B**				**Partner C**				**Partner D**				**Total**			
	Period		**Total**		**Period**		**Total**		**Period**		**Total**		**Period**		**Total**		**Period**		**Total**	
Cost category	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.	Est.	Act.
Direct costs																				
1. Personnel																				
2. Durable equipment																				
3. Subcontracting																				
4. Travel and subsistence																				
5. Consumables																				
6. Computing																				
7. Protection of knowledge																				
8. Other specific costs																				
Subtotal																				
Indirect costs																				
9. Overheads																				
Total																				

Period: Est.: estimated costs in contract for period — Act.: actual costs in period

Total: Est.: estimated cumulative costs to date in contract — Act.: cumulative actual costs to date

PROGRESS OVERVIEW SHEETS

Note that these sheets will not contain any man-month allocations at the time of this report, because of the two pending contract amendments.

The complete sheets will be re-attached to the Cost Breakdown in Januar 2003.

PROGRESS OVERVIEW SHEET

Organisation: Sco

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 3	
Task 3.1	Definition, implementation and use of Stereo Gestalts
Task 3.2	First steps towards the implementation of temporally adaptive Stereo Gestalts
Task 3.3	Integration of the Stereo Gestalts with Rigid Body Motion in order to obtain a recursive process of updating and tracking
WP 7	
Task 7.1	Recording of camera sequences with Ind
WP10	
Task 10.1	Writing the Dissemination and Use Plan
WP11	
Task 11.1	Project Management

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
3.1	Visual templates with Spatial Context - Gestalts in Space	Final
7.1	Calibrated Camera Sequences	Final
10.1	DUP	Final

Dissemination actions (articles, workshops, conferences etc.)

Sco.1) Krüger, N. and Wörgötter, F. (2002). Multi Modal Estimation of Collinearity and Parallelism in Natural Image Sequences. Accepted for 'Network: Computation in Neural Systems'.

Sco.2) Krüger, N., Felsberg, M., Gebken, C. and Pörksen, M. (2002). An explicit and Compact Coding of Geometric and Structural Information Applied to Stereo Processing. Proceedings of the DAGM workshop `Vision, Modeling and Visualization', Zürich

Sco.3) Krüger, N. and Wörgötter (2002) The Gestalt Principle Collinearity and the Multi--Modal Statistics of Natural Image Sequences, Proceedings of the DAGM workshop `Vision, Modeling and Visualization', Zürich

Sco.4) Krüger, N., Th. Jäger and Perwass, Ch. (2002). Extraction of Object Representations from Stereo Image Sequence Utilizing Statistical and Deterministic Regularities in Visual Data. Proceedings of the 2nd Workshop on Biologically Motivated Computer Vision (BMCV 2002).

Sco.5) Krüger, N. and Wörgötter (2002) Statistics of second order multi-modal feature events and their exploitation in biological and artificial visual systems. Proceedings of the 2nd Workshop on Biologically Motivated Computer Vision (BMCV 2002).

Sco.6) Krüger, N. and Wörgötter F. (2002). Different Degree of Genetical Prestructuring in the Ontogenesis of Visual Abilities based on Deterministic and Statistical Regularities. Proceedings of the 'Workshop on Growing up Artifacts that Live', SAB 2002.

Sco.7) Krüger, N. and Wörgötter (2002) Multi-Modal Feature Statistics and Self--Emergence of Feature Constellations, Workshop 'Dynamic Perception' Bochum 14-15 Nov.

Software Package: MoInS: Modality Integration Software (2002).
http://www.cn.stir.ac.uk/ImageAnalysis/MoInS/index.html

Deviations from the planned work schedule/reasons/corrective actions/special attention required

As explained in the main part of this report we expect that WP3.2 and 3.3 will blend into each other. Thus, in year two we will work on both aspects at the same time.

