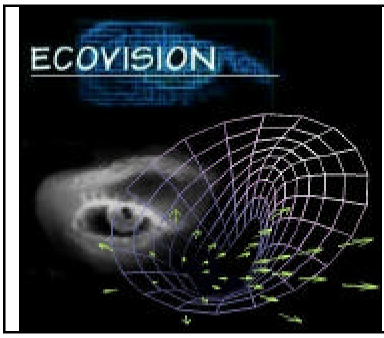


ECOVISION

IST – 2001 - 32114

Artificial vision systems based on early-cognitive cortical processing

Periodic Progress Report N°:2
Covering period 1.1.2003-31.10.2003

Report Version: 1.0
Report Preparation Date: 17.11. 2003
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
(**Mun**) 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)

Table of Contents

Executive Summary

<u>Goals of ECOVISION (copied from the first-year report):</u> 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.

<u>Summary of Achievements:</u> We are at month 22 of the ECOVISION project at the time of compiling this report. As in the last year we can state that we have again in almost all WPs achieved more than what was planned for the first 24 months, in none of them we are behind schedule. All planned deliverables have finished in due time, from some other deliverables (5.2, 6.2) detailed preliminary reports exist already. Table 1 shows the publication activity of the consortium over the first two years, which nicely demonstrates the growing number of results from the project.

	Year 1	Year 2	Total
Rev. Journal Papers	4	23	**27**
Conference Contrib.	12	27	**39**
Book Chapters	0	4	**4**

Table 1: Publication activities of ECOVISION

FPGA implementations (which were not even foreseen in the proposal!!) now exist for **two** optic flow algorithms. Thus we have reached a stage in which we can actually start to design a physical system where the front-end is an FPGA. In addition, Spa has in mind to also implement a new stereo algorithm on FPGA, which is also unplanned work, but which would again lead to a more advanced system.
Furthermore, Partner Spa has raised considerable public interest with their FPGA work which has led to several newspaper publications and TV-appearances.

ECOVISION has prepared a conference which will take place end of May 2004. Cooperations have been continued as documented by several visits. At the moment 6 shared publications exist. The neuroscience part which had been criticized in the last review has been upgraded with two models and one actual application of a new neuroscience based processing principle in ECOVISION. We have:

1) designed an advanced Early visual Attention/Center-of-interest model as well as
2) an Energy model for motion in depth. Both models are directly leading to a design impact for ECOVISION.
3) We are using the receptive field filtering a (probably) performed in area MT to correct flow fields in order to correctly retrieve the heading information.

As a **highlight** we state: Specifically the energy model has led to one remarkable discovery which predicts that cortical motion-in-depth detectors ("MID" detectors which are cells which receptive fields that are sensitive to MID) are **not** sensitive to disparity.

This is paradoxical and unexpected. *So far, there have no MID detectors been found in the cortex*. This may be due to the fact that experimenters have in seemingly natural way looked only at disparity sensitive (i.e. depth-sensitive) cells when searching for MID sensitive cells. The model of Ita predicts that one needs to look at disparity insensitive cells instead. Cooperative experiments to test this are now planned together with an experimental group (not in the ECOVISION consortium).

Furthermore we state that the energy model formalism found by Ita for MID cells now fill the last gap in the class of energy models: Such models have already been formulated for disparity detectors, motion detectors and now for MID detectors. The same set of basis filters can be used for all of them. Thus, only a single set of convolutions needs to be performed. Ita is now compiling an overview sheet which summarized these filtering operations. These filters can then in year three in a unified way replace the currently used independent stereo and motion algorithms. At the same time Ita and Sco are preparing a neurophysiological review article on this field.

The diagrams below (Fig. 1 and 2) summarize the state of the project.

The ECOVISION system consists of 3 main parts: A) Harware front end for motion and stereo, where the latter is existing from older work. B) Motion/flow analysis part and C) Stereo/depth analysis part. **In the second and early third year the central goal is to use 1) motion and grouping to improve stereo analysis and 2) to use stereo analysis and grouping to improve motion analysis.**
Thus, this is a duplicate process where one visual modality serves as the target for improvement by means of the other modalities. We have achieved both of these goals almost completely, the missing parts will be addressed early in year 3 as planned. This shall be described below. In the third year these target modalities will finally be merged with each other.

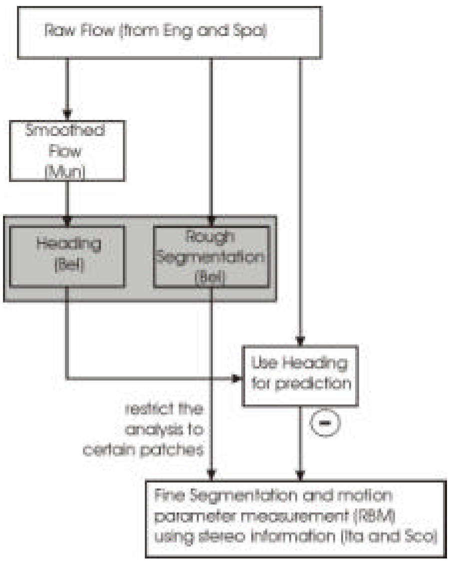


Fig. 1) The motion part of the ECOVISION system

Along the lines of motion analysis we had in the first year encountered an unexpected (overlooked) problem: Flow-field algorithms can only extract "normal flow" (Raw Flow in Fig. 1), hence flow which is orthogonal to the object's edge from which it arises. Thus, only at corners, flow is correct, but this way too few analysis points exist in an image. Thus, we needed to extract true and dense optic flow from the normal flow. This is now done in the following way which also nicely shows the cooperative efforts of this consortium. Raw Flow generated by Spa using FPGAs with an algorithm provided by Eng is smoothed by an algorithm developed by Mun, which uses the receptive field filtering of MT-cells for smoothing. From this we can extract very accurate heading information by an algorithms developed by Bel (described in the last report). The same algorithm can also be used to get a rough segmentation of the flow field into the different moving objects. Essentially this is done by clustering the different flow patterns after having subtracted the heading first. Heading information determines the rigid body motion parameters of each passive object in the scene (Sco). Thus, heading information can be used to predict their RBM. This part is still not working (indicated by the little minus-sign in the figure, work planned for year 3) but a work-around exists at the moment which allows us to finally arrive at a fine image segmentation (Ita and Sco, two different algorithms for this) of the different flow patterns from actively moving and also from passive objects (their flow comes only from the ego-motion/heading).

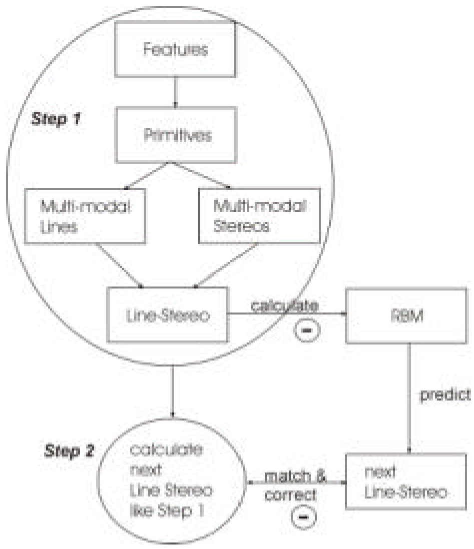


Fig. 2) The Stereo-part of the ECOVISION system.

The central idea behind the stereo part of ECOVISION is to perform multi-modal analysis in order to arrive at accurate stereo information and to further improve this by means of motion information. At the moment we can extract visual features and generate visual primitives from them (These represent a sparse multi-modal data representation of the most relevant image locations, heavily described in the first report, see there). Primitives can be combined into line-Gestalts by using multi-modal visual information to confirm a Gestalt. In the same way we are treating stereo matches: Only if enough multi-modal "support" exists a match is accepted. Now we can combine multi-modal stereo with the multi-modal lines to created lines in 3-D (line stereo). These are true parts of objects, this they can be used to calculate the RBM. This step is not yet working, it's the same step as above (Fig. 1), which does not work yet. With the help of a workaround, we are nonetheless already able to perform the prediction step. RBM can be used to predict the next line-stereo structure in the next camera frame pair. In step 2 we perform the same analysis as in step one but now we can compare the calculated with the predicted results. This step will be put to work early in year three. All these algorithms have been mainly developed by Sco in cooperation with Mun and Ita.

In a truly fruitful co-operation with Hella Hueck KG, new driving scenes have been recorded which contain steering information from car-internal sensors as well. This work was also not planned but has led to far better scenes. Some sub-aspects of this basic research project are now actually being considered at Hella for product development. First and foremost, the optic flow based rear-view-mirror-dead-zone-analysis algorithm (what a word….). This shall be described to some detail below.

Co-operative activities:

Sept. 2002 until Feb 2003, G. G. Botella (Spa) to A. Johnston (Eng) to learn McGM (Ecovision and Marie Curie training support.)

March 2003: One week work visit of N. Pugeault (Sco) to Hella (Ind) to record new camera sequences

May 2003 and August 2003 Visit of N. Krüger (Sco) to M. Lappe (Mun) to work on the problem of intrinisic dimensionality (May) as well as paper preparations on the same topic (August).

September 2003: Visit of Dirk Calow and Farid Kandil (Mun) to N. Krüger and N. Pugeault (Sco) about MT like filtering.

September 2003: Visit of A. Johnston (Eng) to E. Ros (Spa) McGM business ad supervising a PhD student.

November 2003 Visit of M.v.Hulle (Bel) to E. Ros (Spa) about Segmentation of flow fields.

S. Sabatini (Ita) cooperation with E. Ros (Spa) about implementation of stereo algorithms and paper preparation.

3rd ECOVISION MEETING in Granada in May with all participants.

4th ECOVISION MEETING in Genoa in October with all participants

Some examples of specific cross-fertilization effects in brief:

1) Ind provided camera sequences + auxiliary data for everyone.
2) Sco provided “intrinsic dimensionallity Concept” to Mun for MT filtering and to Bel for flow segmentation.
3) Mun sent range images to Sco and Sco sent primitives back to Mun for MT filtering using of the intrinsic dimensionality.
4) Eng and Spa tight cooperation of the McGM algorithm.
5) Ita provided Matlab Code of stereo scene analysis for Spa (to study FPGA implementations)
6) Bel provided flow segmentation to Mun for unbiased heading estimation.
7) Ita: On the web for all partners: Kalman filtering of Motion segmentation and disparity map from phase based filtering approach.
8) Eng uses Mun’s eye movement data for obstacle avoidance task.

There were many more such smaller and bigger exchanges in this truly lively consortium, which cannot all be listed.

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

In these WPs we have addressed how to implement flow-field algorithms on FPGAs. Several algorithms have been analysed how accurate and easy they can be implemented. They were: Simoncelli & Heeger, Lucas & Kanada, mult-channel gradient model (McGM) from partner Eng, and a simple flow-algorithm based on Reichard detectors. Note McGM is related to Lucas & Kanade but more complex and accurate.

All analysis steps have been finished as planned. From this we concluded to centre the implementation on the Lucas & Kanade algorithm and it is planned to extend this to the McGM in year three.

We note that actual FPGA *implementation* has not been foreseen in the work plan. Originally only a feasibility study was planned.

Here we can state that we are going far beyond this workplan, because at the moment we have already implemented the Reichard algorithm and the implementation of the Lucas & Kanade algorithm on FPGA will be completed about next February.

The Reichard algorithm was only meant to be a fall-back safety measure in case the others would have turned out to be too complicated. Thus, Reichard is simple but not very accurate. In spite of this, this algorithm has raised considerable interest at the hands of Ind, because its implementation is so cheap. A subgoal of Ind is to solve the problem of the dead-zone when looking into the rear view mirror. A simple motion (i.e. flow) algorithm together with flow-vector clustering could be enough to induce a warning signal as soon as a car approaches. Spa has implemented exactly such a system and Ind is now considering testing this. This may at the end lead to a viable industrial product, which was initially not at all expected from this FET open research project. Patents have been filed accordingly.

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

(from the first report:) This project part is centred 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) and their interaction

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.

WP3: Two approaches have been pursued in parallel: a) Kalman filter based approach for motion and depth and b) Primitive-base approach for moving 3D-Gestalts. The goal of these approaches is to arrive at real-world coordinates of linked object points. "Linked" here means that these points will belong to the same object, but we cannot name it, because object recognition is not foreseen in mid-level vision to which the ECOVISION concept belongs.

Why are we pursuing two approaches to the same goal? The reason for this is that the Kalman approach renders dense but less accurate motion and stereo maps, while the Primitive-based approach is sparse but very accurate. We hope that joining these approaches in year 3 will lead to dense *and* accurate motion and flow maps.

a) The Kalman filter approach uses flow-field templates of all possible basic flow-field patters (expansion, rotation, translation, etc.) and performs local template matching (this has been done during year 1). The template matching follows a relaxation process which is based on the Kalman filter algorithm. Thus, at this stage only single images had been subdivided into different flow-pattern patches. Now in year two the same process has been performed in a temporally continuous way (thus, now really making use of the temporal prediction properties of a Kalman filter). Thus, flow-patches are updated in a temporally predictive manner. The important and interesting aspect of the Kalman approach is that it can be used in the same way to regularize stereo maps. Thus from a raw stereo map we can now compute in a temporally ongoing way an improve map as shown in Fig. 5.
b) The Primitive based approach first extracts multi-modal image descriptors (Primitives). These primitives capture location, orientation, contrast-transition, colour, disparity and flow into one vector. They are only calculated at high energy locations in the images, thus, they are far less dense as compared to the original number of pixels. This step had been achieved early in the first year of ECOVISION. Primitives can be combined into groups. For example, a statistically very prevalent property of images is the predominance of collinear line segments. This property can be used to link Primitives with the same orientation. We have done this in year 1 and 2 of ECOVISION, but in order to raise confidence in any found collinear structure we have taken the other image information

which Primitives provide also into account. Only when several features (e.g. colour, flow, etc.) coincide with a collinear Primitive-pair we will raise its confidence value. This way multi-modal line segments have been extracted. The same multi-modal confidence mechanism can also been applied to stereo-pair. Thus, this way all stereo-matches are eliminated unless a matching pair also matches in respect to other image features. This algorithm has been written in year 2. Furthermore, we have now combined multi-modal lines and multi-modal stereo to retrieve the true 3-D structure of line segments in space (called line-stereo). Line-stereo information shall be used in year 3 to extract RBM information. Already now we can use such (still hand-computed) RBM information to predict the development of Primitives (or of line-stereo Gestalts) in a scene. This algorithm in functional since month 20. [Hand-computing here means that the tracking of five correspondences necessary correspondences in done by means of marking them with mouse-clicks.]

WP4: In WP4 we have developed a set of space variant filters which can be adapted to different desired resolutions in the visual scene (finished in year 1). These filters are now being used by Spa and Sco in the ECOVISION system. In year 2 we have focussed on methods to alter the spatial resolution of a visual scene in a task dependent way. This is achieved by spatial remapping which magnifies “interesting” parts, while shrinking less important ones. The rear-view mirror problem is a good example here. A mapping function has been designed to guarantee that an approaching car will always have roughly the same size in the image, regardless of its distance. This allows us to recognize the flow pattern of the approaching car already with high accuracy at a great distance and trace it distortion-free through the whole field of (rear-view) vision. This algorithm works fine and is now brought together with the low-level flow-algorithm implementations of Spa as planned for year 3.

WP5: Task 5.1 has after contract amendment (following the move of Prof. Lappe to Münster) now also been completed. Here we where measuring the eye movements during different tasks (heading vs. obstacle avoidance) during driving. These data have influences the image remapping in WP4. At different remapping is suggested during a heading task as compared to an obstacle avoidance task.

Task 5.2 addresses the question of how to optimally extract motion information during different tasks. Here we have in year 2 focused on the heading task. It has been discussed above that flow-field maps are extremely noisy and erroneous because of the aperture problem. Partner Mun has now successfully used a neuronal filtering strategy for improving these maps. By using receptive field filters with a structure similar to that of MT-cells the flow-maps are being corrected and heading is extracted in a very accurate way. To extract heading we are using an algorithm developed last year by Bel. This is a nice example of a new neural technique that could in a rather direct way be transferred to technical image analysis. Once heading is extracted the optical flow which results only from the heading can be “subtracted” and the remaining flow patches can be analysed independently. This step is planned for year 3 also using the algorithm of Bel.

Project Part 3: Computational Neuroscience (WP6)

The computations neuroscience part has been criticized during the last review as being not strong enough. Since the document follows the WP structure only WP6 should be described here. We will however now deviate from this structure, because we are happy to state that neuroscience aspects have arisen in WP3 and WP5 as well which have clear impact on the technical implementation of ECOVISION and/or which provide interesting experimental predictions.

In general we have put substantial effort into the neuroscience part in year 2.

1) We have defined MT-filters for Flow-field smoothing and actually use them to improve heading analysis (already described in WP5 above) [Neuro→Techno]
2) We have made a model if motion-in-depth detectors in V1 based on energy models of cortical cells. This model was inspired by the need to think about detectors that use depth information to perform motion analysis. WP3.3. is centred at the question of how such cross-modal interaction (stereo "helps" motion and vice versa) can be used to improve image analysis. Thus here a technical demand has lead to a new neuroscience model [Techno→ Neuro]. It is, however, also foreseen to use these new receptive field filters in the ECOVISION system in year 3 to improve the Primitive representation.
3) We have designed a model of low-level attentional processing in V1 and LGN. This was the main goal of WP6. The model introduced in Deliverable 6.1: CNS-algorithm for generating Task-Optimised Spatial Representations of Neural Activities, is extended in order to fully explain the experimental results obtained by the lab of Vanduffel et al. [2000]. This model can be used as the basis for task/attention centered remapping in WP4 [Neuro→Techno]. Contrary to the model introduced earlier, this novel model is not restricted to area V1 but includes regions from the thalamus as well. In this way, novel insight is provided with respect to the interaction between contextual effects and attention in the earliest stages of visual processing, in particular, the emergence of suppressed activity surrounding the center of interest (CoI). The dynamical model is recurrent in nature and includes all regions in which attention effects were observed in the physiological experiments, namely in thalamic nuclei lateral geniculate nucleus (LGN) and reticular thalamic nucleus (RTN), and layers 4Cα and 6 of primary visual cortex. The model is capable of reproducing all observed effects by means of a novel mechanism: the diffusion of stimulus driven relay cell activity to RTN (both directly and via V1) and a subsequent injection of inhibition from RTN to LGN relay cells in regions of the LGN surrounding the stimulus representation. In view of the application of this model in the ECOVISION system we observe that in this model, task-relevant information regarding the CoI (either spatial or featural) can now modify the `driving input' at the earliest visual processing stages (task-specific representation), rather than modifying the actual processing. Through dynamic feedback mechanisms, the signal-to-noise ratio of task-relevant features (e.g., motion or disparity) which are computed higher-up in cortex, can be greatly improved.

