Esprit LTR 21914 CUMULI
Month 12 Periodic Progress Report

Introduction

The LTR project CUMULI aims to develop automated methods for accurate 3D industrial measurement from multiple images, with particular emphasis on the geometric aspects of the problem. The consortium is inter-disciplinary, with a mixture of computer vision and geometric reasoning researchers (Grenoble, Sophia, Lund) and applications-oriented partners in close range industrial photogrammetry (IMETRIC), motion measurement from high-speed image sequences (I.V. Image Systems) and 3D modelling for augmented and virtual reality applications (Fraunhofer IGD). It tries to capitalize on the significant recent advances in multi-image geometry in the computer vision community, and apply them in mature, real world visual measurement problems. We also hope that this will lead to some fruitful cross-fertilization between the vision, photogrammetry, reasoning, and modelling communities.

A brief overview of the project is as follows (see the Project Programme for more details). There are three main work packages, each aimed at a different aspect of 3D measurement from multiple images. WP 1 and 2 concentrate on core topics in geometric vision and photogrammetry: projective and uncalibrated structure; constraints, calibration and Euclidean structure; and measurement of non-point-like geometric features (lines, conics, curves, surfaces…). WP 1 focuses on discrete images and the algebraic aspects of these problems. It is directly linked to the IMETRIC demonstrator whose main goal is more flexible methods of recovering multi-camera orientation. WP 2 concentrates on continuous image sequences, tracking, and motion estimation. It is linked to the I.V. Image Systems crash-testing demonstrator, in which the non-rigid 3D structure and motion of a car crashing into a wall will be recovered from high speed image sequences.

WP 3 is more speculative. Geometric models are more than just coordinates, they are complex networks of incidence relations, constraints, *etc*. To measure or build such models from uncertain data, we need methods of recognizing and mobilizing these constraints to ensure consistency, reduce uncertainty, improve registration, generate new matching hypotheses, …. Hence, geometric reasoning under uncertainty will be a key topic in future systems. WP 3 is an initial attempt to adapt existing geometric theorem proving technology to this sort of application. It is linked to the Fraunhofer IGD demonstrator, which focuses on model-image calibration, registration and occlusion problems for augmented reality applications.

For each topic and site, we give man-months (mm) used to date.

WP 1: Multi-Camera Geometry, Discrete Images

Work package 1 concerns the theoretical and practical aspects of deriving very accurate estimates of 3D geometrical structure from a discrete set of images of a static scene. The practical work focuses on the needs of IMETRIC, who also supplied the test data. The theoretical and algorithmic results are used not only in this work package, but also form a basis for work package 2 and 3. Hence, I.V. Image Systems have also followed the work in this package (0.5mm) and have absorbed some of the results, which they expect to use in the near future.

Task 1.1: Projective Multi-camera Geometry

Lund (2mm) have continued their theoretical work on the understanding of multi-camera geometry. One of the key results is an analysis of the differences in using only the bilinear constraints or the trilinear constraints [14]. The connections between algebraic, subspace and shape methods are better understood [12]. This has also resulted in efficient and robust structure and motion algorithms [26, 11]. These shape based factorization methods give estimates of structure and motion that are independent of coordinate systems, ordering of points, and images.

Grenoble (2mm) have developed a new method for multi-image projective reconstruction, which extracts the projection matrices from a set of matching tensors estimated from point or line correspondences [37, 39]. This provides a modest but useful improvement in accuracy over methods that use only two of the images.

They have also developed a general, numerically efficient method for nonlinear, statistically optimal estimation under constraints. This can be used for curve, surface and 3D model fitting, but in this work task it has been applied to the optimal estimation of the fundamental and essential matrices and the trifocal tensor, taking into account the algebraic consistency relations on these objects [36].

Sophia-Antipolis (2mm) have developed a formalism which allows dynamic visual perception of geometric and kinematic 3D information, for a monocular visual system without calibration. To efficiently achieve this, special configurations of points or lines are first detected by considering either projective, affine or Euclidean geometry, with a simple uniform representation of points and lines. Planar and rigid structures are also detected with the same formalism.