Planned actions for the next period

Record more camera sequences under different weather conditions
Merge Stereo (Sco) and Motion (Ita) Gestalts
Utilize Stereo and Motion Gestalts in the context of RBM

PROGRESS OVERVIEW SHEET

Organisation: Bel

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 5	
Task 5.1	• Started background research for the development of an algorithm that detects moving objects (Gestalts) in flow fields • The algorithm will combine clustering techniques with the determination of ego-motion
WP 6	
Task 6.1	• Preliminary results for attention model with lateral, inhibitory interactions • Started compiling complete model

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
	No deliverables due this reporting period.	

Dissemination actions (articles, workshops, conferences etc.)
Bel.1) Van Hulle, M.M. (2002). Kernel-based topographic map formation achieved with an information-theoretic approach. (invited paper) Neural Networks, 15(8-9), 1029-1039. Bel.2) Van Hulle, M.M., & Gautama, T. (2002). Optimal smoothing of kernel-based topographic maps with application to density-based clustering of shapes. J. VLSI Signal Processing Systems for Signal, Image, and Video Technology, in press. Bel.3) Van Hulle, M.M. (2002). Kernel-based topographic map formation achieved with normalized Gaussian competition. 2002 IEEE Workshop on Neural Networks for Signal Processing (Martigny, Valais, Switzerland, September 4-6, 2002), pp. 169-178. Bel.4) Deleus, F.F., De Maziere, P., & Van Hulle, M.M. (2002). Functional connectivity modelling in fMRI based on causal networks. 2002 IEEE Workshop on Neural Networks for Signal Processing (Martigny, Valais, Switzerland, September 4-6, 2002), pp. 119-128. Bel.5) Van Hulle, M.M. (2002). Joint entropy maximization in kernel-based topographic maps. Neural Computation, 14(8), 1887-1906. Bel.6) Van Hulle, M.M. (2002). Kernel-based topographic map formation by local density modeling. Neural Computation, 14(7), 1561-1573.

Bel.7) Gautama, T., & Van Hulle, M.M. (2002). A phase-based approach to the estimation of the optical flow field using spatial filtering. IEEE Trans. Neural Networks, 13(5), 1127 -1136.

Deviations from the planned work schedule/reasons/corrective actions/special attention required

Concerning WP 6, there will be no further 2DG experiments.
To further respond to the needs of the consortium, we will increase our efforts in WP 5 leading to an increased number of man-month covered by an amendment of the contract.

Planned actions for the next period

- WP 6 : Completion of attentional model + study of final connection pattern
- WP 5 : Development of object segmentation algorithm

PROGRESS OVERVIEW SHEET

Organisation: Eng

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 4	
Task 4.1	Developed theory and optimised software implementation for steerable-scaleable filtering.
Task 4.2	Started evaluation of space-variant filtering methods using adaptive recursive strategy and sparse convolution/Taylor expansion representation.

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
4.1	Adaptive Receptive Fields as Steerable-Scalable Filters	final

Dissemination actions (articles, workshops, conferences etc.)

Eng.1) Dale. J. L., Johnston. A., "A Real-Time Implementation of a Neuromorphic Optic Flow Algorithm", European Conference on Visual Perception, Poster Presentation

Eng.2) Johnston, A., McOwan, P.W. and Benton, C.P. (2002) Biological computation of image motion from flows over boundaries, J. Physiol. Paris. (in press).

Deviations from the planned work schedule/reasons/corrective actions/special attention required

Three months ahead of time.

Planned actions for the next period

Two members of our group will be attending the ECOVISION meeting in Munster to present our work on dynamically adapting the filtering operations in a space variant manner. We plan to publish a paper on sharp, high resolution image reconstruction using the Taylor expansion representation using the scale-space filtering techniques developed.

PROGRESS OVERVIEW SHEET

Organisation: Ger/Mun

Workpackage/ Task	Planned effort	Planned Date		Actual Date		Resources employed	Cumulative Resources
	Whole Project	Start	End	Start	End	This Period	Since start
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 3	
Task 3.1	• Description and comparison of receptive field mechanisms for heading detection
WP 5	
Task 5.1	• Experiments on gaze behavior during simulated driving
Task 5.2	• Definition of tasks

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D5.1	Psychophysical Investigations of Gaze and CoI during Simulated Driving in Humans	Draft

Dissemination actions (articles, workshops, conferences etc.)