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.

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 had finished in year 1, but we have recorded several gigabytes of better controlled driving scenes in year 2 to improve on the quality of the data. These scenes also contain other controls, like steering angle of the wheel and velocity.

WP8 concerns benchmarking and testing, mainly to be done in year 3. First steps have been performed by Spa and Sco in testing their motion algorithm implementation against the driving scenes provided by Ind.

WP9 has also started. This is mainly a cooperative effort where we will try to bring the different components of ECOVISION together to build a complete system. In year 2 we have mainly focused on issues of how to actually do this. Currently we can state that, very likely, we will be able to put together a physical demonstrator with FPGA-front end and ECOVISION analysis software to produce advanced image analysis in year 3. This was not planned, but the fact that the FPGA implementations already work to a large degree will permit this extra step.

We are seriously considering dropping WP5.3 and using the free resources to create such a demonstrator. A decision about this shall be made early in year 3. Here we would also ask our reviewers to provide us with some external views on such a possible change of plans.

Project Part 5: Administrative Workparts

WP10 Dissemination
WP11 Management and Evaluation

Scheduling: ongoing
Administrative aspects: Several smaller budget shifts have been made and approved by the coordinator. The work has not been delayed by the two major contract amendments that had been performed end of year 1. 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 Personnel.

Thus in all respects the administration of the ECOVISION project proceeds without problems.

Co-operations and Meetings: See pg. 6 of this report.
Publications: See Deliverable 10.3, attached to the end of this report.

Dissemination: Dissemination has mainly been performed via publications and conference attendance. The FPGA work of Spa has received special coverage in newspapers and the Spanish TV, too.

World-wide State of the Art Update

The central question concerning a state of the art update is: Have there been any advances in the field which would force the ECOVISION consortium to restructure the work. In answer to this we can state that none of the papers published by other groups in the year 2003 represents a challenge to the ECOVISION system.

This concerns the FPGA side of the project, where in the last year only one motion estimation algorithm implemented on an FPGA has been published (Rueda and Estrada 2003). But this was very preliminary work without any results and the computational resources exceeded the FPGA capabilities. In ECOVISION we have already a fully functional system and a second one ready by next spring.

Concerning the use of Gestalt laws in computer vision nothing has been published in 2003 except for one article (Desolneux et al. 2003) which has become available to us only as a preprint. This paper uses a different Gestalt approach towards grouping and applies different Gestalt laws in an exemplary manner. Little is done to try this approach on real image sequences. In spite of this, this paper is nice support for the ideas of ECOVISION.

Thus, at the moment we are convinced that ECOVISION represents the state of the art with no serious competing approach in its specific sub-fields.

Selena Maya Rueda, Miguel Arias Estrada, FPGA Processor for Real-Time Optical Flow Computation. LNCS 2778, pp 1103-1106, 2003, Springer-Verlag Berlin Heidelberg.

Desolneux, A., Moisan, L., Morel, J-M. (2003) Gestalt theory and Computer Vision" preprint CMLA N° , to appear

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

Motion Algorithm

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.

The goal of this WP is to find a flow-field algorithm well suited for FPGA implementation.
Schedule: This WP finished on schedule in month 15.
Central Achievements:

- D1 (due month 13). A working integer version has been implemented on a Matrox board.
- A low-level simulator of Eng's algorithm (McGM) has been implemented to evaluate the bit-cutting procedure.
- The implementation feasibility of two algorithms have been studied: Lucas & Kanade approach and McGM.

Planned and performed steps:
Step1)

A type of motion algorithms called "gradient models" have been chosen as good candidates to be implemented on specific hardware. Among all the approaches available of the shelf we have concentrated on the Lucas & Kanade Algorithm (whose implementability is fairly obvious from its structure and Eng's algorithm (that requires a higher computational cost but provides better accuracy).

Step 1a) Assessment of Eng's motion algorithm (called Multi-channel Gradient Model) if it will be in principle feasible for an FPGA implementation.

Schedule: Finished in month 13.
Performed actions: Eng and Spa collaborative work with data exchanges and lab visits.
Results: The Multi-channel Gradient Model (McGM) was found to be feasible but of high computational cost. Therefore, due to its high accuracy and robustness its implementability study goes on, but other simpler model needs to be considered to assure the FPGA implementation of a complete motion algorithm.
Documentation: Internal Technical Report: *The Multi-Channel Gradient Model and its Real-Time Implementation*.
Publications: not in this second year period

Step 1b) Assessment of other motion algorithms
Simoncelli and Heeger's motion algorithm:

Schedule: Finished in month 13.
Performed actions: Ita and Spa collaborative work.

Result: The algorithm is highly regular and parallel but it requires a high computational cost and it is contrast-dependent which a crucial drw-back in real-world, open-air (automotive) sequences.
Documentation: Short Report: *Some notes on the Simoncelli and Heeger algorithm*
Publications: Spa1, Spa3

Lucas and Kanade motion Algorithm:
Schedule: Finished in month 13.
Performed actions: Spa study and state of art update.
Result: The algorithm is of relative low computational cost and it achieves good results. Therefore it has been chosen as a very valid candidate to be completely implemented on a FPGA.
Documentation: none
Publications: none

Step 2)
Actual mapping of the integer version onto the FPGA architecture study and accuracy estimation.
Schedule: finished in month 15.
Performed actions: Bit cutting procedure simulation version of the McGM and computational resource study. Lucas & Kanade implementation with a hardware friendly C code.
Results: Some interesting experiments with different bit cutting strategies have been done. Also the computational and storage requirements of the McGM have been evaluated. The Lucas and Kanade algorithm has been implemented in C to evaluate its computational cost, and a first block diagram and pipeline stages have been evaluated.
Documentation: Internal Technical Report produced by Spa and Eng: *McGM algorithm: Some results with different bit cutting strategies*. Short Report produced by Spa: *Implementability of the Lucas & Kanade*
Publications: none

Step 3)
Quantification of the accuracy of the different cutting depth allowed by the algorithms with the sequences provided by the Ind partner.
Schedule: finished in month 15
Performed actions: The sequences provided by the Ind partner (overtaking sequences in diverse weather conditions with a standard CCD camera and a High Dynamic Range camera) have been used to evaluate the performance of the motion estimation algorithms under bit depth restriction constraints.
Results: The bit cutting strategy used in different stages of the McGM algorithms has a strong impact in the final motion accuracy. Therefore the multiple-stage bit cutting strategy needs to be studied in detail and optimised globally.
Documentation: Internal Technical Report produced by Spa and Eng: *McGM algorithm: Some results with different bit cutting strategies*.
Publications: none

Deliverables:
D1 was produced on month 13.
D1 consists of a software version of the motion/flow-field algorithm with well defined bit-depth for every operation. A computer program and a Technical Report/Documentation have been produced.

WP2

FPGA simulation

After the results obtained in WP1, the work carried out in WP2 addresses the complete simulation (and implementation) of a motion/flow field algorithm onto an FPGA platform. We have focused on the Lucas & Kanade algorithm that requires less computational resources than the McGM.
The motion estimation algorithm has been implemented through a Hardware Description Language (HDL), concretely Handel-C. This high level description language facilitates the prototyping and design change.
The simulations of the complete motion/flow-field algorithm have been carried out within the DK2 environment of Celoxica.
With the simulations have been studied different possible projections onto an FPGA platform to optimise the FPGA functionality. The speeding up and silicon cost have been evaluated with different processing configurations. Therefore in WP2 we have addressed the following points:

1. Define technical specifications (gates consumption, embedded memory resources, other specific purpose blocks such as multipliers, etc), that may be crucial for the algorithm operability.
2. Identify the crucial steps and evaluate their relevance in terms of silicon cost and processing speed.
3. Study different processing schemes and their implications in the final efficiency.

In this WP we have gone one step forward the original ECOVISION proposal. After HDL code has been finished we have not only simulated it, but also implemented (or compiled) onto a FPGA platform. Concretely, we have implemented it onto a PCI board with a FPGA device of 2000 million gates. This allows us the better study of the implementation performance and specification.

Schedule: This WP is on schedule and will finished on month 27.
Central Achievements:

- A “Fully Functional HDL simulation” (corresponding to D2) will be available in month 25.
- The implementation of the HDL code onto a FPGA prototyping platform for the specifications study.
- The real-time motion processing unit feasibility assessment.

Publications: publications about the work of WP2 have been intentionally delayed to next period in order to better proceed with the patent procedure (protection of knowledge).
Patents: One patent protecting the implementation of motion algorithm is currently in process.

Performed steps:
Step 1) definition of the technical specifications that may be crucial for the algorithm implementability.
Step 2) identification of the crucial stages in terms of silicon cost and processing speed.
Step 3) Simulation of the HDL code to address the study of different processing configurations and their implications in the final efficiency.

WP3

Adaptive receptive fields as context sensitive visual filters – Gestalts in space-time.
The final goal of this WP is to provide a context-based adaptive analysis of a three-dimensional (3D) dynamic scene. The solutions proposed at the end of the first year have allowed us to point out, through data-driven processes, 3D visual segments as well as structural regularities in optic flow fields. On this ground, the research activities of the second year, mainly conducted by Sco and Ita (in collaboration also with Mun) have consolidated and extended the approach to consider multimodal integration of stereo and motion information. Specific results achieved in this WP are: (1) specification of recurrent statistical processes for disambiguation of middle level processing such as stereo, grouping and ultimately 3D reconstruction; and (2) the specification of early motion-in-depth detectors built on computational resources available at early vision stages (cf. simple and complex cells in the primary visual cortex), suggesting a possible unifying phase-based framework to disparity, motion and motion-in-depth measurements by a single set of cortical-like filters.

<u>Task 3.2: Visual templates with spatio-temporal context - Gestalts in space-time.</u>

a) Statistically defined stereo-gestalts for motion computation

In Task 3.1 has been proposed a middle level image representation constituted of a sparse map of primitives. Those primitives are extracted from local measurements at high energy points (according to the monogenic signal). We have shown that stereopsis can be processed in this sparse space, and that a 3D entities can be reconstructed from the stereo-correspondences.

Due to geometric distortion between both cameras the local manifestation of a 3D feature might be widely different between both views. Also, a given primitive can have a large number of potential stereo-correspondences. Finally the local representation of a 3D feature is intrinsically ambiguous, due to repetitiveness in natural sequences. All this leads to ambiguity in the stereopsis, and to a number of erroneous 3D entities. Those outliers are often difficult to identify by local methods (although global heuristics can remove number of them). The grouping process faces the same kind of uncertainty, if at a lesser level, due to ubiquitous noise and fine (i.e. finer than the present scale) texturing in natural sequences. Those uncertainties cause problem for the RBM computation, where reliable 3D lines are desired.

To handle those decisions more finely, we have developed a probabilistic framework joining dynamic grouping and stereopsis. All potential stereo correspondences are considered and their likelihood is evaluated, according to the stereo matching defined in task 3.1. Also all potential gestalts are being considered and evaluated, according to their compliance to the gestalt laws. From the confidences in those individual events and their consistency, those hypothesis likelihoods are re-evaluated, down rating the inconsistent ones and enhancing consistent groups of hypothesis.

To evaluate this consistency over stereo and gestalt assumptions, we defined a Basic Stereo Consistency Event (BSCE, Fig. 3), being the minimal possible case of confirmation between several of our hypothesis. By integrating this consistency over a primitive neighbourhood, we get an estimate on how a stereo hypothesis is consistent over the potential gestalt structures involving this primitive.

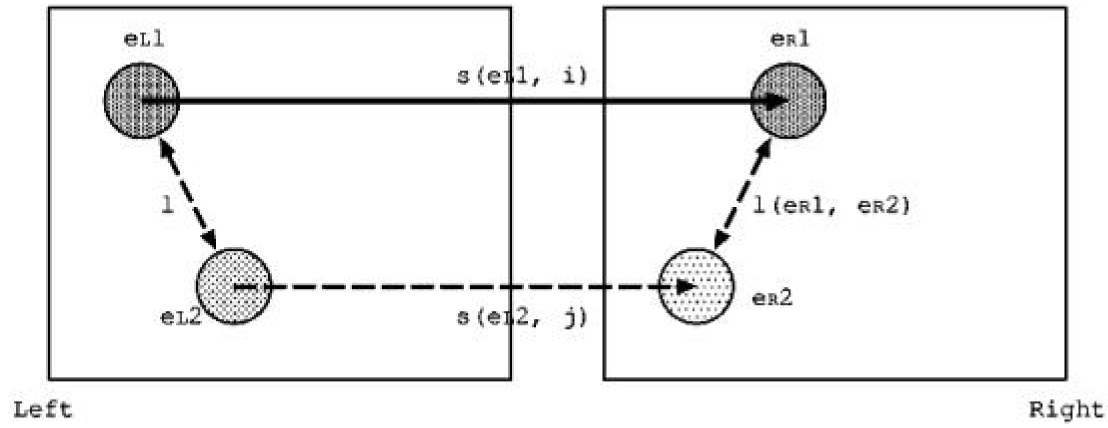


Fig. 3) Basic Stereo Consistency Event: in this case if we consider the hypothesis that eR1 is the stereo-correspondence of eL1. If eL2, part of a common putative gestalt l with eL1, has a candidate correspondent eR2, and if a gestalt l' also exists between eR1 and eR2, then we consider that those 4 stereo-gestalt hypothesis are consistent, and confirms each other.

The effect of this process on stereo has been shown, through the improvement of the reconstructed 3D-entities (see figure 4).

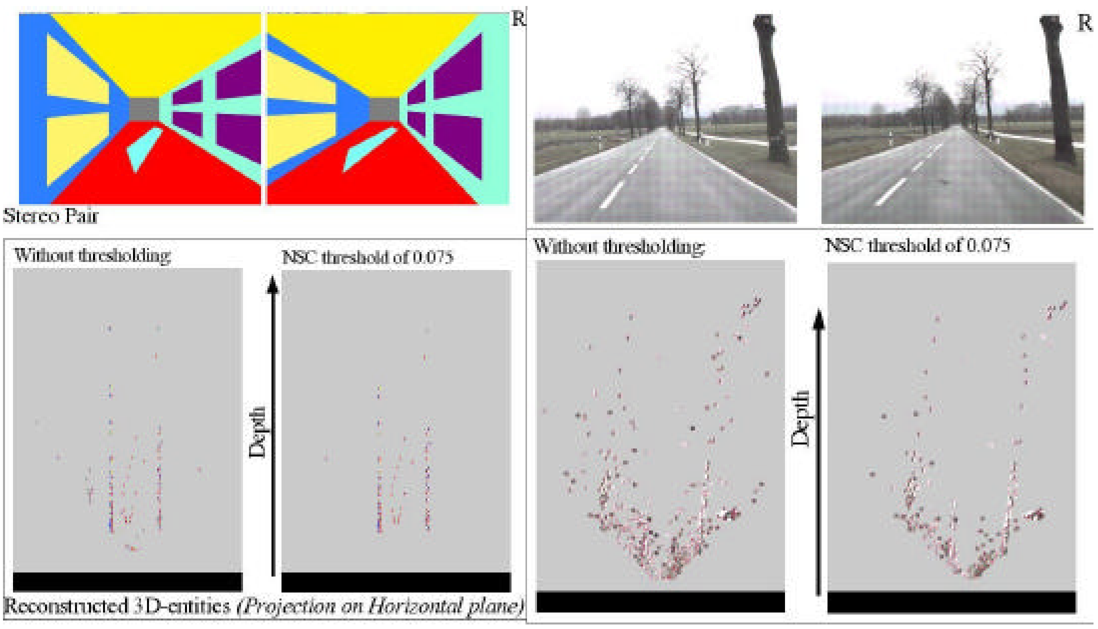


Fig. 4) In both figures, the 3D entities are reprojected using only the most likely stereo correspondence of the primitive. The lower figures show an orthographic reprojection of those on the horizontal plane (the depth is the horizontal axis, and the x axis is vertical). The left one shows the original reprojection, using only local multi-modal information for discriminating between the hypotheses. The right one shows the same reprojection after applying a threshold on this stereo-gestalt consistency. The left hand side features an artificial test sequence. Here all erroneous 3D entities are being removed. The right side shows the same process applied to one of the sequences recorded with partner Ind in March.