Task 1.2: Euclidean and Affine Multi-camera Geometry

Lund (2mm) have continued their work on autocalibration, *i.e.* the recovery of internal and external calibration and Euclidean structure from the knowledge that some of the internal parameters of a moving camera are constant or known. The case of all internal parameters being constant has been solved in a new formulation using the projection of the absolute quadric [13].

They have also developed a theory that enables autocalibration under very weak assumptions on camera calibration, for example, only that the skew is known to be zero [15, 16].

The connection between Euclidean structure and motion can be regarded as a special case of projective structure and motion with one extra quadric (the absolute quadric). Thus it has been shown that projective structure and motion from 5 points and one conic in two images has 10 solutions [17], similar to the case of 5 point in two calibrated images [19].

Grenoble (2mm) have also continued their work on autocalibration. They have reformulated the problem in terms of a new object — the 'absolute quadric' — and developed an efficient and stable numerical method based on the rectification of a prior projective reconstruction [38]. Initial tests show that this method is much stabler than previous ones, having comparable accuracy to conventional calibration against a precision 3D calibration grid.

They have also developed a complete catalogue of the degeneracies intrinsic to *any* autocalibration method [34, 33], which should help in the development of future algorithms. Finally, they have generalized the autocalibration problem to a zoom-lens camera, under the restriction that a prior model of the variation of principal point with zoom is available [32, 35].

Sophia-Antipolis (1mm) have analyzed the use of specific displacements (fixed axis rotation, pure translations, ...) or specific sets of camera parameters for self-calibration from un-calibrated image sequences. This allows: *(i)* a minimal parameterization of the camera and motion parameters for improved estimation; *(ii)* the use of affine or metric constraints when available; *(iii)* the automatic choice of appropriate kinds of displacements for the visual system. Since these correspond to "industrial displacements", they are of practical use for robotic or vehicle visual systems.

Task 1.3: Geometric Features and Uncertainty

Lund (2mm) have continued their theoretical work on the recovery camera motion using only the deformation of apparent contours in images [5, 27, 18]. They have implemented algorithms which calculate both projective and Euclidean motion from multiple images containing silhouettes of unknown general surfaces [18].

They have also extended their shape based factorization methods for points to the case of curves. This theory enables efficient structure and motion estimates using only the images of general curves [2]. As in the point case, the shape based algorithms are independent of coordinate systems, the ordering of the images and of the initial parametrization of the image curves. In fact, the algorithm also solves the point correspondence problem for curves.

The shape based factorization methods for points can be refined using nonlinear 'bundle adjustment' techniques [25]. Lund have generalized these from points to curves. Thus it is possible to obtain a statistically optimal estimates of structure and motion for curves if the characteristics of the noise are known [3].

This requires the development of techniques for calculating and handling noise characteristics for extracted points and edge curves. This has been possible due to new developments on stochastic models of image acquisition and low level image processing [28, 30, 3].

Grenoble (1mm) have developed a linear algorithm to recover 3D affine structure and motion of 3D lines from three uncalibrated affine views [22, 21, 24]. The key idea is the introduction of a one-dimensional projective camera, which converts the affine 3D line reconstruction problem into a projective 2D point one. The solution is unique up to a re-ordering of the views.

Task 1.4: Automatic Estimation of Camera Pose

Lund (2mm) have put considerable effort into the implementation and transfer of existing techniques as well as the more recent results from the group. An initial MATLAB-toolbox for structure and motion, resection and intersection of points, lines, conics and curves has been developed [31].

They have continued the research and implementation of routines for Euclidean structure and motion in minimal or close to minimal situations and under degenerate configurations. In particular the methods of Philip (for 6 or more non-coplanar points) and of Wunderlich (for 4 or more coplanar points) have been implemented, tested and transferred to the industrial partner.

The MATLAB toolbox they have developed [31] has been tested on data provided by IMETRIC, to solve the relative orientation problem. This is expected to have a substantial impact on the industrial application.

Grenoble (1mm) have developed a family of stable quasi-linear algorithms to estimate the 3D position and orientation of a calibrated camera from the images of 4 or more known 3D points [23]. The methods guarantee a unique solution (even from coplanar points) provided the points do not lie in a critical configuration, and allow additional points to be incorporated to further enhance stability.