Ger.1) J. A. Beintema, A. V. v. d. Berg, and M. Lappe. Receptive field structure of flow detectors for heading perception. In: T. G. Dietterich and S. Becker and Z. Ghahramani (eds), Advances in Neural Information Processing Systems 14, MIT Press, 2002

Ger.2) M. Lappe. Building blocks for time-to-contact estimation by the brain. In: H. Hecht and G. Savelsberg, Theories of Time-to-Contact (Advances in Psychology Series), Elsevier, 2003 (in press)

Ger.3) Rolf P. Würtz, Markus Lappe (eds.). Dynamic Perception. Workshop of GI section 1.0.4 "Image Understanding" and the European Networks MUHCI and ECOVISION. Infix Verlag, St. Augustin, Germany, 2002

<u>Organization of Workshop</u>

Dynamic Perception. Workshop of GI section 1.0.4 "Image Understanding" and the European Networks MUHCI and ECOVISION. Held at the Ruhr-University of Bochum November 14-15, 2002.

This workshop was organized by Rolf Würtz and Markus Lappe and co-sponsored by ECOVISION. Its interdisciplinary theme "Dynamic Perception" covered important aspects of perception in computer science, neurobiology, and psychology. Specific topics ranges from optical flow to the coordination of perception and action, the analysis of human motion and multimodal sensory integration. The workshop consisted of 20 talks and 32 posters.
Deviations from the planned work schedule/reasons/corrective actions/special attention required
In early 2002 Markus Lappe took a new job as a Professor at the University of Münster, Germany. Because of this, an amendment to the contract was requested in order to transfer the project from Bochum to Münster University. Because of the change of labs, which involved changes to the personnel and to the resources, the implementation of the planned recruitment of researchers and the acquisition of equipment was delayed. It is hoped that the acquisition and recruitment processes can be completed until the end of the year or the beginning of next year, once the amendment is finalized. This led to a delay in WP5.1 of about 6 months.
Planned actions for the next period
We will begin work on the use of space variant mapping (WP4) and on the combination of stereo and motion cues for robust flow field representation (WPs 3 and 4). We will also start investigations of the use of statistical regularities for the definition of Gestalt-like flow descriptions in cooperation with SCO (WPs 3 and 5).

PROGRESS OVERVIEW SHEET

Organisation: Ita

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	
Task 1.1	Contribution to the choice of the motion algorithm that is most optimally suitable to FPGA implementation by a comparative analysis of different motion algorithms from the literature. Specific consideration on the Simoncelli and Heeger's neuromorphic algorithm.
WP 3	
Task 3.1	Contribution to the definition of motion Gestalts in space through hierarchical combination of 1st-order differentials of the optic flow. Formulation of a Kalman-based approach to design adaptive filters for early cognitive analysis of outdoor dynamic scenes.

Deliverables due this period

Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D1	Best motion algorithm	Draft
D3.1	Visual templates with spatial context – Gestalts in space	Final

Dissemination actions (articles, workshops, conferences etc.)

Ita.1) S.P. Sabatini, F. Solari and G.M. Bisio (2002) Detection of First-order Elementary Components in Noisy Optic Flow Fields Through Context Sensitive Recurrent Filters. at "4th Workshop on Dynamic Perception" to be held in Bochum, 14-15 Nov.

Ita.2) S.P. Sabatini, F. Solari, G. Andreani, C. Bartolozzi, and G.M. Bisio. A hierarchical model of complex cells in visual cortex for the binocular perception of motion-in-depth . In: T. G. Dietterich and S. Becker and Z. Ghahramani (eds), Advances in Neural Information Processing Systems 14, MIT Press, 2002.