Note that the same threshold as for the first sequence has been applied here, removing a large number of noise, demonstrating the stability of the process over widely different

cases.
The interaction of stereo and gestalts allow here to identify erroneous correspondences where local information is ambiguous. This process allows not only to improve our local processes of stereo and grouping, but also to create more complex and reliable structures: 3D pseudo-gestalts. More specifically they allow us to identify reliably the 3D line structures, crucial for RBM computation.

b) Kalman-based context sensitive filters based on spatio-temporal Gestalts (Ita)
According to the general framework developed in Task 3.1, the structural constraints necessary to model spatial context (Gestalts) are embedded in a lattice network with proper anisotropic interconnections and boundary conditions. Such networks represent "Markovian" affine models for structural motion regularities. For each combination of the cliques and of the boundary conditions, the network reaches an equilibrium state that corresponds to one of the elementary flow component (EFC) that can be used to approximate the 1st-order structure of the optic flow. For each EFC we need a specific Markov network. To smooth the variation of affine transformations over time a relaxation term has been introduced. On this basis, the resulting context sensitive filters (CSFs) (1) represent multiple competitive hypotheses in space and (2) integrate information over time. The results obtained on real-world sequences provided by Hella (see http://www.pspc.dibe.unige.it/ecovision/private/index.html) have evidenced the good efficiency of the Kalman-based CSFs, even in the case where the hypothesis of white Gaussian noise does not apply, suggesting that a "real" noise does not affect de facto the filters' performances. It is worth noting that the inclusion of time in the Kalman filter (KF) equations is the natural extension of what developed in Task 3.1. Such an extension, indeed, meets the temporal iterative character of the KF, which continuously adapts to the new visual data by exploiting the whole past history of the input signal.
By exploiting the same formalism developed for motion Gestalts, we have investigated the possibility of regularizing the dense (but rough) disparity maps obtained by phase-based techniques (Fleet and Jepson, CVGIP 53(2):198-210, 1991; Solari et al, Elect. Letters 37(23):1382-1383, 2001). More specifically, by casting a spatio-temporal continuity constraint of binocular disparity as the KF's model equation we have obtained depth map estimates, which are regularized in both space and time. Figure 5 shows snapshots of the results obtained by the context-sensitive disparity filter, on real-world sequences, in comparison to the initial non-regularized rough data.
The results of these activities have been partially published (see list at end of report).

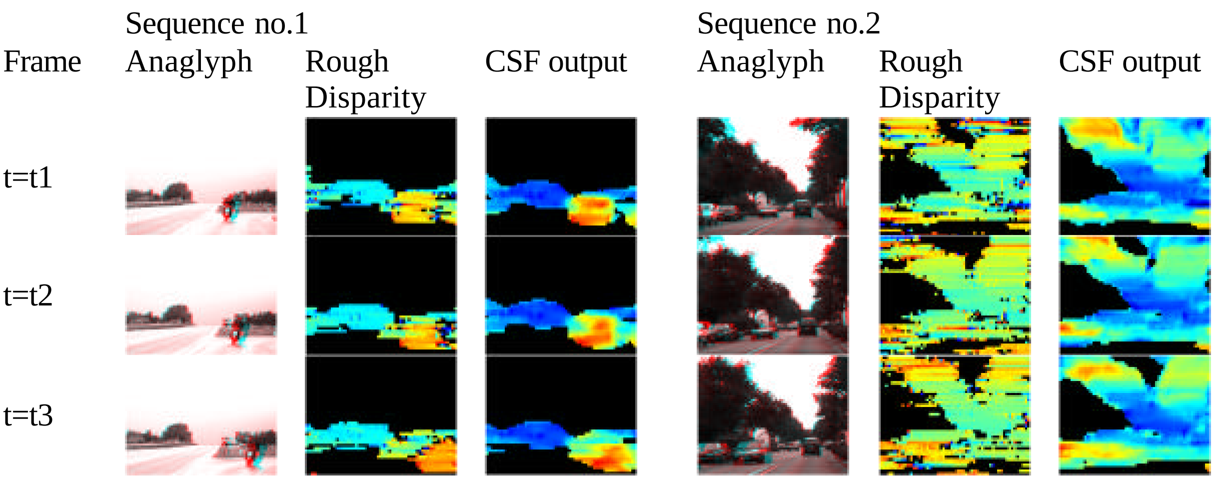


Fig. 5) Examples of regularized disparity maps obtained by Kalman-based context sensitive filters on real-world driving sequences. The rough disparity maps have been

obtained by a modified phased-based technique (Solari et al., Elect. Letters 37(23):1382-1383, 2001). Red-coding for near objects and blue-coding for far ones. Black pixels represent points discarded according to a confidence measure.

Task 3.3: Cross-Modal interactions between Gestalts define the first level of recursive scene analysis.

a) Spatial – temporal gestalts for stereo disambiguation

Based on the multi-modal Primitives (the Primitives as well as their biological motivation by hypercolumns are described in (Krüger et al., 2003)) we have developed temporal-spatial Gestalts. Rigid body motion (RBM) is the underlying regularity that binds Primitives derived from single frames together (a detailed review about RBM, its formalisation, estimation and utilisation as well as its potential combination with grouping processes is described in (Krüger and Wörgötter, 2004)). The process that leads to temporal-spatial Gestalts is schematically described in figure 6. In this scheme the change of visual entities across different frames is predicted and correspondences lead to an increase of confidences (while non-correspondences lead to a decrease in confidences) as well as to an interpolation of parameters of entities. This scheme has already been described in the last report and in (Krüger et al., 2002b). However, this scheme has not been used within our framework of multimodal Primitives. This we have achieved now. By doing this we have realised some problems that needed to be solved as described below.

We will first give a short description of the application of the scheme, the specific problems we have encountered as well as their solution. Then we will describe the results that we have achieved.

Formalization of Spatial-Temporal Gestalts and their Utilization for Disambiguation of Stereo Information

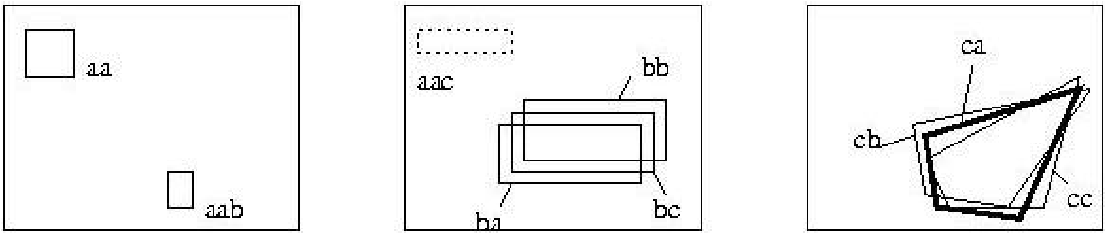


Fig. 6) The accumulation scheme: the entity e^1 (here represented as a square) is transformed to $T^{(1,2)}(e^1)$. Note that without this transformation it is barely possible to find a correspondence between the entities e^1 and e^2 because the entities show significant differences in appearance and position. Here a correspondence between $T^{(1,2)}(e^1)$ and e^2 is found because a similar square can be found close to $T^{(1,2)}(e^1)$ and both entities are merged to the entity $\hat{e}^2$. The confidence assigned to $\hat{e}^2$ is set to a higher value than the confidence assigned to e^1 indicated by the width of the lines of the square. In contrast, the confidence assigned to e'^1 is decreased because no correspondence in the second frame is found. The same procedure is then applied for the next frame for which again a correspondence for e_1 has been found while no correspondence for e'^1 could be found. The confidence assigned to e^1 is increased once again while the confidence assigned to e'^1 is once again decreased (the entity has disappeared). By this scheme information can be accumulated to achieve robust representations.

This scheme is of rather generic nature. However, for its application a number of crucial details were to be solved:

Change of entities across frames: The transformation of entities across frames must be formalized. In our scenes the dominated change is caused by ego-motion which can be described by an RBM (see the circled 1 in figure 7). However, the RBM can only be applied directly to 3D entities and not to image entities such as the Primitives (see also (Krüger and Wörgötter, 2004)). Therefore 3D entities must be reconstructed from different perspective views of a stereo pair of images (see the circled 2 in figure 7). For this we make use of the different modalities coded in the Primitives. Furthermore, after applying the RBM to the 3D entity, this entity must be reprojected to the stereo image pair of the second frame to be comparable to the extracted Primitives. That means that beside the formalisation of the change of entities during an RBM also the reconstruction and reprojection problem needs to be addressed (see the circled 3 in figure 7). Reconstruction is done from stereo correspondences that have been found by using a multi-modal stereo matching that makes use of all aspects coded in the Primitives. We have also investigated the importance of the different aspects for stereo matching (see (Pugeault and Krüger 2003, Krüger et al. 2002a, Krüger and Felsberg 2002)). Reprojection addresses beside the geometric information also all other modalities coded in the Primitives.

Comparison of entities: When we want to find correspondences of transformed entities and Primitives in a frame a comparison of entities according to some metric is required. Here we have a couple of choices that may lead to quite different results. For example we can perform a comparison of 3D entities, i.e., we formalise the process in Euclidian space. However, we found out that such a formalisation (as done, e.g., in (Krüger et al. 2002b)) leads to problems since reconstruction acuity depends on depths, i.e., we would need to apply a metric in an inhomogeneous space. Entities that have large depths would tend to find less likely correspondences then entities that are close to the camera. The solution to this problem that we have chosen is a formalisation of the metric in the image space in which errors reflect a more homogenous behaviour (see the circled 4 in figure 7).

Handling of different Modalities: The visual Primitives carry beside the geometric information position and orientation also non-geometrical information such as phase and colour. However, since Rigid Body Motion only describes only the change of the geometric components we need to approximate the change of phase and colour. Furthermore, in the comparison step (see above) we have so far only used position and orientation. However, the comparison becomes more efficient when we also use the other modalities. For example a transformed red/green edge might be similar in orientation and position to an extracted Primitive but very different in its colour attributes and should then not be seen as a correspondence. Therefore, we now use a comparison that takes all these modalities into account (see the circled 4 in figure 7).

Update Rule: When a correspondence has been found it needs to be decided how the parameters of the entities influence each other. Moreover, in the scenes entities can be out of frame after a motion when the objects have been passed by the camera. The naive application of the scheme described in figure 6 would lead to a decrease of confidences of such hidden entities and valuable already generated knowledge would be lost. We therefore adapted the scheme to these out-of-frame situations such that entities for which the position is predicted as being out of frame are not altered once they have achieved a certain confidence.

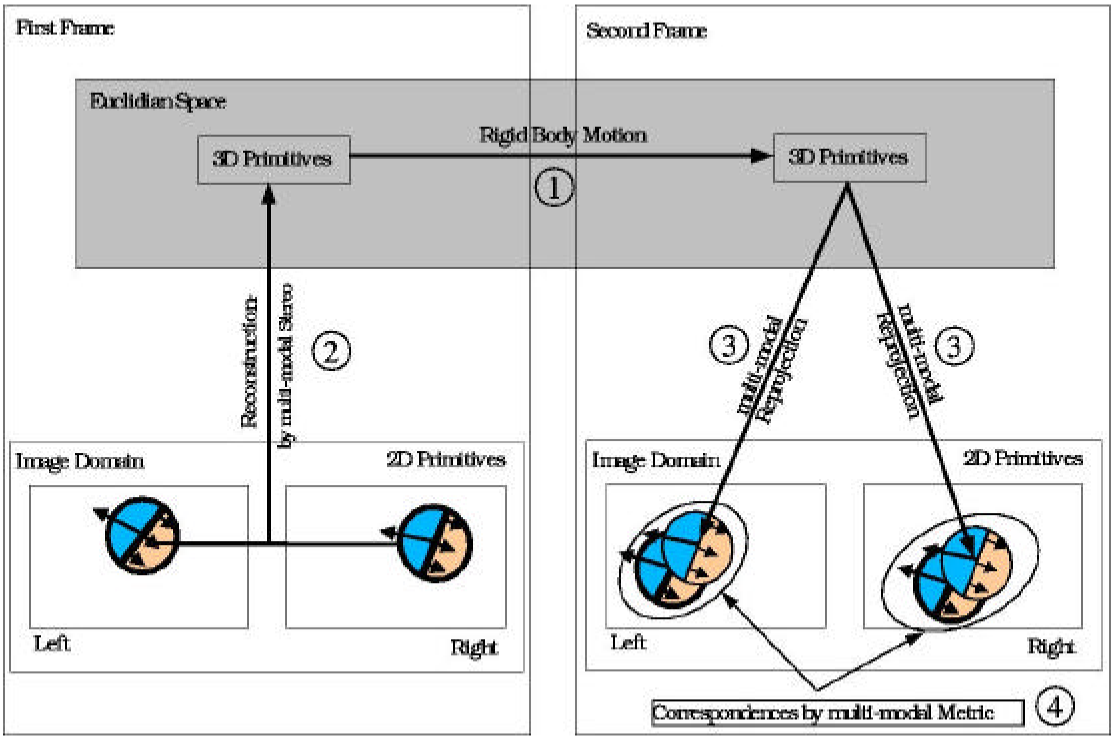


Fig.7) A more detailed description of an iteration of the scheme shown in figure 6 that points to problems concerning specific subaspects.

Results:
We have applied spatial-temporal Gestalts to stabilize ambiguous stereo information for artificial and natural scenes. The figures 8 and 9 show the results.

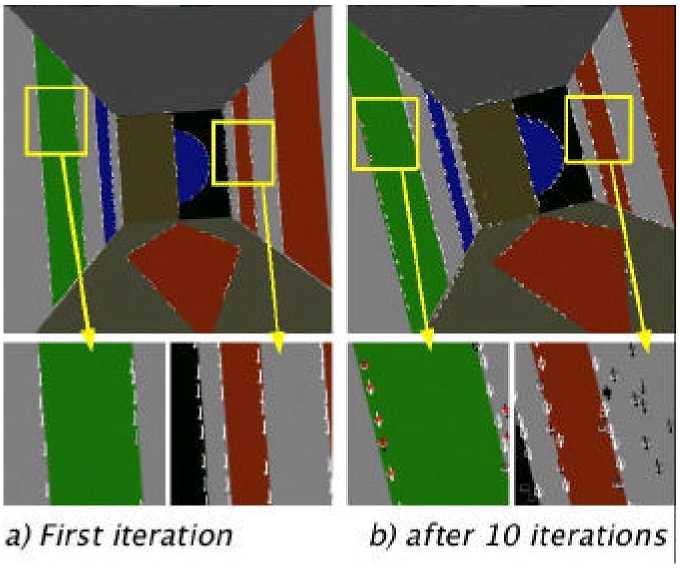


Fig. 8) Development of spatial-temporal Gestalts across frames for the artificial sequence for frame 1 and 10. The RBM in the artificial sequence is described by a translation that has significant term in x, y and z direction and a rotation around the z axis. All visual spatial-temporal Gestalts that have been generated in the process described here are shown on the left side while only the spatial-temporal Gestalts with high confidence are displayed on the right side. All displayed entities have been developing over different frames by transforming the entities according to the computed RBM and updating according to the scheme. The grey level of the orientation bar

represents the confidence that has been accumulated over time. Entities which display a black square are "dead" entities, i.e. entities that have not been updated for a certain number of frames. Red entities are newly generated entities, i.e. entities for which no correspondences have been found and therefore a new hypothesis has been created.

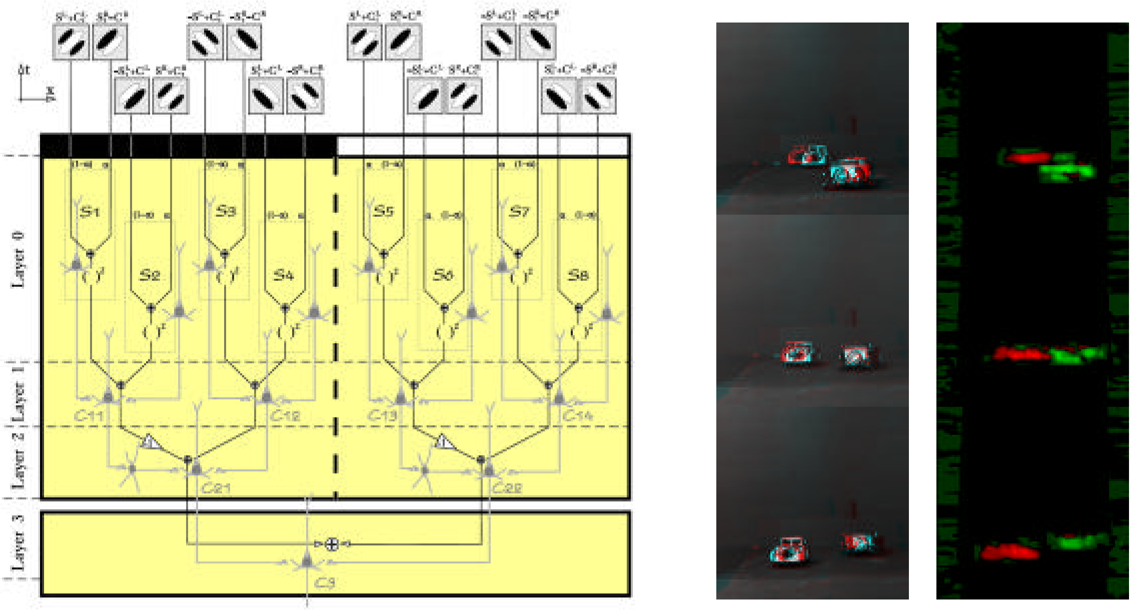
Fig. 9) Development of the spatio-temporal Gestalts across frames and projection on the horizontal plane of the reconstructed 3D entities.

b) Early motion-in-depth detectors

When the stereopsis problem is extended to include time-varying images, one has to deal with the problem of tracking the monocular point descriptions or the 3-D descriptions which they represent through time. Therefore, in general, dynamic stereopsis is the integration of two problems: static stereopsis and temporal correspondence (Jenkin and TsoTsos, CVGIP 33:16-32, 1986). Considering the stereo dynamic correspondence problem, a moving object in the 3-D space projects different trajectories onto the left and right images. The differences between the two trajectories carry information about motion-in-depth. From this perspective, dynamic stereopsis implies, in general, the knowledge of the position of objects in the scene as a function of time. We have demonstrated that motion-in-depth estimation can be obtained on a local basis and without tracking, from cyclopean motion energy measurements.