Grenoble have also developed a new method for relative camera orientation from images of four or more unknown but coplanar points. This is similar to Wunderlich's 1982 method, but uses SVD rather than eigen-decomposition to make the results more stable for distant planes. This has been combined with a method for non-coplanar points in an attempt to recover relative orientation from 6 or more points whatever their degree of coplanarity. The result works well in the coplanar and non-coplanar limits, but more work is required to provide a reliable method for intermediate cases.

Imetric (3mm) has delivered two test data sets to the partners. The first consists of 6 Kodak DCS460 images (3072×2048 pixels) of a network of nearly coplanar points. Image point coordinates, exterior orientation parameters, camera calibration parameters and 3D object point coordinates were also provided. This data is being used by the academic partners to develop algorithms for relative orientation using minimal point configurations and for obtaining 3D coordinates for the points.

The second data set consists of 6 Kodak DCS460 images of 8 cylindrical objects placed at various attitudes inside a room. Exterior orientation parameters, camera calibration parameters and three dimensional data describing the cylinders were also provided. This data set is being used by the partners to develop algorithms for automatic orientation of images using 3D features and for the determination of 3D data about the features.

IMETRIC is also negotiating with one of its clients to obtain an industrial data set from a nuclear power plant. At the moment security issues are delaying the release of the data. IMETRIC are still working with the customer to see whether these problems can be overcome, or whether another test site will have to be selected.

IMETRIC has been evaluating the industrial relevance of the techniques under development in CUMULI. Recently it has appeared that it may be possible to eliminate the "orientation cross" from the current measurement process. This is potentially of great value for the industry, as in automated production environments the placement of an orientation cross gives rise to security constraints which adversely affect the performance of the overall system. The university partners have provided information on algorithms which may allow the cross to be eliminated, and IMETRIC will strive to implement these and perform initial tests in Q1 of 1998.

WP 2: Image Streams and 3D Motion

This work package closely parallels WP 1, but emphasizes image streams and motion rather than discrete images of static scenes. Virtually all of the scientific work listed under WP 1 is also applicable to WP 2, but to avoid duplication we have listed it *only* under WP 1. Please keep this in mind and refer to the WP 1 section as necessary. Here we describe only work aimed specifically at image streams and motion estimation. I.V. Image Systems have followed the scientific work in this package (0.5mm) and expect to use some of the results in the future.

Task 2.1: Incremental Projective Camera Geometry

Grenoble (2mm) have begun a study of numerical methods and organizations for large scale bundle adjustment (nonlinear statistical optimization of structure and motion estimates) and its recursive and constrained cousins. Statistical near-optimality is a prerequisite for accuracy, so bundle methods will always be central to measurement applications. However the volume of data available from image streams makes efficient numerical methods essential, and the need to handle complex problems (independent, sometimes non-rigid motions, complex prior information, ...) and to improve automation and reliability (automated model/method choice, blunder detection, ...) all make the optimization problem highly nontrivial. This applies equally to projective reconstruction (task 2.1) and to constrained or calibrated problems (task 2.2).

The new projective reconstruction method [37, 39] detailed under task 1.1 is also directly applicable to this work task.

Lund (2mm) have developed a theory that generalizes the multilinear constraints to the case of a continuous stream of images. In this framework, the epipolar constraint, as well as the trilinear, quadrilinear and all higher order ones have been generalized to the continuous case. A key result is that the third order continuous constraints are needed to reconstruct the scene and the camera motion, *i.e.* optical flow is not sufficient [29].

They have also developed an iterative and recursive algorithm that solves for projective structure and motion from closely sampled image sequences. This algorithm calculates the camera matrices and reconstructs the scene incrementally [1].

Sophia-Antipolis (2mm) have developed first order motion models for structure and motion from monocular image sequences. The camera may have unknown and varying intrinsic parameters, and both the camera and objects in the scene may be moving. To achieve this, they have revisited the problem of retinal motion parameterization in uncalibrated image sequences, including a mechanism for image stabilization in which the rotational disparity is iteratively canceled. This continuous model has been compared experimentally to a discrete approach on various sets of real images.