Deviations from the planned work schedule/reasons/corrective actions/special attention required
none
Planned actions for the next period
WP1: almost finished. WP3.1: Consolidation of the proposed methodology on the basis of experimental tests performed on real sequences. Inclusion of the statistical models for motion Gestalts discussed with the other partners. WP3.2: The grounds to support the research activities of WP3.2 have been substantially laid. The activities conducted during the past period, focused on spatial context, can be indeed extended to consider temporal context, e.g., the (spatially local-) temporal properties of the visual signal related to the constraints posed by rigid body motion. We expect that a proper specification of the matrix Phi in the process equation of the Kalman filter will be used to this purpose.

PROGRESS OVERVIEW SHEET

Organisation: Spa

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	
Task 1.1	• Motion Algorithm study. The McGM is implemented in a high level language to gain expertise with the details of the model. • It's Hardware implementation viability is at first time evaluated.
Task 1.2	• Resolution and bit precision requirement assessment. • Factors that may affect the accuracy of the model.
WP 2	
Task 2.1	• Define technical specifications, that may be crucial for the algorithm operability, to be taken into account in the first VHDL approach.
Task 2.2	• 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.
Task 2.3	• Study different possible projections on an FPGA board and their implications in the final efficiency.

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D1	Best Motion Algorithm	Draft

Dissemination actions (articles, workshops, conferences etc.)
Spa.1) Ros, E., Pelayo, F.J., Prieto, A., del Pino, B.: Ingeniería Neuromórfica: El papel del hardware reconfigurable, pp. 89-92, jcra'2002, (ISBN: 84-699-9448-4), 2002. Spa.2) S. Mota, Ros, E. Pino, B., Botella, G. Pelayo, F.J., Prieto, A.: Sistema de monitorización de adelantamientos en tiempo real, pp. 93-96, jcra'2002, (ISBN: 84-699-9448-4), 2002. Spa.3) Morillas, C., Morillas, S., Ros, E., Díaz, A.F., Pino, B., Pelayo, F.J.:Real-time vision guided movement with reconfigurable Hardware, accepted for the Workshop Dynamic Perception, Bochum, Germany, November 2002. a) Ros, E.: "FPGA Implementation feasibility of the McGM". Technical Report, October, 2002. (internal report) b) Andreani, G., Cavalleri, C., Sabatini, S.P.: "Some Notes on the Simoncelli and Heeger's Algorithm", Technical Report, September, 2002. (internal report). **Workshop sponsor:** As part of the local committee we let ECOVISION be sponsor of the National Conference of Reconfigurable Hardware (II Jornadas sobre Computación Reconfigurable y Aplicaciones: jcra'2002) held in Almuñecar, Granada.
Deviations from the planned work schedule/reasons/corrective actions/special attention required
Stereo vision algorithms will also be taken under consideration to study their implementation viability by means of FPGAs.
Planned actions for the next period
Going on mainly with the tasks of WP2.

PROGRESS OVERVIEW SHEET

Organisation: Ind

Workpackage/ Task	**Planned effort**	**Planned Date**		**Actual Date**		**Resources employed**	**Cumulative Resources**
	Whole Project	**Start**	**End**	**Start**	**End**	**This Period**	**Since start**
Total							
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 7	
Task 7.1	Blindspotmovies were recorded and placed on the members' area page on the ECOVISION homepage. Also with the help of Norbert Krüger stereo sequences were record and also placed on the same page. A camera setting for permanent installation in the test car were defined and the installation was built.
Task 7.2	For the I/O parameters of the driver-assistant-systems, there were two reports written. One for a pedestrian-recognition-system and one for a lane-change-assistant-system.
WP 8, WP9	No action to be taken in this period

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
7.1 7.2	Well adjusted driving sequences (Movie) with well-defined scene parameters Set of agreed-on I/O parameters	Final Final

Dissemination actions (articles, workshops, conferences etc.)
First movies from WP 7.1 were published for the partners on the members' area on the home page. Two specifications papers were also published in the middle of October for the WP 7.2.
Deviations from the planned work schedule/reasons/corrective actions/special attention required
A number of movies for the WP 7.1 is existing, for better tests more movies will be recorded next year, to have more weather situations covered, like snow or fog.
Planned actions for the next period
To record more movies with the permanently mounted camera rig.