Psychophysical studies evidenced that there are at least two binocular cues that can be used to determine the motion of an object toward or away from an observer (Harris and Watamaniuk 1995, Vision Res. 35:885-896). First, the visual system might use binocular combination of monocular velocity signals. Second, the motion-in-depth signal could arise by the rate of change of retinal disparity. The predominance of one measure on the other one corresponds to different hypotheses on the architectural solutions which evolved, by nature, in the visual cortex. More specifically, the question of the degree to which the brain mechanisms for detecting motion-in-depth are independent of the mechanisms for detecting static disparities has an important bearing on the issue of whether the brain adopts a specialized non-Fourier motion channel to compute motion-in-depth, or relies upon the hierarchical combination of the motion signals conveyed by the Fourier channel (Chubb and Sperling, 1988 J. Opt. Soc. Am. A 5

1986-2006). We demonstrated analytically that information hold in the interocular velocity difference is the same of that derived by the evaluation of the total derivative of the binocular disparity, if a phase-based disparity encoding scheme is assumed. This result could partially resolve the apparent dichotomy on the structural mechanisms underlying cortical perception of motion-in-depth, suggesting that both static disparity and motion-in-depth computations could be based on the same computational primitives, interpretable as spatiotemporal receptive fields of simple and complex cells in V1. In particular, we have demonstrated that binocular energy complex cells relay phase temporal derivative components that can be combined, at a higher level, to yield a specific motion-in-depth (MID) selectivity.
On this basis, an architectural cortical model for MID selectivity in the visual cortex is proposed. By hierarchical combinations of the same signals provided by spatio-temporal frequency channels, the resulting cortical units actively eliminate sensitivity to a selected set of parameters, thus becoming specifically tuned to different features, such as disparity but not MID, or MID but not disparity. The emergence of MID tuning is pointed out in relation to the unbalanced ocular dominance of the afferent binocular contributions (see Fig. 10a).
From a computational point of view, the chain rule in the evaluation of the temporal derivative of phases allows us to gain information about MID directly from the convolutions of the two stereo images with complex spatio-temporal band-pass filters. This formulation eliminates the need for an explicit trigonometric function to compute the phase signal, thus avoiding also problems arising from phase unwrapping and discontinuities. Moreover, the approximation of temporal derivatives by temporal filtering operations yields to regularized solutions in which noise sensitivity is reduced (see Fig. 10b).
The results of these activities have been partially published (see list at end of report).



(a) (b)

Fig. 10) (a) Functional representation of the proposed cortical-like architecture for early motion-in-depth detection. Each branch groups cells belonging to an ocular dominance column (black and white slabs in the figure). The afferent signals from left and right *ocular dominance columns are combined in layer 3. The basic units are binocular simple cells tuned to motion directions (s1, …, s8). The responses of the complex cells in layers 1, 2 and 3 are obtained by linear combinations of the outputs of those basic units. (b) Experimental results on a natural scene. Two toy cars are moving in opposite directions respect to the observer. Left and right frames at three different times are shown. The color in the MID maps codes the motion-in-depth of the two cars: the red blob represents the car moving toward the observer, whereas the green blob represents the car moving away. Black pixels represent points discarded according to the confidence measure.*

Task 4: Context-based confidence maps

The results obtained by this task are the whole set of data representing the outputs of the context sensitive filters on the benchmark sequences. Such data, shared by all the members of the consortium, provide highly confident stereo and motion feature to be used by WP4 and WP5.

Presently on the web:

? Kalman-based CSF outputs 1, 2003-02-07 by S.P. Sabatini and F. Solari [4394 KB]

? Kalman-based CSF outputs 2, 2003-02-07 by S.P. Sabatini and F. Solari [1690 KB]

References

N. Krüger and M. Felsberg. An explicit and compact coding of geometric and structural information applied to stereo matching. *Submitted*

N. Krüger, M. Felsberg, C. Gebken and M. Pörksen. An explicit and compact coding of geometric and structural information applied to stereo-processing. *Proceedings of the Workshop 'Vision, Modeling and VISUALIZATION 2002'*, 2002a

N. Krüger, T. Jaeger and Ch. Perwass. Extraction of object representations from stereo image sequences utilizing statistical and deterministic regularities in visual data. *DAGM Workshop on Cognitive Vision*, pages 92-100, 2002b

N. Krüger, M. Lappe and F. Wörgötter. Biologically motivated multi-modal processing of visual primitives. *The Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour*, 1(5), 2004.

N. Krüger and F. Wörgötter. Statistical and deterministic regularities: Utilization of motion and grouping in biological and artificial visual systems. *Advances in Imaging and Electron Physics*, 131, 2004.

N. Pugeault and N. Krüger. Multi-modal matching applied to stereo. *Proceedings of the BMVC 2003*, 2003

WP4

WP4 is now in its second phase: WP4.2-General space-variant mapping methods. The objective is to develop transform techniques capable of mapping the raw image acquired by the camera into a new, more biologically plausible image, that is, the image as it could be represented in the visual cortex. It is expected that the new mapping will be more amenable to high-level information extraction by making more explicit meaningful relations that are not apparent in the raw image data. WP4.2 will therefore provide a useful pre-processing step to the motion computation of WP1. WP4.2 should also facilitate the extraction of visual cues for self-motion in WP5.

<u>WP4.2.1 – Space-variant filters</u>

Before space-variant mappings can be explored the problem of space-variant filtering has to be addressed. The straightforward method, that is, convolving the image with a specific kernel for each pixel would lead to prohibitive amounts of computation and memory storage making any FPGA implementation all but impossible.

A first method that can be described as 'steering in scale' was tried. It involves Taylor series of the scale variable and using the diffusion equation to replace derivatives in respect to scale by the more usual spatial derivatives (Florack et al., 1996). The range and accuracy of the extrapolation achieved with this technique was however not sufficient to make it a practical solution.

An alternative method, recursive filtering (Deriche, 1992), was investigated and found to possess very attractive features. Its main strengths are its ability to approximate Gaussian kernels of any size with a fixed number of operations per pixel (24 multiplications) and its high degree of accuracy when compared with the output of real Gaussian convolutions. The method is general in that the most widely used filters in computer and biological vision can be decomposed as a Gaussian convolution, that our method allows to be space-variant, and another filter (such as a derivative) that does not need to be space variant. We published a detailed account of our method in the journal Real-Time Imaging.

Deliverables:
The deliverable is a C++ DLL that allows the user to specify a scale for each pixel in the image. The user simply inputs an array of floating point numbers that corresponds to the scale chosen for each pixel. If a circular symmetric linear or sinusoidal scale map is desired, the scale map can be generated by the DLL. The deliverable will be finished in time.

WP4.2.2 – Space-variant mappings

The log-polar mapping has been investigated. In the fovea where samples in the original image are less numerous than in the mapped image, Taylor expansion has been used to extrapolate between sparse image points. The original image was also space-variantly blurred with increasing Gaussian scales towards the periphery in order to avoid aliasing problems.

An even more promising space-variant mapping has been found in the form of a radial affine transform that undoes the perspective effect. This inverse perspective property can be seen in Figure 11.

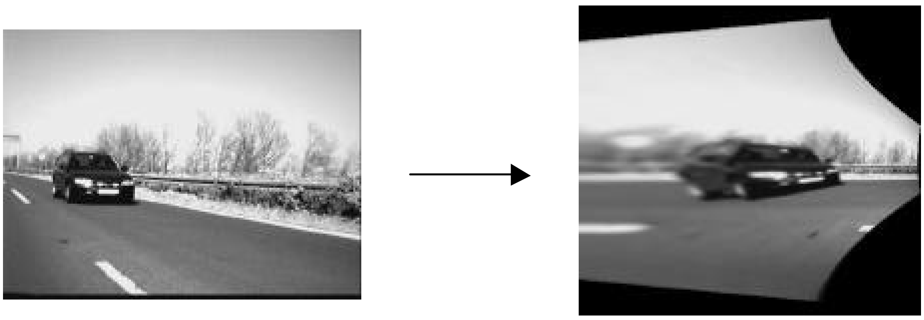


Fig.11 Original image and its inverse perspective transform

The main advantages of this mapping are:

1) The velocity of an object coming towards the camera does no longer appear to increase as it approaches. It is therefore easier to extract the speed of oncoming cars from the optic flow data.
2) The expansion effect of an object travelling towards the camera is significantly attenuated. This means that the translational motion that is the important information in our application is decoupled from the motion due to expansion and is therefore easier to detect.

It should be noted that here again the space-variant algorithm developed in WP4.2.1 is necessary because aliasing can occur in the periphery.

References:
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)
Deriche, R., "Recursively implementing the Gaussian and its derivatives", In: *Proc. Second Int. Conf. On Image Processing*, Singapore, September, pp. 263-267(1992).

WP5

Task 5.1: Psychophysical investigations of driving behavior and center of interest (CoI) in humans

Task 5.1 has been successfully concluded.

The goal of the conducted psychophysical experiments was to study the patterns of eye movements of human observers performing driving-related tasks in pre-recorded and computer-generated driving scenes. We wanted to analyze, on the one hand, which parts of the scenes attract the viewer's interest and, on the other hand, how the eye movement behavior can be characterized. Our findings about the interplay between optic flow and eye-movements are described in detail in deliverable 5.1.
In brief, eye movements were recorded with the EyeLink system while subjects viewed a stimulus that was generated on a Silicon Graphic Indigo2 Extreme computer and back projected onto a transparent screen with a video projector. In different conditions, the subjects were asked for performing three tasks: looking on the stimulus without specific intention, keep gaze directed towards the heading direction while heading changed unpredictably from time to time, and determine whether the current heading is towards any obstacle in the scene.
We found that when subjects viewed the stimulus without any specific task, gaze was clustered near the focus of expansion. When subjects were explicitly instructed to look in to the heading direction gaze also clustered near the focus of expansion. In contrast, when the subjects were required to identify obstacles along the simulated path of self-motion saccades were directed to the obstacles or to the ground in front of the subject and not to the focus of expansion, even though the task clearly required to monitor heading throughout the trial.
In the two cases that involved heading estimation, the 'heading task' and the 'obstacle task', subjects had to react quickly and accurately to changes in heading direction. However, performance differed in the two tasks. Reaction time in the obstacle task was shorter than in the heading task, on average 294ms as compared to 325ms. Surprisingly, the reaction time in the obstacle task did not depend on time-to-contact with the obstacle.

A difference between the two tasks was also apparent in the accuracy of the saccadic reactions to heading change. Subjects need about three saccades to align gaze with the new heading during the heading task and only one saccade to reach obstacles near the new path during the obstacle task. Thus, the accuracy of saccade behavior is much higher in the obstacle than in the heading task.
Despite that the obstacle task involves two sub-tasks, monitoring the direction of heading and finding the obstacles based on the current heading, performance in the obstacle task is faster and more accurate than in the heading task. The results suggest that during driving the center of interest of the observer is mainly directed to obstacles. Saccades towards the focus of expansion are not needed to monitor heading and, more importantly are rather difficult to perform. In contrast, during obstacle tracking heading estimation is easier and more accurate.

Task 5.2: Task optimized representation and the CoI

Optimized representation for robust three-dimensional motion-estimation and obstacle avoidance

As the results from WP5.1 suggest, an important task during driving is to estimate the self-motion during obstacle tracking and avoidance. Visual self-motion estimation relies on the analysis of the global optical flow patterns. Unfortunately, problems arise when optical flow is measured by biological or technical detectors from moving visual stimuli. First, detectors covering small view fields are affected by the aperture problem. The aperture problem occurs for biological detectors like motion selective cells in the striate visual cortex (V1) as well as for optical flow techniques applied to image sequences obtained from pixel-based camera images. Second, biological and technical detectors produce inevitably noise in addition to meaningful flow. Despite these limitations, humans are able to estimate self motion rather accurately. Thus, the brain must have developed methods to remedy the shortcomings of the flow detectors. Processing of self-motion estimation in the brain proceeds from V1 via the middle-temporal area (MT) to the medial superior temporal area (MST). MST is supposed to perform heading detection by analyzing the optic flow provided by area MT. Therefore, area MT is likely the area where noise reduction is performed and a solution to the aperture problem is implemented. In WP 5.2 we have implemented and tested a space variant filtering method of optic flow fields that is derived from properties of area MT and implements an optimized representation for the task of self-motion estimation during obstacle tracking. The method and the tests performed are described in an attached report (*Space variant filtering of optic flow for robust three dimensional motion estimation*). The following section provides a brief overview.

The filtering technique improves the stability of the flow representation by averaging flow vectors within local neighborhoods to stabilize the motion signal. Based on the properties of area MT this filtering method decreases noise by averaging flow vectors over image areas, which increase in size d proportional to the eccentricity ε from the center of the field of view

$$d=0.018+0.61\varepsilon.$$

While averaging over large areas is more favorable for noise reduction and smoothing, averaging over small areas saves information such as local speed and velocity or local motion parallax. The spatial integration over peripherally increasing image areas is a compromise between both goals and well adjusted to the typical structure of the flow field elicited by self-motion. Small areas surrounding the center of the view field contain sets of vectors with large deviations in the local flow direction, whereas the flow field in the periphery is more homogeneous allowing spatial averaging over a large scale without loosing information. The main goal of our analysis was to apply this *MT-like filtering* model to optical flow fields obtained from image sequences with an optical flow algorithm and to test its implication on the quality of heading detection.

We used for our investigations the Brown Range Image Database available on http://www.dam.brown.edu/ptg/brid/range/. It contains three-dimensional data of natural scenes. The knowledge of 3 dimensional data of a given environment allowed us to simulate the view of a moving camera in this scene and calculate both the image on the camera as well as the true motion field. Thus we have ground truth about the simulated camera self-motion. The components of the camera motion, translation and rotation, simulate translation and rotation such that the point in the center of the image is stabilized. This form of motion is similar to the combination of forward motion and eye movement in the obstacle task in WP5.1. The fixation directions that we used in our simulations were obtained from measuring fixations of observers that viewed the same scenes statically on a computer monitor. We obtained optical flow fields from the 3D-generated image sequences by a flow algorithm provided by partner Sco. We used the Heeger-Jepson heading detection algorithm to estimate the influence of the MT-like filtering on the quality of heading detection. Alternatively the IMO detection algorithm from BEL was used (this algorithm is described also under 5.2 below).

Figures 12 and figure 13 present an example of the results. Figure 12 shows for a specific self-motion and environment the correct optic flow, the optic flow as obtained from the image sequence, and the optimized representation of this flow after the MT-like filtering stage. Figure 13 shows the results of the heading estimation applied to the original flow and optimized representation after MT-like filtering. To estimate performance improvements we randomly selected 150 flow vectors for a single run of the heading estimation and computed mean and standard deviation over 30 runs. In Figure 13, the estimated headings in the different runs are given by green dots superimposed on the first image of the sequence, the mean is shown by a blue dot, and the true heading is identified by a red dot. For the original flow (left picture of figure 13) the width of distribution is 25 degree, the deviation of the mean value (blue point) from correct heading (red point) is 12 degree. In the optimized representation (right picture of figure 13) the width of distribution is 6 degree and the deviation of the mean value (blue point) from correct heading (red point) is 4 degree.

Our results thus demonstrate that MT-like filtering is a reasonable strategy to decrease noise in optical flow fields and to improve heading detection. The method works well on optical flow fields based on natural scenes affected by strong noise and the aperture problem. Furthermore, the results show the MT-like filtering method is optimized for flow fields generated by self-motion, which has the same form as the observed motion in the obstacle task in WP5.1. The stability of the heading detection algorithm is increased, the spread of the resulting heading directions is decreased and the mean is near to the correct heading. In the following year we plan to further improve the MT-like optimized representation by including stereo depth information (in conjunction with WP 3.3) and local information about the confidence of a flow vector derived from intrinsic dimensionality (in conjunction with WP. 3.2). We also plan to further investigate the statistical properties of natural optic flow fields in collaboration with partner Sco in order to refine the filter characteristics.

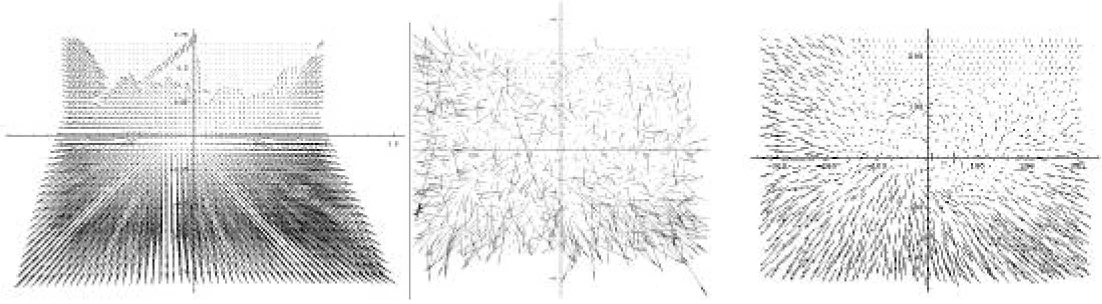


Figure 12: Left: correct optic flow, Middle: Optic flow obtained from the image sequence, Right: MT-like optimized flow representation.