Task 2.2: Continuous Constraints and Euclidean Structure

Grenoble (1mm): The new autocalibration method [38] detailed under task 1.2 is well adapted to continuous image problems, as it is linear in the number of images and can work from any initial projective reconstruction. The work on methods for bundle adjustment mentioned under task 2.1 is also very directly applicable to task 2.2.

Lund (1mm) have developed the theory of continuous constraints for image streams, mentioned in Task 2.1, also in the case of Euclidean structure from calibrated cameras [29].

The work on Euclidean structure from weak assumptions on the camera geometry, mentioned in Task 1.2, is also applicable in this task [13, 15, 16].

Sophia-Antipolis (2mm) have developed a model of dynamic foveal 3D sensing for analyzing the Euclidean structure of a dynamic scene from uncalibrated monocular image sequences, using an active steerable camera. They have shown that the camera motion can always be chosen to produce a particularly simple form for the retinal displacement, in which the equations in the uncalibrated case become almost identical to the calibrated ones. Very interesting properties occur, for instance, fusion along the image sequence and reconstruction of the observed scene are easy, *etc*. If the motion constraints are only approximately satisfied, the model is still valid for points close to the retina. A small experimental implementation takes an image sequence as input, estimates the retinal motion fields, and calculates the reconstruction up to a particular affine transformation of the scene.

Task 2.3: On-line Calibration and 3D Motion from Image Streams

This task is devoted to work directly associated with the I.V. Image Systems demonstrator: rigid and non-rigid motion estimation for points on a car crashing into a wall during a safety test.

A significant effort was required during the first 6 months to define the demonstrator scenario (see month 6 report). 3D information is currently little used in the crash-testing industry, so there was no established industry practice to base the demonstrator on. The markings on the cars are currently large and sparse — designed for manual rather than automated tracking. We have suggested that smaller, denser markings would provide denser, more accurate 3D information and make automated tracking easier. To prove this concept, we have had to make special arrangements for an appropriately marked car to be crash tested. This has delayed the delivery of the formal test-data for WP 2.3 (although interim data of a crash test dummy's head was delivered at month 6 as promised), however we now expect the final data to be available by the end of the year. This should not delay the final demonstrator, although much of the work on the tracking library has been postponed until we have the images to test it on.

I.V. Image Systems (3mm) were actively involved in the discussions to define the WP 2 demonstrator. They provided the first set of test data (month 6) and made preparations for obtaining further test data from a full-scale crash test to take place in autumn 1997. They have also developed a new tracking algorithm for use in the demonstrator and commercial applications. This is currently undergoing testing and evaluation.

Grenoble (1mm) played an active part in refining the demonstrator scenario, and made a start on the tracking library for the demonstrator. More effort will be concentrated on this task in the next few months, when the test images become available.

Lund (1mm) have, in a master's thesis, developed algorithms for feature extraction, tracking and incremental reconstruction from closely sampled image sequences [9].

WP 3: Algebraic Symbolic Reasoning

Task 3.1

IGD (5mm): IGD's work in this work package focused on the elaboration of application scenarios, data acquisition and solution specification in discussion with the other partners. The exploitation results of REALISE showed that visually reconstructed models of 3D environments are not usually competitive with hand-built ones for most Virtual Reality applications [20]. So we decided to shift the focus from building static Virtual Reality models to on-line 3D reconstruction for Augmented Reality applications, in which the camera-scene calibration is an intrinsic part of the task.

The Millennium Bridge in London has been selected as an application scenario for this work package. We acquired video sequences and static images of the site from several different points of view, focus and illumination. These have been delivered to all WP 3 partners. We also acquired and prepared a detailed 3D CAD model of the designed bridge. In this application, the architects are likely to provide us with both new image data and new CAD models rather frequently, which we will have to augment in short turn around time.