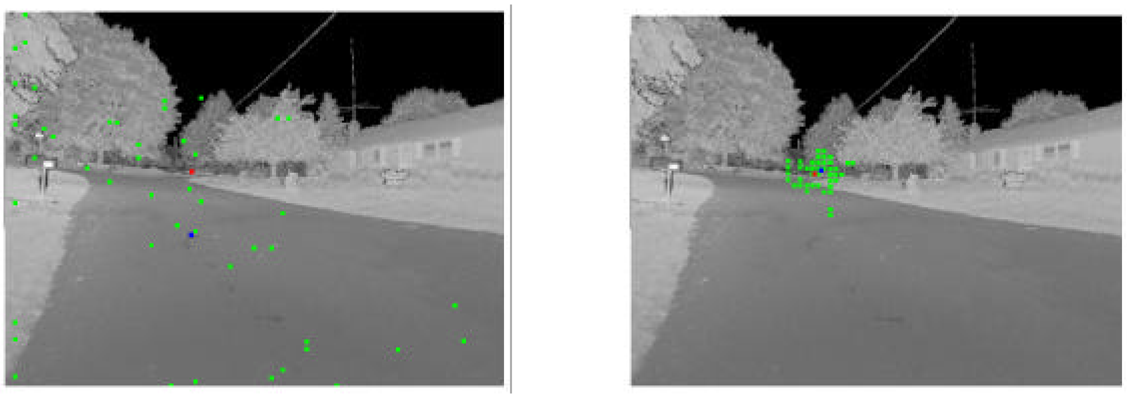


Figure 13: Left: Results of heading detection based on the original flow, Right: Results of heading detection based on MT-like filtered flow representation

Optimized representation for the overtaking task

A further task that was addressed within the consortium is the task of detecting approaching cars from the perspective of the rear view mirror and to determine velocity and impending danger. Partner Eng has implemented an optimized space-variant remapping of the rear mirror image to simplify the task.
The problem that arises in the overtaking task is that we do not a priori know the location of the car or its direction. Foremost, the task is a problem of segmentation. Moreover, the aperture problem causes deviations of the detected flow vectors from the real motion direction. Further complications for the segmentation are that the approaching object expands over time and that the speeds of the flow vectors depend on the distance of the car from the camera. This makes it difficult to measure the real velocity.
Partner Eng has implemented a topologically isomorphic transformation of the two-dimensional image, the inverse perspective transform, which remedies these difficulties. After this transformation, the velocity is no longer depending on the distance to the camera, the expansion of the approaching car is reduced and the translational component of the car is the predominant flow direction. These effects facilitate the car-background segmentation and therefore the localization of the overtaking car. Furthermore the detection of the direction and of the magnitude of the approaching velocity of the overtaking car is improved.

Part of task 5.2 is also concerned with the data-driven segmentation of Independently Moving Objects (IMOs) from optic flow fields. This part is performed by BEL in cooperation with MUN and SCO.

Progress to date

Based on the preliminary background research, performed in the previous reporting period, further work on WP5 has resulted in the development of a complete architecture for the identification of IMOs. First, a novel robust and efficient (real-time) procedure for the extraction of all egomotion parameters from optic flow has been developed. This algorithm is able to deal with the nonlinear interaction between translation and rotation parameters. A second component corrects the intrinsic bias in the translation estimate that originates from the error norm used by nearly all instantaneous-time egomotion algorithms. Finally, IMOs are considered as outliers and techniques from robust statistics applied to make the egomotion estimation insensitive to their presence. Once the observer's motion parameters are extracted, IMOs are identified by evaluating the residual errors of all flow vectors against the egomotion model. Note that this approach can also cope with non-rigid IMOs. Currently this approach is used by MUN to extract heading from MT-filtered flow-fields.

Task 5.3: Implementing early cognitive control in a technical environment

WP5.3 has begun as planned at the end of the current reporting period. Results are not expected before the next period.

WP6

Contextual effects in human and non human primate visual system

The goal of this work package is the development of algorithmic procedures, restricted by neurophysiological constraints that help understanding the contextual effects of attention on early visual processing that have been observed in physiological experiments performed in the laboratory of partner Bel.

The following results have been obtained in this reporting period:
1) building on the results of the previous reporting period, the proposed model mechanism for attentional processing in primary visual cortex, has now been validated in simulation. This concludes WP 6.1.
2) in WP 6.2 this mechanism has been further extended to incorporate regions of the thalamus as well. The complete thalamocortical model, which complies with neurophysiological constraints, constitutes a novel approach and can recreate all experimental results.

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

Building on the preliminary results obtained in the previous reporting period, the novel model architecture, introduced to investigate attentional effects observed in early visual cortex during 2DG experiments, has been completed. The model involves a new center-surround kernel and requires a novel learning rule to be trained. Using a synthetic dataset, which closely resembles the observed physiological data, it has been shown that a specific spatial configuration of long-range inhibitory connections (originating at the location of the representation of the attended stimulus) can generate a ring of suppression surrounding the focus of attention.

Synthetic, one-dimensional data has been created which closely resembles a cross-section of the observed 2DG-data. This dataset is shown in Fig. 14. The stimulus is represented on the left-hand side of the figure. Since, in accordance with the 2DG data, there is less visual stimulation outside the representation, the activation is slightly lower at those locations. In the featural dataset, activation is lowered with respect to the spatial activation in a limited region (ring) outside the stimulus representation.

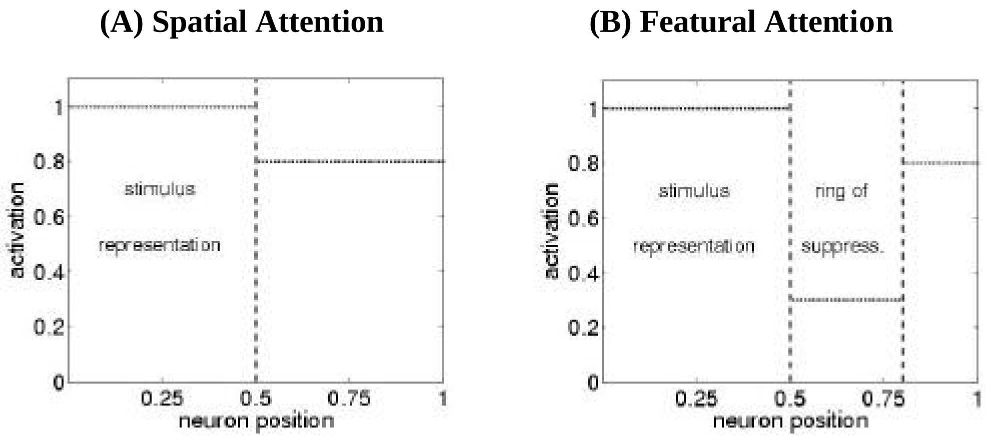


Fig. 14) Synthetic datasets for the spatial (A) and featural (B) attention cases.

Recall that the connectivity pattern for every neuron is parameterised by a splines-based center-surround kernel. The two parameters that need to determined for every neuron are the span of the inhibitory (S_i) and the span of the excitatory (S_e) kernel. Goal of the training is to parameterise the model in such a way that its activation evolves to the featural attention case when initialised with the spatial attention (or stimulus-driven) case. The model configuration, obtained by fitting the model to the synthetic dataset of Fig. 14, is shown in Fig. 15.

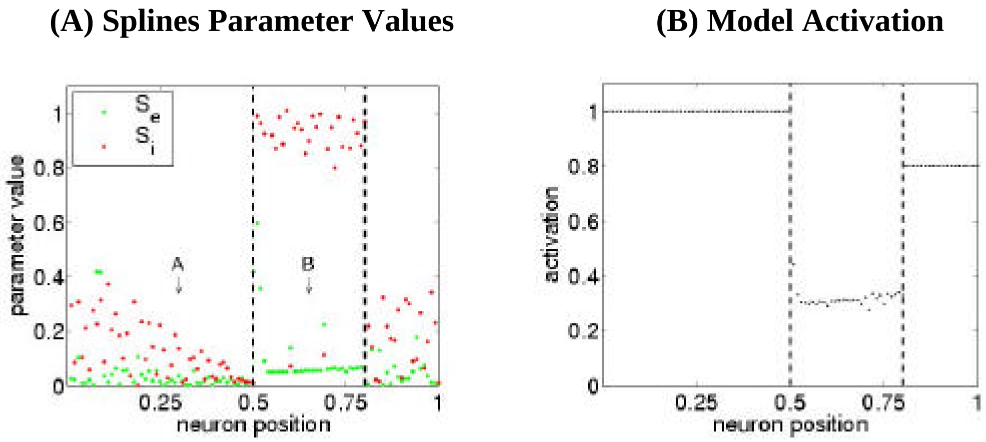


Fig. 15) Featural attention activation (B) obtained by the model with trained excitatory and inhibitory parameter values shown in (A)

Since the activation remains constant in the areas corresponding to the stimulus representation and outside the ring of suppression, different parameter configurations are possible. For this reason, the fit is perfect and the parameter values are always relatively small. No modulatory mechanisms are required here since the neurons are already driven to saturation. The apparent noise, present in the parameter estimates, is due to the discrete nature of the kernel parameterisation and the consequent difficulties in training the model. The higher values on the left-hand side of the figure are caused by border effect. The area outside the stimulus representation is more interesting and clearly shows a consistent parameter configuration. The excitatory center is always relatively small and the inhibitory surround is always large. Fig. 16 shows the connection pattern for the two kernels marked with arrows in Fig. 15.

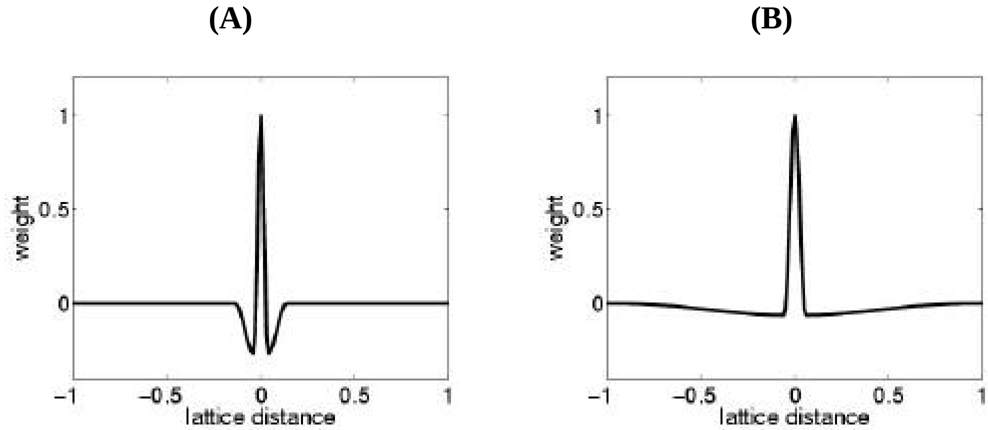


Fig. 16) Representative kernel configurations in the region corresponding to the stimulus representation (A) and the suppressive region surrounding the stimulus representation (B)

It is clear that long-range inhibitory connections allow for the emergence of the desired attentional effects. For further details concerning this model, we refer to **D6.1 "CNS - algorithm for generating task-optimized spatial representations of neural activities"**

WP6.2) Task-optimized representations - general case

The model introduced in WP 6.1 is extended in order to fully explain the experimental results obtained by Vanduffel et. al, 2000. Contrary to the model introduced earlier, this novel model is not restricted to area V1 but includes regions from the thalamus as well. In this way, novel insight is provided with respect to the interaction between contextual effects and attention in the earliest stages of visual processing, in particular the emergence of suppressed activity surrounding the center of interest (CoI).

<u>Experimental Results</u>

By comparing metabolic activity for the two conditions, attentional effects were observed in areas as early as the lateral geniculate nucleus (LGN) and the magnocellular-recipient layers 4Cα and 4B of the striate cortex. In all these areas, attention manifests itself as a retinotopically specific band of suppressed activity, peripheral to the representation of the stimulus.

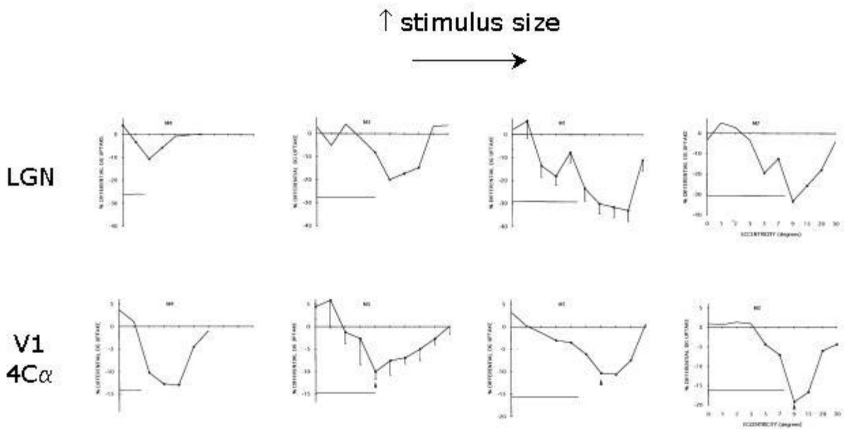


Fig. 17) Normalised differential DG uptake (using the spatial attention task as baseline) in LGN (top row) and V1 4Cα (bottom row) as a function of eccentricity is plotted for different stimulus sizes. Stimulus size increases from left to right. The standard deviations are shown for monkey M1 in LGN and for monkey M3 in V1. The radius of the gratings which were presented in each experiment is indicated by the lines at the bottom of each panel.

The experimental results are summarised in Fig. 17. This figure indicates that the ring of suppression changed in diameter with that of the grating both in LGN and V1, confirming that the suppression surrounded the stimulus representation. The strength of the suppression increased with stimulus diameter in the magnocellular layers of the LGN. This effect was less pronounced in V1 4Cα. In addition to these results, another subcortical change was observed in the visual thalamus: metabolic activity of the reticular thalamic nucleus (RTN) increased in the attention-to-the-grating condition relative to the attention-away condition. This increase strengthened with increasing stimulus size. Due to the absence of precise data on the retinotopy of RTN, the location of this increased activity remains unclear.

Computational Modelling

Before constructing the model, an extensive comparison with existing models has been made. The novel model introduced here can generate the desired effects with a global, uniform attention signal. In this way, the attention signal that enters the model is the simplest one conceivable. Since the experiment is concerned with featural attention, this uniformity assumption is necessary. The main mechanism involves a diffusion of stimulus-driven relay cell activity to RTN (both directly and via V1) and a subsequent injection of inhibition from RTN to LGN relay cells in regions of the LGN surrounding the stimulus representation.

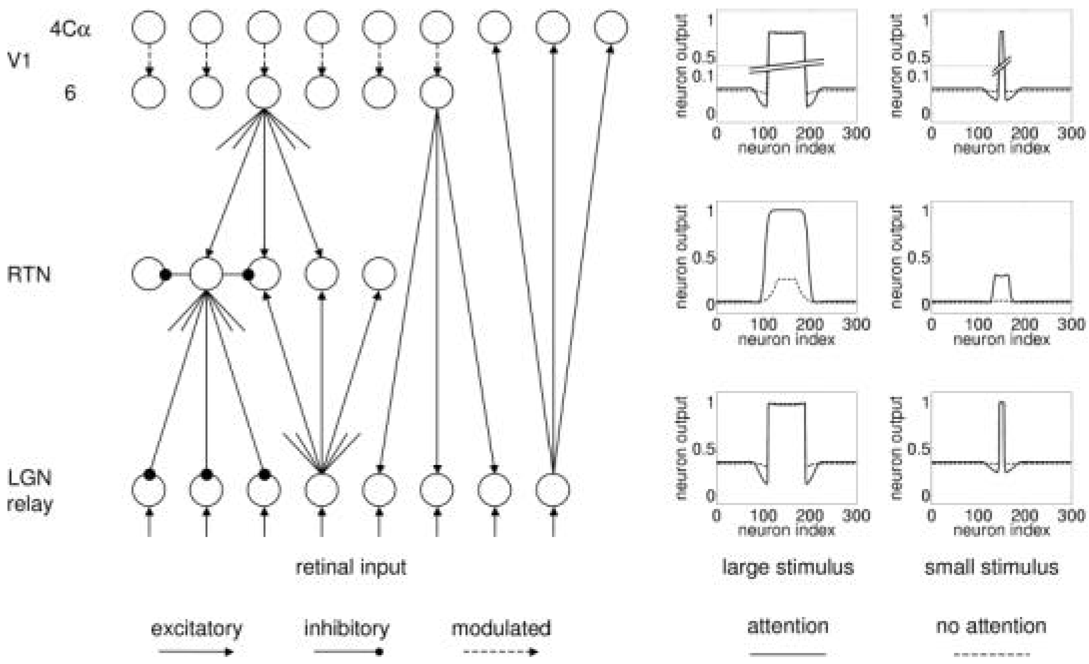


Fig. 18) Model for attentional modulation in the thalamocortical network. Left panel shows synaptic connections between LGN relay neurons, RTN and V1 layers 4Ca and 6 neurons. For each layer, a chain of 300 model neurons with sigmoidal output functions is used. Connectivity to and from RTN is broader (indicated by larger numbers of connections). All connections are excitatory except those coming from RTN; the dashed lines indicate the uniform attentional input in layer 6. Right panel shows the resulting neural output as a function of the neuron's position in the layer in the case of attention (solid lines) and no attention (dashed lines), and for two stimulus sizes (left and right columns). Upper row corresponds to layer 4Ca, middle row to RTN, and bottom row to LGN relay neurons. The thalamocortical model forms a recurrent network with fixed-point iteration dynamics. The plots in the right panel represent the fixed-point states, which are reached after 10 iterations.

As can be seen in Fig. 18, both for large and small stimuli, a ring of suppressed activity surrounds the representation of the stimulus both in LGN and V1. The differential effects in V1 are smaller than in LGN. At the representation and outside the ring, there is no significant change in activity. In correspondence with the experimental results and the inhibitory nature of area RTN, an increased activation can be observed during attention in this layer. When comparing the activity patterns for large and small stimulus size it becomes apparent that the effects scale with stimulus size. The suppression in LGN and V1 is smaller for the small stimulus and so is the increase in activation in RTN. In conclusion, all model results are in correspondence with the experimental results.

Parts of the model have been published in:

Orban, G.A., Pauwels, K., Van Hulle, M.M., & Vanduffel, W. (2003). Attention dependent suppression of metabolic activity at the early stages of the macaque visual system. In *Neurobiology of Attention*, L. Itti, G. Rees, J. Tsotsos (Eds.), Academic Press, Elsevier, in press.

WP7

Has finished in year 1. A set of new camera sequences, however, has been recorded together with Ind to improve the quality.

WP8

Benchmarking and Testing

Only little work has been planned in the WP for year 2. In spite of this some benchmarking has been done in Spa, testing the implemented motion algorithms against the car scenes of HELLA.

WP9

Putting the ECOVISION Systems together

Mainly planned for year 3. Only the first steps have been performed here. These were concerned with assuring that data structures will match and other technical issues.

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 of year 1 and can be downloaded at:
http://www.pspc.dibe.unige.it/ECOVISION/private/index.html (from "Official report")

Task 10.3: Scientific Publications
Have been summarized in D10.3 (which is attached at the end of this report). See also Table 1.

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 booked.
Title of the WS: *Early cognitive vision*
See http://www.cn.stir.ac.uk/ecovision-ws/ for details.

WP11

Management has proceeded without any problems in year 2. There is really nothing much to report here. We are encountering a slight underspending this year and the number of allocated person month has now caught up in year 2 according to plan.

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]	Classifi-cation [2]	Due Date	Issue Date
1	1	**Best Motion Algorithm**	**R+S**	**Rest.**	**31.1.03**	**31.1.03**
3.2	1	**Visual templates with spatial-temporal context - Gestalts in Space-Time**	**R+S**	**Rest.**	**30.6. 03**	**30.6. 03**
4.2	0	**Space Variant Mapping Methods**	**S**	**Rest.**	**31.12.03**	**Planned for 31.12.02**
5.1	1	**Psychophysical investigations of gaze and CoI during simulated driving in humans**	**R**	**Publ.**	**31.12.02**	**31. 01.03**
6.1	1	**CNS algorithm for generating task-optimized spatial representations of neural activities**	**R**	**Publ.**	**30.06.03**	**30.06.03**
10.3	1	**Publication List**	**R**	**Publ.**	**30.6.03**	**30.6.03**

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: D1

Project Number: IST-2001-32114
Project Acronym: ECOVISION
Title: Best Motion Algorithm

Deliverable N°: D1
Due date: January 2003
Delivery Date: January 2003

Short Description:

Low level integer software version of a motion/flow-field to study the bit cutting strategies and their impact on the motion estimation accuracy.

A working integer version of McGM algorithm exists on a Matrox Board (performance examples can be viewed at: http://www.electra.psychol.ucl.ac.uk/jason/). A Technical Report: *The Multi-channel Gradient Model and its Real-Time Implementation*. Describes different topics studied with the implemented version of the McGM.

A software version of the McGM algorithm with embedded bit-cutting restrictions has been implemented to better study the bit-depth in different stages. Some documentation explaining the instructions to use the program and illustrating some results have been added.

Partners owning: Eng and Spa
Partners contributed: Eng and Spa
Made available to: Partners, Comission

DELIVERABLE SUMMARY SHEET D3.2

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

Deliverable N°: D3.2
Due date: 30-06-2003
Delivery Date: 30-06-2003

Short Description: The research activities conducted in WP3.2, and reported in several papers by Sco and Ita have extended and consolidated the approach defined in the first year of the Project. In particular: (1) the rigid body motion (RBM) constraint has been exploited to predict and track visual correspondences of stereo edges, and (2) the temporal prediction properties of a Kalman filter have been exploited to update optic flow and disparity Gestalts across frames, in a temporally continuous way. In this way, we have obtained in the end true motion and stereo Gestalts *both* in space and time

Partners owning: Ita, Mun, Sco

Partners contributed: Ita, Mun, Sco

Made available to: Commission only

DELIVERABLE SUMMARY SHEET D5.1

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: 1.1. 2003

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: Mun
Partners contributed: Mun, Sco, Ind
Made available to: Partners, Comission

DELIVERABLE SUMMARY SHEET D6.1

Project Number: IST-2001-32114
Project Acronym: ECOVISION
Title: Computational Neuroscience Algorithm for generating Task-Optimized Spatial Representations of Neural Activities

Deliverable N°: D6.1
Due date: 30-06-2003
Delivery Date: 30-06-2003

Short Description:

A novel model architecture has been introduced to investigate attentional effects observed in early visual cortex during 2DG experiments. The model involves a new center-surround kernel and requires a novel learning rule to be trained. Using a synthetic dataset which closely resembles the observed physiological data, it has been shown that a specific spatial configuration of long-range inhibitory connections (originating at the location of the representation of the attended stimulus) can generate a ring of suppression surrounding the focus of attention.

Partners owning: Bel.

Partners contributed: Sco., Bel., Ita

Made available to: Pub.

DELIVERABLE SUMMARY SHEET D10.3

Project Number: IST – 2001 - 32114
Project Acronym: ECOVISION
Title: List of Publications

Deliverable N°: 10.3
Due date: July 2003
Delivery Date: July 2003

Short Description: This deliverable lists all publications in journals, at conferences as well as book chapters by the consortium between months 0 and month 18 of the ECOVISION project

Summary:	Reviewed Journal Papers:	31
	Conference Contributions:	41
	Book Chapters	3

Partners owning: Sco

Partners contributed: all

Made available to: Partners, Commission

PROGRESS OVERVIEW SHEETS

Small letters (f) indicate that this Task has finished (either in year one or now).

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**
1	5.0	0	12	0	12	--	5.0 (f)
3.1	10.5	0	15	0	15	2.0	10.5 (f)
3.2	10.5	6	21	6	30	7.0	10.5
3.3	10.0	15	27	15	27	7.0	7.0
3.4	3.0	21	33	21	33	2.0	2.0
4.2	8.0	12	24	12	24	6.0	6.0 (f)
5.3	13.0	24	36	--	--	--	--
7.1	2.0	0	6	0	6	--	2.0 (f)
7.2	1.0	0	12	0	12	--	1.0 (f)
8	5.0	18	36	18	36	3.5	3.5
9	13.0	18	36	18	36	6.0	6.0
10.2	3.0	0	6	0	6	--	3.0 (f)
10.3	1.0	12	36	12	36	1.0	1.0
10.4	3.0	12	36	12	36	1.0	1.0
11	25.0	0	36	0	36	9.0	15.0
Total	113.0					44.5	73.5
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP3	
Task 3.1	Definition, implementation and use of Stereo Gestalts
Task 3.2	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 4	
Task 4.2	Analysis how spoace variant maps can be implemented into the image analysis software of ECOVISION
WP8	Benchmarking of the stereo and motion algorithm using Ind's car movie sequences
WP9	Coordination of the plan to put the different components of ECOVISION together. Planning to implement also stereo on FPGA with Spa.
WP10	
Task 10.3	Compiling a publication list
Task 10.4	Preparing the ECOVISON WS on the Isle of Skye
WP11	Project Management

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
3.2 10.3	Visual templates with Spatial-Temporal Context - Gestalts in Space-Time Publication List	Final, but shall be extended Final

Dissemination actions (articles, workshops, conferences etc.)

Journals

1. Krüger, N., Lappe, M. and Wörgötter, F. (2004).Biologically motivated multi-modal processing of visual primitives. The Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour, 1(5).
2. Krüger, N. and Wörgötter, F. (2003). Symbolic Pointillism: Computer art motivated by human perception. Proceedings of the AISB 2003 Symposium on Biologically inspired Machine Vision, Theory and Application, Wales, 63-69.
3. Krüger, N., Lappe, M. and Wörgötter, F. (2003). Biologically motivated multi-modal processing of visual primitives. AISB Journal 1(4)

Conferences

4. Calow, D., Krüger, N., Wörgötter, F. and Lappe, M. (2004).Space variant filtering of optic flow for robust three dimensional motion estimation. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.
5. Calow, D., Krüger, N., Wörgötter, F. and Lappe, M. (2004).Space variant filtering of optic flow for robust three dimensional motion estimation, submitted to the conference on 'Engineering in Intelligent Systems' 2004 (inpress)
6. Krüger, N. and Felsberg, M. (2003).A continuous formulation of intrinsic dimension. Proceedings of the British Machine Vision Conference.
7. Krüger, N., Lappe, M. and Wörgötter, F. (2003).Biologically motivated multi-modal processing of visual primitives. Proceedings of the AISB 2003 Symposium on Biologically inspired Machine Vision, Theory and Application, Wales, pages 53—59.
8. Krüger, N.,Felsberg, M. and Wörgötter, F.(2004).Processing multi-modal primitives from image sequences. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.
9. Pugeault, N. and Krüger, N. (2003).Multi--modal matching applied to stereo. Proceedings of the BMVC 2003.
10. Pugeault, N.,Wörgötter, F. and Krüger, N. (2004).A non-local stereo similarity based on collinear groups. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.

Book Chapter

11. Krüger, N. and Wörgötter. F.(2004).Statistical and deterministic regularities: Utilisation of motion and grouping in biological and artificial visual systems. Advances in Imaging and Electron Physics, 131.

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 last report WP3.2 and 3.3 blend strongly into each other. Thus, in year three we will continues tol work on both aspects at the same time. **Therefore, nominally we are extending WP3.2 to month 30**

Planned actions for the next period
Merge Stereo (Sco) and Motion (Ita) Gestalts Create Cross-modal interations 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**
5.2	10.0	12	36	12	36	4.0	4.0
6.1	10.0	0	24	0	24	7.0	12.0 (f)
6.2	21.0	12	36	12	36	8.0	8.0
Total	41.0					19.0	24.0
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 5	
Task 5.2	• development of a complete architecture for the identification of IMOs
WP 6	
Task 6.1	• completion of novel model architecture and learning rule for investigation of attentional effects in primary visual cortex • validation of model in simulation
Task 6.2	• incorporation of regions of the thalamus in attention model of Task 6.2

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D6.1	Computational Neuroscience Algorithm for generating Task-Optimized Spatial Representations of Neural Activities	Final

Dissemination actions (articles, workshops, conferences etc.)

Papers published in 1/10/2002 - 31/12/2003 for which Ecovision funding is acknowledged

Journal papers

1.Orban, G.A., Pauwels, K., Van Hulle, M.M., & Vanduffel, W. (2003). Attention dependent suppression of metabolic activity at the early stages of the macaque visual system. In Neurobiology of Attention, L. Itti, G. Rees, J. Tsotsos (Eds.), Academic Press, Elsevier, in press.

2. Mutihac, R. & Van Hulle, M.M. (2003). Biological relevance of features extracted from natural images. Romanian Reports in Physics, 56(3-4), in press.
3. Mutihac, R. & Van Hulle, M.M. (2003). A comparative survey on adaptive neural network algorithms for independent component analysis. Romanian Reports in Physics, 55(1), 43-67.
4. Gautama, T., Mandic, D.P., & Van Hulle, M.M. (2003). On determining the nature of some biomedical signals. IEEE Trans. Biomedical Engineering, in press.
5. Gautama, T., Mandic, D.P., & Van Hulle, M.M. (2003). On the indications of nonlinear structures in brain electrical activity. Physical Review E, 67, 046204.
6. Van Hulle, M.M., & Gautama, T. (2003). 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.
7. Gautama, T., Mandic, D.P., & Van Hulle, M.M. (2003). Signal nonlinearity in fMRI: A

comparison between BOLD and MION. IEEE Transactions on Medical Imaging, 22(5), 636-644.

8. Deleus, F.F., & Van Hulle, M.M. (2003). Monitoring elasticity between science and technology domains and its visualization. Scientometrics, 56(1), 147-160.
9. Van Hulle, M.M. (2002). Joint entropy maximization in kernel-based topographic maps. Neural Computation, 14(8), 1887-1906.
10. 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.
11. 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.
12. Van Hulle, M.M. (2002). Kernel-based topographic map formation by local density modeling. Neural Computation, 14(7), 1561-1573.

Conference papers

1. Mutihac, R., Van Hulle, M.M. (2003). Statistics of feature extraction by topographic independent component analysis from natural images. Proceedings 2nd Int. Conf. on Electronics, Control and Signal Prcessing (World Scientific and Engineering Academy and Society (WSEAS)) (Singapore, 7-12 December, 2002), CD ROM, 451-289.
2. Mutihac, R., Van Hulle, M.M. (2003). Bayesian restoration of medical X-ray digital images. Proceedings 2nd Int. Conf. on Electronics, Control and Signal Prcessing and E-Activities (World Scientific and Engineering Academy and Society (WSEAS)) (Singapore, 7-12 December, 2002), CD ROM, 451-288.
3. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003), A Non-parametric Test for Detecting the Complex-Valued Nature of Time Series, Proceedings of Knowledge-Based Intelligent Information and Engineering Systems (KES'2003) (Oxford (GB), 3-5 September, 2003) Part I, V. Palade, R.J. Howlett and L. Jain, Eds., Germany: Springer-Verlag Berlin Heidelberg, pp. 1364-1371.
4. Gautama, T., & Van Hulle, M.M. (2003). Surrogate-based test for Granger causality. 2003 IEEE Workshop on Neural Networks for Signal Processing (Toulouse, France, September 17-19, 2003), pp. 799-808.
5. Gautama, T., Mandic, D.P., & Van Hulle, M.M. (2003). A Differential Entropy based Method for Determining the Optimal Embedding Parameters of a Signal, Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2003), (Hong Kong, April 6-10, 2003), Vol. VI, pp. 29-32.
6. 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.
7. Deleus, F.F., De Mazière, 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.
8. Mutihac, R., & Van Hulle, M.M. (2002). Neural network implementations of independent component analysis. 2002 IEEE Workshop on Neural Networks for Signal Processing (Martigny, Valais, Switzerland, September 4-6, 2002), pp. 505-514.
9. Pauwels, K., Gautama, T., Van Hulle, M.M., & Mandic, D.P. (2002). Towards model-independent mode detection and characterisation of very long biomedical time series. Proc. 4th International Conference on Recent Advances in Soft Computing (RASC2002) (Nottingham (GB), 12-13 December 2002), pp. 77-78.

Participation to Conferences

1. Brain Connectivity Workshop, 1-3 May 2003, Cambridge, England (F. Deleus)
2. 4th International Conference on Recent Advances in Soft Computing (RASC2002, 12-13 December 2002, Nottingham (GB) (K. Pauwels)
3. Workshop on Neuromorphic Engineering, 28 June - 20 July, Telluride, Colorado, USA (K. Pauwels)
4. 2002 IEEE Workshop on Neural Networks for Signal Processing, 4-6 September 2002,

Martigny, Valais, Switzerland (M. Van Hulle & F. Deleus)

5. 2003 IEEE Workshop on Neural Networks for Signal Processing, 17-19 September 2003, Toulouse, France, (M. Van Hulle)
6. Workshop on Self-Organizing Maps (WSOM'03), 11-14 September 2003, Hibikino, Kitakyushu, Fukuoka, Japan (M. Van Hulle)

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

Planned actions for the next period

Task 5.2

- development of IMO algorithm
- combination of IMO algorithm with task optimised representations for heading detection (WP 5.2)
- employment of context-sensitive filters (spatial or spatiotemporal) from WP 3.1/3.2 to outline the IMO and to determine its (rigid) motion parameters

Task 6.2

- completion of computational neuroscience algorithm for the interaction of attention and context

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**
1	9.0	0	14	0	14	1.0	6.0 (f)
2	4.0	6	18	6	18	2.0	4.0
4.1	21.0	0	18	0	18	8.0	17.0 (f)
4.2	19.0	6	36	6	36	5.5	8.5
8	2.0	18	36	18	36	0.5	0.5
9	5.0	18	36	18	36	1.0	1.0
Total	60.0					18.0	37.0
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 1 WP2	Helping with the implementation of the McGM algorithm
WP 4	
Task 4.1	1. Development and evaluation of 'steering in scale' with taylor series. 2) Development and evaluation of space-variant Gaussian filtering with recursive filtering.
Task 4.2	1) Development of a log-polar mapping using Taylor series and space-variant blurring. 2) Development of an affine inverse perspective mapping using Taylor series and space-variant blurring.
WP 8 WP9	Minor contributions

Deliverables due this period		
Deliverable number	**Title of Deliverable**	**Status (Draft Final, Pending)**
D4.2	Space Variant Mapping Methods	Expected on plan in month 24

Dissemination actions (articles, workshops, conferences etc.)

1) S. Tan, J. L. Dale, A. Johnston, (2003), "Performance of three recursive algorithms for fast space-variant Gaussian filtering", Real-Time Imaging, 9(3), 215-228.
2) Dale, J. and Johnston, A. (2002) A real-time implementation of a neuromorphic optic-flow algorithm. Perception,31, 136.
3) Johnston, A., Durant, S., & Dale, J. L. (2003). A labile representation of spatial information in the visual cortex. *Journal of Vision*, *3*(9), 224a, http://journalofvision.org/3/9/224/, doi:10.1167/3.9.224.
4) Johnston, A., McOwan, P.W. and Benton, C.P. (in press) Biological computation of image motion from flows over boundaries, *J Physiol. Paris*

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

None
Planned actions for the next period
1) Improve the inverse perspective mapping by automating parameter calibration 2) Investigate other space-variant mappings 3) Investigate the benefit of space-variant mappings for velocity information extraction 4) Investigate how space-variant mappings affect self-motion perception (with Ger) 5) Continue work on hardware implementation with Spa

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**
3.2	4.0	12	21	12	30	2.0	2.0
5.1	14.5	0	26	6	24	8.0	9.5 (f)
5.2	21.0	12	36	12	36	9.0	9.0
8	1.0	18	36	18	36	--	--
9	3.0	18	36	18	36	--	--
Total	43.5					19.0	20.5
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	Action
WP 3	
Task 3.2	1) Description and comparison of receptive field structures for flow analysis and heading detection 2) Contributions to a biologically motivated model describing the smallest multi-modal units used by visual systems to perform grouping in collaboration with Sco. 3) Investigations of an optical flow filtering model using weights steered by local intrinsic dimensionality (in collaboration with Sco). First insights in locally statistic properties of optical flow fields justifying the filter model in WP5.2. 4) Implementation of a task optimized optical flow filtering method using static inference stereo information to improve detection of three dimensional self motion (in conjunction with WP5.2).
WP 5	
Task 5.1	Analysis of oculomotor human behaviour during simulated driving with different tasks (Deliv. 5.1). Investigations of the human ability to solve distance estimation tasks from optic flow during simulated self-motion.
Task 5.2	Formulation and testing of a task optimized optic flow filtering method using space-variant mapping to decrease noise and improve heading detection.