This scene has proven to be very difficult to autocalibrate with the tools of REALISE (TotalCalib from INRIA Sophia). Some configurations seem to be unstable and no results can be generated. There are still problems concerning the inaccuracy of the generated model: errors in planarity of surfaces, orthogonality and alignment. Yet, the use of a manual tool that IGD developed in the project CCIC has shown that with some extra knowledge (scene maps, position of the photographer, orthogonalities and coplanarities in the scene), a 3D model can be constructed. For this reason, the London scene is a very good test-case for work package 3.

INRIA Sophia (9mm) has studied various different methods (Chou-Gao-Zhang, Richter-Gebert Crapo, Wu) for geometric theorem proving and identified some of their strengths and drawbacks. These are powerful methods that can handle many kinds of geometric hypotheses in 2D (projective, affine or Euclidean) and generalize the ad-hoc methods that have been designed for some particular types of properties (alignments, coplanarities, ...). The methods of Richter-Gebert and Crapo deal with binomials of determinants and directly manipulate (projective) invariants, but seem difficult or costly to generalize to the 3D case. Those based on characteristic sets (Wu) add constraints on how to choose the coordinate system and the parameters to use. The final approach by Chou-Gao-Zhang also deals with invariants, but depends heavily on the order in which hypotheses are entered into the system. Combining elements borrowed from all these methods, we are currently working on a new method that attempts to correct their shortcomings. Integration of uncertainty information seems to be non-trivial.

Simultaneously, we have tried to use the geometric proof methods in a "generative way": the problem here is to find all the conclusions that can be drawn from some given geometric configuration. We obtained very encouraging results for plane geometry. The method is time consuming, but it is highly parallel and could be used off-line to find all properties of a given template situation. The 3D case is under investigation.

Finally, work was done in conjunction with IGD and Lund to analyze the problems posed by the London sequence. We plan to test these symbolic techniques for the analysis and reconstruction from an image and a map.

Grenoble (1.5mm) have developed a general approach based on Clifford algebra and term-rewriting for geometric construction, calculus and reasoning in 2D and 3D [41]. It has been applied to several example problems from computer vision. The feasibility and efficiency of the approach has also been demonstrated by our preliminary experiments on automated theorem proving in plane Euclidean geometry.

Lund (2mm): in this task Lund are concentrating on numerical approaches to geometric reasoning under uncertainty. Their work relies heavily on basic tools from WP 1 and 2, adapted to the particular demands on model building raised by VR and AR. Examples of such requirements are accurate depth recovery and occlusion handling, often with complications caused by the fact that the images are taken under far from ideal circumstances, *e.g.* with scenes that are approximately flat, and with all viewpoints from the same vicinity.

Starting from the affine shape formalism for the recovery of projective structure and motion [26, 11, 12], their work so far has been directed on enforcing affine and Euclidean constraints on the reconstruction (*e.g.* that certain points are coplanar, that the line connecting two points is parallel to *e.g.* the ground plane, that certain lines are orthogonal). The methods allow simultaneous handling of any number of points in any number of images, where a new issue to cope with occlusions is that all points need not be detected in all images. Future developments will include the use of more complex information, like maps in models for city planning, and tools for automatic hypothesis generation (*e.g.* that certain points are coplanar) and verification.

Another line of work, involving geometric reasoning, aims at the development of tools for image correction, to make them consistent with constraints in the scene. This work emanates from previous work in the ESPRIT-project VIVA, where analogous problems for single-view images of polyhedral scenes were dealt with [10].

Changes in work plan for WP 3

Formal geometric proving techniques offer a powerful, general framework in which to handle arbitrary geometric constraints. However, given the difficulties we have identified with the current generation provers, it seems unrealistic to hope to achieve all of the goals of task 3.2 using *only* these tools. In particular, the London sequence presents a real challenge and there is a lot that can be done in this direction in tasks 3.1 and 3.2. In particular, much needs to be done to integrate prior geometric knowledge of the scene with the current tools. We thus propose the following changes to the work plan of package 3. These proposals are detailed in a joint document that will constitute a 12 month deliverable for work package 3 (included as an appendix of this report).

Task 3.1:
• Continue to study and improve geometric provers within task 3.1 (Grenoble,Sophia) with the goals of using them for maintaining the coherence of a geometric data base and generating off-line the properties of common template situations.
• Develop more specialized methods for identifying and handling certain sub-classes of geometric constraints (partial 3D knowledge from maps, coplanarities, . . .).

Task 3.2:
The focus of task 3.2 will change from virtual reality to augmented reality applications, and will use combinations of the new tools of task 3.1 (including geometric proving if available) to achieve the specified goals.

WP 4: Management

The management of the project is going smoothly and there have been no major upsets. During the first six months, the key industrial problems around which technology transfer in CUMULI will focus were defined and clarified. The WP1 and 2 demonstrator scenarios have been refined to reflect this, but the changes are small: WP1 transfer will focus on orientation (camera pose) from geometric primitives, and WP2 transfer on 3D motion estimation, with calibration from known 3D points in the images. WP 3 has required more significant a re-alignment to ensure that it focuses on practically important problems, however we believe that it is now on track and that no further reorganization should be necessary.

Deliverables

To minimize the amount of material the reviewers have to read, we have not included copies of the CUMULI month 12 deliverable papers and code with this report (which is CUMULI deliverable 4.1/12). However we would be very happy to supply copies of any of the below material, should the reviewers or the commission wish to inspect it.

European Community support for CUMULI has aided and been acknowledged in (at least) the following papers: [1, 2, 3, 4, 6, 7, 8, 14, 15, 16, 17, 18, 20, 21, 22, 23, 31, 36, 38, 40]. Of these, [26, 11, 39] have been selected as deliverable 1.1/12*(i)* ‘a short article on algebraic multi-image projective geometry’ and [1, 29] as deliverable 2.1/12*(i)* ‘a short article on incremental multi-image projective reconstruction.

The associated code deliverables 1.1/12*(ii)* ‘initial implementation of algorithms for discrete multi-image projective reconstruction’ and 2.1/12*(ii)* ‘. . . for incremental multi-image projective reconstruction’ are contained in MATLAB libraries supplied by Lund and Grenoble. Each site has written routines for each problem. Routines for relative orientation of calibrated cameras from unknown 3D points have also been supplied for 1.1/12*(ii)*, as IMETRIC feel that this problem has more immediate industrial applications.

Both the IMETRIC and I.V. Image Systems industrial data sets (1.4/12 and the second half of 2.3/6*(ii)*) have been delayed by the reluctance of their clients to release data which might possibly be commercially sensitive. However both I.V. Image Systems and IMETRIC are actively pursuing these negotiations and expect to be able to provide suitable test data within the next few months. We do not expect this to cause delays to the overall project workplan.

As WP 3 has been significantly reoriented to improve its industrial relevance, a short unscheduled report discussing the changes has been included as an appendix of this document.

Summary

At present CUMULI is on target and running to budget. The man-months consumed in the first 12 months are listed in table 1, by work task and by site. Overall, there has been a substantial transfer of technology in which all partners have played an active part, both through technical discussion at CUMULI meetings and through experiments on the test data provided by the industrial partners.

WORK TASK	MONTH	Gren	Soph	Lund	I.V.	Imetric	Fraun	TOT.
1 Discrete Images		6/17	4/12	8/17	0.4/2	4/17		22.4/65
1.1 Multi-camera Geometry	0–24	2/4	2/5	2/4	0.2/1	0.4/1		6.6/15
1.2 Euclidean Structure	0–36	2/4	1/5	2/4		0.3/1		5.3/14
1.3 Features & Uncertainty	0–36	1/4		2/4	0.2/1	0.3/1		3.5/10
1.4 Transfer Imetric	0–36	1/5	1/2	2/5		3/14		7/26
2 Image Streams		4/15	4/11	4/14	2.6/20			14.6/60
2.1 Incremental Geometry	0–24	2/5	2/5	2/5	0.2/1			6.2/16
2.2 Continuous Constraints	6–30	1/5	2/5	1/5	0.1/1			4.1/16
2.3 Transfer I.V.	0–36	1/5	–/1	1/4	2.3/18			4.3/28
3 Reasoning		1.5/3	9/24	2/7			5/22	17.5/56
3.1 Geometric Reasoner	0–18	1.5/3	9/12	2/4			5/11	17.5/29
3.2 Visual Applications	12–36		–/12	–/3			–/11	–/27
TOTAL RESEARCH		11.5/35	17/47	14/38	3/22	4/17	5/22	54.5/181
4 Management	0–36	4/13	1/3	1/3	1/3	0.3/1	0.5/2	7.8/25
TOTAL PROJECT		15.5/48	18/50	15/41	4/25	4.3/18	5.5/24	62.3/206

Table 1: Man-months consumed in year 1 *vs.* total budgeted man-months for entire task, per work task and per site.