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	Final

Dissemination actions (articles, workshops, conferences etc.)
1) J. A. Beintema, A. V. van den Berg, M. Lappe. Circular receptive field structures for flow

analysis and heading detection, in L. Vaina, S. Beardsley, & S. Rushton, Optic flow and beyond. Amsterdam: Kluwer Academic Press. 2003, 2) H. Frenz, F. Bremmer and M.Lappe, Discrimination of travel distances from situated optic flow, Vision Research 43 (2003) 2173-2183 3) N. Krüger, M.Lappe and F. Wörgötter, Biologically motivated multi-modal processing of visual primitives, AISB Journal 1(4) (2003) 4) D. Calow, N. Krüger, F. Wörgoetter and M. Lappe, Space variant filtering of optic flow for robust three dimensional motion estimation, submitted to the conference on 'Engineering in Intelligent Systems' 2004
Deviations from the planned work schedule/reasons/corrective actions/special attention required
None
Planned actions for the next period
We will continue the analysis of optic flow filter models using stereo information, intrinsic dimensionality and implications from statistics in order to provide task optimized representations. A second emphasis during the next period is to continue and to extend the investigations of the first and second order statistic properties of flow fields to get implications for improved filtering methods and to reveal global and local motion gestalt laws.

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**
1	7.0	0	12	0	12	--	7.0 (f)
3.1	9.5	0	15	0	15	--	9.5 (f)
3.2	5.0	6	21	6	30	3.0	5.0
3.3	9.0	15	27	15	27	4.0	4.0
3.4	16.0	21	33	21	33	3.0	3.0
4.1	6.0	0	18	0	18	1.0	5.5 (f)
5.2	10.5	12	36	12	36	7.0	7.0
7.1	1.0	0	6	0	6	--	1.0 (f)
7.2	1.0	0	12	0	12	--	1.0 (f)
8	1.5	18	36	18	36	--	--
9	1.5	18	36	18	36	1.0	1.0
10.1	3.0	0	36	0	36	1.0	2.0
10.3	1.0	12	36	12	36	0.5	0.5
Total	72.0					20.5	46.5
One person month is equal to		126		Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP1	Minor contributions
WP 3	
Task 3.2	Implementation of temporally adaptive Motion Gestalts making use of the temporal prediction properties of a Kalman filter to smooth over time the context-based regularized patch motion segmentation of noisy optic flow fields. Formulation of a Kalman-based approach to regularize rough disparity maps.
Task 3.3	Definition of a possible unifying phase-based framework to disparity, motion and motion-in-depth measurements by a single set of cortical-like filters.
Task 3.4	**Highly confident stereo and motion feature maps to be used by WP4 and WP5.**
WP 4	
Task 4.1	Discussion and implementation help as to the aspect of scalable steerable motion filters.
WP 5	
Task 5.2	Contribution to the definition of a strategy for data driven segmentation of independently moving objects by exploiting the Kalman-based CSF outputs.
WP9	Minor contribution
WP 10	
Task 10.1	Web page maintenance
Task 10.3	Publications and participation to conferences
Deliverables due this period	

Deliverable number	Title of Deliverable	Status (Draft Final, Pending)
D3.2	Visual templates with spatio-temporal context – Gestalts in space-time	Final but shall be extended

Dissemination actions (articles, workshops, conferences etc.)

Ita.1) Paolo Cavalleri, Silvio P. Sabatini, Fabio Solari and Giacomo M. Bisio *Centric-minded Templates for Self-motion Perception* . Vision Research, 43(13): 1473-1493, June 2003.

Ita.2) Silvio P. Sabatini, Fabio Solari, P. Cavalleri and Giacomo M. Bisio *Phase-based Binocular Perception of Motion-in-depth: Cortical-like Operators and aVLSI Architectures* . EURASIP Journal on Applied Signal Processing, 7: 690-702, 2003.

Ita.3) Silvio P. Sabatini, Fabio Solari and Luca Secchi *A Continuum-field Model of Visual Cortex Stimulus-driven Behaviour: Emergent Oscillations and Coherence Fields.* Neurocomputing, *in press 2003*.

Ita.4) S.P. Sabatini and F. Solari *Emergence of Motion-in-depth Selectivity in the Visual Cortex through Linear Combination of Binocular Energy Complex Cells with Different Ocular Dominance.* The annual Computational Neuroscience Meeting, 5-9 July 2003 Alicante, Spain

Ita.5) S.P. Sabatini, F. Solari, and G.M. Bisio *Lattice Models for Context-driven Regularization in Motion Perception.* WIRN03, 5-7 June 2003, Vietri, Italy

Ita.6) S.P. Sabatini and F. Solari *An Early Cognitive Approach to Visual Motion Analysis.* AIIA03, 23-26 September 2003, Pisa, Italy

Ita.7) S.P. Sabatini, F. Solari, and L. Secchi *Emergence of Oscillations and Spatio-temporal Coherence States in a Continuum-model of Excitatory and Inhibitory Neurons* International Workshop Neuronal Coding, 20-25 September 2003, Aulla, Italy

Ita.8) S.P. Sabatini, F. Solari, and G.M. Bisio *Emergence of motion-in-depth selectivity in the visual cortex: An evidence of phase-based second-order motion mechanisms?* European Conference on Visual Perception, 1-5 September 2003, Paris, France

Ita.9) F. Solari, S.P. Sabatini and G.M. Bisio *Context-based structuring action on optic flow fields by generative models of first-order motion primitives: velocity likelihoods and Gestalt detection.* European Conference on Visual Perception, 1-5 September 2003, Paris, France

Ita.10) J. Diaz, E. Ros, S. Mota, G. Rotella, A. Cañas and S.P. Sabatini *Optical flow for cars overtaking monitor: the rear mirror blind spot problem.* Proc. of the 10th Conference on Vision in Vehicles, Granada, September 2003.

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

none

Planned actions for the next period

WP3.3: (1) Analysis of robustness and stability of the local information extracted (e.g., phase and phase derivatives) by the motion-in-depth computational primitives, in order to evidence the best set of local spatio-temporal filters (i.e., with the highest coding capability of joint depth and motion information). This study will extend to spatio-temporal problems the method used in the literature to validate vision algorithms based on phase measurements (Fleet and Jepson 1993, IEEE Trans. on PAMI 15(12):1253-1268).
(2) Characterization of the cortical model of motion-in-depth detectors: plausibility and predictions eventually suggesting new experiments to test the theoretical hypotheses.
(3) Comparison of different architectures based on Gabor-like (cf. energy models) and derivatives of Gaussian kernels (cf. multi-channel gradient models).

WP3.4: Generation of cross-modal confidence maps by using depth (disparity) information to improve the reliability of motion (optic flow) maps.

WP5.2: The problem of segmenting, on the basis of the "common-cause" paradigm, an optic flow field that contains multiple motions is complicated by the presence of ego-motion components. We plan to explore the efficacy of Kalman-based context sensitive filters (CSFs) to obtain a fine segmentation by working on a flow field cleaned from the component due to heading on the basis of the clustering technique developed by Bel.

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
1	19.0+10*	0	14	0	14	2.0 + 5.0*	14.0 + 5.0*
2	38.0+10*	6	25	6	25	16.0 + 5.0*	24.0 + 5.0*
5.3	5.0	24	36	--	--	--	--
8	4.5	18	36	18	36	1.0	1.0
9	4.5	18	36	18	36	1.0	1.0
10.3	1.0	12	36	12	36	--	--
Total	72.0					20.0 + 10.0*	40.0 + 10.0*
One person month is equal to				Person hours			

*Spa has 20 additional MM dedication time (without affecting the budget). These MM have in part been used in years 2 and will continue in year 3 of the project.

Main contribution during this period	
Workpackage/Task	**Action**
WP 1	
Task 1.1	• McGM implemented with embedded bit cutting strategies to evaluate their impact in the final flow-field. • Lucas and Kanade implementability assessment. • Bit cutting evaluation
WP 2	
Task 2.1	Toward a "fully functional HDL simulation" of a flow-field algorithm (Lucas and Kanade approach): • Define technical specifications • Identify the crucial steps and evaluate their relevance in terms of silicon cost and processing speed. • Study different processing schemes and their implications in the final efficiency • Hardware description of the Lucas and Kanade algorithm. • First steps towards the implementation of the McGM algorithm as an extension of the L&K algorithm • First steps towards the implementation of a stereo algorirthm
WP8, WP9	Minor contributions

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

Dissemination actions (articles, workshops, conferences etc.)

Journals:

Spa. 1) J. Díaz, S. Mota, E. Ros and G. Botella: Neural competitive structures for sengmentation based on motion features, *Lecture Notes in Computer Science*, vol. 2686, pp.

710-717 (**ISSN:** 0302-9743), Springer-Verlag, 2003.
Spa. 2) S. Mota, E. Ros, E. M. Ortigosa, and F. J. Pelayo: Bio-inspired motion detection for blind spot overtaking monitor. *International Journal of Robotics and Automation, Special Issue of Neuromorphic Systems* (In Press).

Conferences:
Spa. 3) J. Díaz, E. Ros, S. Mota, G. Botella A. Cañas and S. Sabatini : Optical flow for cars overtaking monitor: the rear mirror blind spot problem, *10th. International Conference on Vision in Vehicles (VIV'2003)*, Granada (Spain), 7-10 September 2003.
Spa. 4) S. Mota, E. Ros, J. Díaz, G. Botella, F. Vargas and A. Prieto: Motion driven segmentation scheme for car overtaking sequences. *10th. International Conference on Vision in Vehicles (VIV'2003),* Granada (Spain), 7-10 September 2003.

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

Going on with the deviation introduced in the last PPR:
Stereo vision algorithms will also be taken under consideration to study their implementability by means of FPGAs.

Planned actions for the next period

- Debugging the "fully functional HDL code", which means finishing D2 due to month 25. And write the specifications. (WP2)
- Evaluation of the migration of the implemented algorithm to a stand-alone platform in order to be used as an embedded system (Beyond WP2)
- Benchmarking and testing (WP8)
- Evaluation of the feasibility of real-time stereo to be embedded on an FPGA platform (according to the deviation introduced in the last PPR)

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**
7.1	5.0	0	6	0	6	1.0	5.0 (f)
7.2	8.0	0	12	0	12	--	7.0 (f)
8	6.0	18	36	18	36	1.0	1.0
9	8.0	18	36	18	36	--	--
Total	27.0					2.0	13.0
One person month is equal to				Person hours			

Main contribution during this period	
Workpackage/Task	**Action**
WP 7	
Task 7.1	• As planed in the last year, a permanently mounted camera rig was mounted in the car. In the end of March Norbert Krüger and Nicolas Pugeault visited Hella to generate more records. A collection of movies and results are published on the EcoVision webpage in the private area.
WP 8	
	• We start with the preparation for a visit from the partner Spa in spring of the next year to benchmark the FPGA-System.

Effort in person months for reporting period 1/1/2003 -31/10/2002
Cumulative (i.e. “total”) since 1.1.2002

	Sco				Bel				Eng				Mun				Ita				Spa				Ind				Total			
	Period		Total		Period		Total		Period		Total		Period		Total		Period		Total		Period		Total		Period		Total		Period		Total	
WP/T	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act**	**Est.**	**Act**	**Est.**	**Act**	**Est.**	**Act**	**Est.**	**Act**	**Est.**	**Act**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**	**Est.**	**Act.**
1	**0.0**	**0.0**	**5.0**	**5.0**					**1.0**	**1.0**	**7.0**	**6.0**					**0.0**	**0.0**	**7.0**	**7.0**	**5.5**	**2.0+5**	**19.5**	**14.0+5**					**6.5**	**3.0+5**	**38.5**	**32.0+5**
2									**2.5**	**2.0**	**5.0**	**4.0**									**19.0**	**16.0+5**	**29.0**	**24.0+5**					**21.5**	**18.0+5**	**34.0**	**28.0+5**
3.1	**1.5**	**2.0**	**10.0**	**10.5**													**0.0**	**0.0**	**9.5**	**9.5**									**1.5**	**2.0**	**19.5**	**20.0**
3.2	**7.0**	**7.0**	**11.0**	**10.5**									**1.5**	**2.0**	**1.5**	**2.0**	**2.0**	**3.0**	**4.0**	**5.0**									**10.5**	**12.0**	**16.5**	**17.5**
3.3	**5.0**	**7.0**	**5.0**	**7.0**													**3.5**	**4.0**	**3.5**	**4.0**									**8.5**	**11.0**	**8.5**	**11.0**
3.4	**2.0**	**2.0**	**2.0**	**2.0**													**3.0**	**3.0**	**3.0**	**3.0**									**5.0**	**5.0**	**5.0**	**5.0**
4.1									**8.0**	**8.0**	**20.5**	**17.0**					**1.5**	**1.0**	**7.0**	**4.5**									**9.5**	**9.0**	**26.5**	**22.5**
4.2	**6.0**	**6.0**	**6.0**	**6.0**					**3.0**	**5.5**	**6.0**	**8.5**																	**9.0**	**11.5**	**12.0**	**14.5**
5.1													**9.0**	**8.0**	**10.5**	**9.5**													**9.0**	**8.0**	**10.5**	**9.5**
5.2					**5.0**	**4.0**	**5.0**	**4.0**					**9.5**	**9.0**	**9.5**	**9.0**	**7.0**	**7.0**	**7.0**	**7.0**									**21.5**	**20.0**	**21.5**	**20.0**
5.3																																
6.1					**4.0**	**7.0**	**9.0**	**12.0**																					**4.0**	**7.0**	**9.0**	**12.0**
6.2					**6.5**	**8.0**	**6.5**	**8.0**																					**6.5**	**8.0**	**6.5**	**8.0**
7.1	**0.0**	**0.0**	**2.0**	**2.0**													**0.0**	**0.0**	**1.0**	**1.0**					**1.0**	**1.0**	**5.0**	**5.0**	**0.0**	**1.0**	**9.0**	**8.0**
7.2	**0.0**	**0.0**	**1.0**	**1.0**													**0.0**	**0.0**	**1.0**	**1.0**					**1.0**	**0.0**	**9.0**	**7.0**	**1.0**	**0.0**	**11.0**	**9.0**
8	**3.0**	**3.5**	**3.0**	**3.5**					**0.0**	**0.5**	**0.0**	**0.5**									**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**5.0**	**6.0**	**5.0**	**6.0**
9	**4.0**	**6.0**	**4.0**	**6.0**					**1.0**	**1.0**	**1.0**	**1.0**					**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**					**7.0**	**9.0**	**7.0**	**9.0**
10.1																	**1.0**	**1.0**	**2.0**	**2.0**									**1.0**	**1.0**	**2.0**	**2.0**
10.2	**0.0**	**0.0**	**3.0**	**3.0**																									**0.0**	**0.0**	**3.0**	**3.0**
10.3	**1.0**	**1.0**	**1.0**	**1.0**													**0.5**	**0.5**	**0.5**	**0.5**									**1.5**	**1.5**	**1.5**	**1.5**
10.4	**1.5**	**1.0**	**1.5**	**1.0**																									**1.5**	**1.0**	**1.5**	**1.0**
11	**8.0**	**9.0**	**14.5**	**15.0**																									**8.0**	**9.0**	**14.5**	**15.0**
Total	**30.0**	**44.5**	**30.0**	**73.5**	**5.0**	**19.0**	**5.0**	**24.0**	**24.0**	**18.0**	**24.0**	**37.0**	**1.5**	**19.0**	**1.5**	**20.5**	**26.0**	**20.5**	**26.0**	**46.5**	**24.0**	**20+10**	**24.0**	**40+10**	**13.0**	**2.0**	**13.0**	**13.0**	**139.0**	**153.0**	**262.5**	**264.5**

Note that Partner Spa has allocated 20MM more of dedication time only (i.e; without affecting the budget of Spa). These MM will appear in years 2 and 3 in WP1 and WP2 and are marked in red. As compared to year 1 where less work had been done due to the contract amendment and the move of M.Lappe to Münster, this year 10% more effort has been spent, which brings the total expected effort now close to the actually used one.

Costs in euro for reporting period 1/1/2003 -31/10/2003

Only the costs for the reporting period are given. Thus, the table deviates for the template in that the previous reporting period has not been added. This is due to the fact in the first PPR costings had not been provided. Thus, only the official cost statement had been issued in January and approved by the Commission. The reason for this was the somewhat confusing situation with the ongoing contract amendments.
The table below shows that we are in general slightly underspending. Deviations from the plan are small, though.

	Sco		**Bel**		**Eng**		**Mun**		**Ita**		**Spa**		**Ind**		**Total**	
	Est	Act	Est	Act	Est	Act	Est	Act	Est	Act	Est	Act	Est	Act	Est	Act
Cost category																
Direct Costs																
1 Personnel	182312	162085	74850	69603	85418	68181	93053	78310	89281	115797	46381	46423	3250	11701	574545	552100
2. Durable Equipment							41278	28471	2083	4407			0	5756	43361	38634
3 Subcontracting									11583	0					11583	0
4 Travel and Subsistence	4166	3188			2500	4331	5000	7327	4750	5320	4167	4719	0	22	20583	24907
5 Consumables	2667	1016	1764	2328	1667	1500	4042	4550	1667	2834	9083	2129	0	100	20890	14457
6 Computing																
7 Protection of Knowledge																
8 Other specific costs					0	1333									0	1333
Subtotal	189147	166289	76614	71931	89585	75345	143373	122004	109364	128358	59631	53271	3250	17579	670964	631431
Indirect costs																
Overheads	37829	31247	15323	14386	17917	2058*	7884	6726	71425	92638	11926	10654	2600	10248	164904	167957
Total	226976	197536	91937	86317	107502	77403	151257	128730	180789	220996	71557	63925	5850	27827	835868	799388
Deviation %		-13%		-6%		-28%		-15%		+22%		-11%		+476%**		-5%

NOTES:
*: The overhead number of Eng is preliminary, because their cost department has still not charged the overheads to the project, We expect that this number will at the end essentially match the estimate
**: There was a budget shift between year1 and 3 into year 2 for Ind. This leads to the apparent high percent-overspending by Ind. Note the total used by Ind in year 2 is still very small, because the role of Ind in the second year was rather limited.

Appendix 6 – Project's Achievements Fiche

Questions about project's outcomes	Number	Comments
1. Scientific and technological achievements of the project (and why are they so ?)		
Question 1.1. Which is the 'Breakthrough' or 'real' innovation achieved in the considered period	N/A	Hardware: Real-time extraction of normal flow. Software 1) Solution for combining motion and stereo information into the joint ECOVISON system. Software 2) Robust estimation of heading in normal flow fields At the current time, these components are already being combined into the ECOVISION system.
2. Impact on Science and Technology: Scientific Publications in scientific magazines		
Question 2.1. Scientific or technical publications on reviewed journals and conferences		**Please see list attached at the end of this document**
Question 2.2. Scientific or technical publications on non-reviewed journals and conferences		**Please see list attached at the end of this document**
Question 2.3. Invited papers published in scientific or technical journal or conference.		**Please see list attached at the end of this document**
3. Impact on Innovation and Micro-economy		
A - Patents		
Question 3.1. Patents filed and pending		1. Device for real-time moving object detection. Authors: S. Mota, J. Díaz, E. Ros. Country: Spain, November, 2003. 2. FPGA device applied to driving assistance during overtaking processes. Authors: S. Mota, J. Díaz, E. Ros. Country: Spain, November, 2003. 3. Device for optical flow estimation through FPGAs. Authors: J. Díaz, E. Ros, S. Mota. Country: Spain, November, 2003. 4. Alert device for lateral collision avoidance in vehicles based on computer vision algorithms implemented in FPGAs. Authors: J. Díaz, E. Ros, S. Mota. Country: Spain, November, 2003.
Question 3.2. Patents awarded		None

Question 3.3. Patents sold		None
Questions about project's outcomes	**Number**	**Comments or suggestions for further investigation**
B - Start-ups		
Question 3.4. Creation of start-up	No	
Question 3.5. Creation of new department of research (ie: organisational change)	No	
C – Technology transfer of project's results		
Question 3.6. Collaboration/ partnership with a company ?		Only as planned with HELLA Hueck KG in the course of the project
4. Other effects		
A - Participation to Conferences/Symposium/Workshops or other dissemination events		
Question 4.1. Active participation[1] to Conferences in EU Member states, Candidate countries / NAS. (specify if one partner or "collaborative" between partners)		**Conferences attended match with the publication list, see there.**
Question 4.2. Active participation to Conferences outside the above countries (specify if one partner or "collaborative" between partners)		**Conferences attended match with the publication list, see there.**

[1] 'Active Participation' in the means of organising a workshop / session / stand / exhibition directly related to the project (apart from events presented in section 2).

B – Training effect		
Question 4.3. Number of PhD students hired for project's completion		Computer Science 5 Psychology 5
Questions about project's outcomes	**Number**	**Comments or suggestions for further investigation**
C - Public Visibility		
Question 4.4. Media appearances and general publications (articles, press releases, etc.)		Media Appearances: Newspapers: Ideal (15 of July 2003). Area of Visibility: Spain Andalucía Investiga (25 of August 2003). Area of Visibility: Andalucía. News in the Web Media, digital Media: Andalucía Investiga (25 of August 2003). Digital Version: http://www.andaluciainvestiga.com/contenidos/espanol/1/rosvidal.asp Consumer: http://www.consumer.es/web/es/noticias/salud_y_seguridad/2003/07/15/63308.php Ideal Digital http://servicios.ideal.es/granada/pg030715/prensa/noticias/Vivir/200307/15/ALM-SOC-046.html TV News: ECOVISION appeared in the TV news on the 16th of July. Channels: Antena 3 (National Channel) and Canal Sur (Autonomic Channel).
Question 4.5. Web-pages created or other web-site links related to the project		http://www.pspc.dibe.unige.it/ecovision/ and the ECOVISION Workshop page:: http://www.cn.stir.ac.uk/ecovision-ws/
Question 4.6. Video produced or other dissemination material		None
Question 4.7. Key pictures of results		We suggest to wait and take some pictures from the presentation at Brussels. At the moment the material is still spread out too much.
D - Spill-over effects		
Question 4.8. Any spill-over to national programs	No	

Question 4.9. Any spill-over to another part of EU IST Programme	Yes	ECOVISION has fundamentally triggered the submission of an IP proposal in the IST call on Cognitive Systems. The next step beyond ECOVISION is to perform higher level (cognitive) vision and action and this is the goal of PACO-proposal (Perception, Action and Cogntion)
Question 4.10. Are other team(s) involved in the same type of research as the one in your project ?	Yes	**MANY !!** **COGVIS** - Cognitive Vision Systems http://cogvis.nada.kth.se **COGVISYS** - Cognitive Vision Systems http://cogvisys.iaks.uni-karlsruhe.de/mainpage.html **ACTIPRET** - Interpreting and Understanding Activities of Expert Operators for Teaching and Education http://actipret.infa.tuwien.ac.at **DETECT** - Real Time Detection of Motion Picture Content in Live Broascasts http://www.detect-tv.com/ **VISATEC** - Vision-based Integrated Systems Adaptive to Task and Environment with Cognitive Abilities http://www.ks.informatik.uni-kiel.de/~visatec/ **LAVA** - Learning for Adaptable Visual Assistants http://www.l-a-v-a.org **CAVIAR** - Context Aware Vision using Image-based Active Recognition http://www.dai.ed.ac.uk/homes/rbf/CAVIAR/ **VAMPIRE** - Visual Active Memory Processes and Interactive REtrieval http://www.TechFak.Uni-Bielefeld.DE/ags/ai/projects/VAMPIRE/

Deliverable D10.3 Publication List months 0-18 of the ECOVISION Project

Listed are:
Reviewed Journal Papers: 1-27
Conference Papers (in part reviewed) 28-66
Book Chapters 67-70

Journals

1. Cavalleri, P., Sabatini, S.P., Solari, F. and Bisio, G.M. (2003). Centric-minded templates for Self-motion Perception. Vision Research, 43(13), 1473-1493.

2. Deleus, F.F. and Van Hulle, M.M. (2003). Monitoring elasticity between science and technology domains and its visualization. Scientometrics, 56 (1), 147-160.

3. Díaz, J., Mota, S., Ros, E. and Botella, G. (2003).Neural competitive structures for segmentation based on motion features, Lecture Notes in Computer Science, vol. 2686, pp. 710-717 (ISSN: 0302-9743), Springer-Verlag.

4. Frenz, H., Bremmer, F. and Lappe, M. (2003). Discrimination of travel distances from situated optic flow. Vision Research, 43, 2173-2183.

5. Gautama, T. and 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.

6. Gautama, T., and 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.

7. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003). On determining the nature of some biomedical signals. IEEE Trans.~Biomedical Engineering, in press.

8. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003). Signal nonlinearity in fMRI: A comparison between BOLD and MION. IEEE Transactions on Medical Imaging, 22(5), 636-644.

9. Gautama, T., Mandic, D.P., and Van Hulle, M.M. (2003). On the indications of nonlinear structures in brain electrical activity. Physical Review E, 67, 046204.

10. Johnston, A., Durant, S. and Dale, J. L. (2003). A labile representation of spatial information in the visual cortex. Journal of Vision, 3(9), 224a

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

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

13. Krüger, N., Lappe, M. and Wörgötter, F. (2004).Biologically motivated multi-modal processing of visual primitives. The Interdisciplinary Journal of Artificial Intelligence and the Simulation of Behaviour, 1(5).

14. Krüger, N. and Felsberg, M. (2002).An explicit and compact coding of geometric and structural information applied to stereo matching, Pattern Recognition Letters, in press.

15. Krüger, N. and Wörgötter, F. (2003). Symbolic Pointillism: Computer art motivated by human perception. Proceedings of the AISB 2003 Symposium on Biologically inspired Machine Vision, Theory and Application, Wales, 63-69.

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

17. Krüger, N., Lappe, M. and Wörgötter, F. (2003). Biologically motivated multi-modal processing of visual primitives. AISB Journal 1(4)

18. Mota, S., Ros, E., Ortigosa, E.M. and Pelayo, F.J. (2004) Bio-inspired motion detection for blind spot overtaking monitor. International Journal of Robotics and Automation, Special Issue of Neuromorphic Systems (In Press).

19. Mutihac, R. and Van Hulle, M.M. (2003). A comparative survey on adaptive neural network algorithms for independent component analysis. Romanian Reports in Physics, 55 (1), 43-67.

20. Mutihac, R. and Van Hulle, M.M. (2003). Biological relevance of features extracted from natural images. Romanian Reports in Physics, 56, 3-4, in press.

21. Sabatini, S.P., Solari, F. and Secchi, L. (2003).A continuum-field Model of visual cortex stimulus-driven behaviour: Emergent oscillations and coherence fields, Neurocomputing (in press).

22. Sabatini, S.P., Solari, F., Cavalleri, P. and Bisio, G.M. (2003). Phase-based binocular perception of Motion-in-depth: Cortical-like operators and aVLSI Architectures. EURASIP Journal on Applied Signal Processing, 7, 690-702.

23. Tan, S., Dale, J.L. and Johnston, A.(2003). Performance of three recursive algorithms for fast space-variant Gaussian filtering. Real-Time Imaging, 9(3), 215-228.

24. Van Hulle, M.M. (2002). Joint entropy maximization in kernel-based topographic maps. Neural Computation, 14(8), 1887-1906.

25. 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.

26. Van Hulle, M.M. (2002). Kernel-based topographic map formation by local density modeling. Neural Computation, 14(7), 1561-1573.

27. Van Hulle, M.M. and Gautama, T. (2003). Optimal smoothing of kernel-based topographic maps with application to density-based clustering of shapes. VLSI Signal Processing Systems for Signal, Image, and Video Technology, in press.

Conference Papers:

28. Beintema, J.A., Berg, A.V.v.d. and Lappe, M.. (2002).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,

29. Calow, D., Krüger, N., Wörgötter, F. and Lappe, M. (2004).Space variant filtering of optic flow for robust three dimensional motion estimation. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.

30. Calow, D., Krüger, N., Wörgötter, F. and Lappe, M. (2004).Space variant filtering of optic flow

for robust three dimensional motion estimation, submitted to the conference on ‘Engineering in Intelligent Systems’ 2004 (inpress)

31. Dale, J. and Johnston, A. (2002).A real-time implementation of a neuromorphic optic-flow algorithm. Perception, 31, 136.

32. Dale. J. L. and Johnston. A., (2002)."A Real-Time Implementation of a Neuromorphic Optic Flow Algorithm", European Conference on Visual Perception, Poster Presentation

33. Deleus, F.F., De Maziere, P. and 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.

34. Diaz, J., Ros, E., Mota, S., Rotella, G., Cañas, A. and Sabatini, S.P. (2003).Optical flow for cars overtaking monitor: the rear mirror blind spot problem, 10th. International Conference on Vision in Vehicles (VIV’2003), Granada (Spain), 7-10 September 2003.

35. Felsberg, M. and Krüger, N. (2003). A probablistic definition of intrinsic dimensionality for images. Pattern Recognition, 24th DAGM Symposium.

36. Gautama, T. and Van Hulle, M.M. (2003). Surrogate-based test for Granger causality. 2003 IEEE Workshop on Neural Networks for Signal Processing (Toulouse, France, September 17-19, 2003), in press.

37. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003). A Differential Entropy Based Method for Determining the Optimal Embedding Parameters of a Signal, Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP 2003), (Hong Kong, April 6-10, 2003), Vol. VI, pp. 29--32.

38. Gautama, T., Mandic, D.P. and Van Hulle, M.M. (2003), A Non-parametric Test for Detecting the Complex-Valued Nature of Time Series, Proceedings of Knowledge-Based Intelligent Information and Engineering Systems (KES'2003).(Oxford (GB), 3-5 September, 2003).Part I, V. Palade, R.J. Howlett and L. Jain, Eds., Germany: Springer-Verlag Berlin Heidelberg, pp. 1364--1371.

39. Krüger, N. and Felsberg, M. (2003).A continuous formulation of intrinsic dimension. Proceedings of the British Machine Vision Conference.

40. Krüger, N., Lappe, M. and Wörgötter, F. (2003).Biologically motivated multi-modal processing of visual primitives. Proceedings of the AISB 2003 Symposium on Biologically inspired Machine Vision, Theory and Application, Wales, pages 53—59.

41. Krüger, N.,Felsberg, M. and Wörgötter, F.(2004).Processing multi-modal primitives from image sequences. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.

42. 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

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

44. Krüger, N. and Wörgötter, F. (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).

45. 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.

46. 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

47. Krüger, N., Jäger, T. and Perwass, C. (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).

48. Morillas, C., Morillas, S., Ros, E., Díaz, A.F., Pino, B. and Pelayo, F.J.(2002).Real-time vision guided movement with reconfigurable Hardware, Workshop Dynamic Perception, Bochum, Germany, November.

49. Mota, S., Ros, E., Pino, B., Botella, G., Pelayo, F.J. and Prieto, A. (2002).Sistema de monitorización de adelantamientos en tiempo real, pp. 93-96, jcra'2002, (ISBN: 84-699-9448-4).

50. Mota, S., Ros, E., Díaz, J., Botella, G.,Vargas, F. and Prieto, A. (2003). Motion driven segmentation scheme for car overtaking sequences. 10th. International Conference on Vision in Vehicles (VIV'2003), Granada (Spain), 7-10 September 2003.

51. Mutihac, R. and Van Hulle, M.M. (2002). Neural network implementations of independent component analysis. 2002 IEEE Workshop on Neural Networks for Signal Processing (Martigny, Valais, Switzerland, September 4-6, 2002), pp. 505-514.

52. Mutihac, R., Van Hulle, M.M. (2003). Bayesian restoration of medical X-ray digital images. Proceedings 2nd Int. Conf. on Electronics, Control and Signal Processing and E-Activities .World Scientific and Engineering Academy and Society (WSEAS).(Singapore, 7-12 December, 2002), CD ROM, 451-288.

53. Mutihac, R., Van Hulle, M.M. (2003). Statistics of feature extraction by topographic independent component analysis from natural images. Proceedings 2nd Int. Conf. on Electronics, Control and Signal Processing ,World Scientific and Engineering Academy and Society (WSEAS)).(Singapore, 7-12 December, 2002), CD ROM, 451-289.

54. Pauwels, K., Gautama, T., Van Hulle, M.M. and Mandic, D.P. (2002). Towards model-independent mode detection and characterisation of very long biomedical time series. Proc. 4th International Conference on Recent Advances in Soft Computing (RASC2002) (Nottingham (GB), 12-13 December 2002), pp. 77-78.

55. Pugeault, N. and Krüger, N. (2003).Multi--modal matching applied to stereo. Proceedings of the BMVC 2003.

56. Pugeault, N.,Wörgötter, F. and Krüger, N. (2004).A non-local stereo similarity based on collinear groups. Fourth International ICSC Symposium on ENGINEERING OF INTELLIGENT SYSTEMS, 2004.

57. Ros, E., Pelayo, F.J., Prieto, A. and del Pino, B.: (2002).Ingeniería Neuromórfica: El papel del hardware reconfigurable, pp. 89-92, jcra'2002, (ISBN: 84-699-9448-4).

58. Sabatini, S.P. and Solari, F. (2003).An Early Cognitive Approach to Visual Motion Analysis. AIIA03, 23-26 September 2003, Pisa, Italy

59. Sabatini, S.P. and Solari, F. (2003).Emergence of Motion-in-depth Selectivity in the Visual Cortex through Linear Combination of Binocular Energy Complex Cells with Different Ocular Dominance. The annual Computational Neuroscience Meeting, 5-9 July 2003 Alicante, Spain

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

61. Sabatini, S.P., Solari, F. and Bisio, G.M. (2003).Emergence of motion-in-depth selectivity in the visual cortex: An evidence of phase-based second-order motion mechanisms? European Conference on Visual Perception, 1-5 September 2003, Paris, France

62. Sabatini, S.P., Solari, F. and Secchi, L. (2003).Emergence of Oscillations and Spatio-temporal Coherence States in a Continuum-model of Excitatory and Inhibitory Neurons International Workshop Neuronal Coding, 20-25 September 2003, Aulla, Italy

63. Sabatini, S.P., Solari, F., Andreani, G., Bartolozzi, C. and Bisio, G.M. (2002).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.

64. Sabatini, S.P., Solari, F. and Bisio, G.M. (2003).Lattice Models for Context-driven Regularization in Motion Perception. WIRN03, 5-7 June 2003, Vietri, Italy

65. Solari, F., Sabatini, S.P. and Bisio, G.M. (2003).Context-based structuring action on optic flow fields by generative models of first-order motion primitives: velocity likelihoods and Gestalt detection. European Conference on Visual Perception, 1-5 September 2003, Paris, France

66. 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.

Book Chapters

67. Beintema, J.A.,van den Berg, A.V. and Lappe, M. (2003).Circular receptive field structures for flow analysis and heading detection, in L. Vaina, S. Beardsley, and S. Rushton, Optic flow and beyond. Amsterdam: Kluwer Academic Press.

68. Lappe, M. (2003).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, (in press)

69. Krüger, N. and Wörgötter. F.(2004).Statistical and deterministic regularities: Utilisation of motion and grouping in biological and artificial visual systems. Advances in Imaging and Electron Physics, 131.

70. Orban, G.A., Pauwels, K., Van Hulle, M.M. and Vanduffel, W. (2003). Attention dependent suppression of metabolic activity at the early stages of the macaque visual system. In Neurobiology of Attention, L. Itti, G. Rees, J. Tsotsos (Eds.), Academic Press, Elsevier, in press.