Appendix: WP 3 Reorganization (unscheduled deliverable)

Introduction

The goal of work package WP3 of CUMULI is to investigate the use of automatic geometric reasoning to aid the reconstruction and management of 3D model data bases.

Formal geometric proving techniques offer a powerful, general framework in which to handle arbitrary geometric constraints. However, given the difficulties we have identified with the current generation provers, it seems unrealistic to hope to achieve all of the goals of work package 3 using *only* these tools. This is mainly because of the unstructured nature of the hypotheses we have to deal with: whereas conventional mathematical theorems are described in a very structured way (only the relevant hypotheses and objects given, in a specified “constructive order”), objects and hypotheses generated from images may or may not be relevant and are given in an arbitrary order. This data structuring problem makes it very difficult to apply existing 3D reasoning methods.

On the other hand, we have recognized a need to integrate prior geometric scene knowledge into existing vision tools, to allow them to deal with difficult situations. One example is IGD's "London" scene, which appears to be very difficult to autocalibrate with the tools of REALISE.

This document details the changes in the workplan of the package WP3 that were summarized in the first year CUMULI report.

Task 3.1

The original task 3.1 was exclusively devoted to a study of available geometric provers, and to the development of a system that would overcome the limitations of existing ones. Due to the considerations explained in the introduction, the main emphasis of the task toward the development of an automatic 3D theorem prover has been slightly shifted and the following directions have been added:

- The use of maps: in many situations, maps of the scene are available. These are commonly used (*e.g.* by IGD) when building models by hand. Moreover, they potentially (at least in their numerical form) constitute a very rich source of constraints that can be used to recover the geometry of the scene more accurately. Thus, it seems interesting to develop automated tools able to deal with this source of information, and to integrate them into a vision based system. For example, maps might be used to stabilize the computation of the epipolar geometry, constrain and check the calibration, *etc*.
- Develop tools for the automatic generation of some specific types of hypotheses. There are various geometrical constraints that can be derived directly from the images (for example, incidences).
- Integrate geometrical knowledge into computer vision methods. This knowledge can be used in various ways: finding the model most appropriate for a given situation, selecting configurations of geometric features that lead to better stability of the algorithms, informing the user of degenerate situations…

This does not mean that we propose to do more in the task with the same effort, but rather that we will spend less of our efforts on the 3D prover and more on the above three directions, while keeping the overall number of man months constant.

As a consequence we propose to keep deliverable 3.1/18 as a demonstration of software, but with the focus shifted from the general 3D prover to the results of investigations on the above three topics.

Task 3.2

Originally, task 3.2 was devoted to the integration and testing of the new tools within the vision based modelling system built in REALISE. The four main topics that need to be considered are:

- The coherence of the geometric data base.
- The prediction of properties of the environment.
- To keep the coherence between the geometric data base and uncertainties.
- Better tools to help the correspondence process.

The goals of this task will remain the same under the change of workplan, but the means will include all the new tools developed in task 3.1 in addition to those that were originally planned.

Conclusion

The main thrust of the task has been shifted from a highly theoretical one to a more applied one with a stronger integration of the needs of the end-user (IGD).

References

[1] R. Berthilsson, A. Heyden, and G. Sparr. Recursive structure and motion from image sequences using shape and depth spaces. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 444–449. IEEE Computer Society Press, 1997.

[2] R. Berthilsson and K Åström. Reconstruction of 3d-curves from 2d-images using affine shape methods for curves. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 476–481. IEEE Computer Society Press, 1997.

[3] R. Berthilsson, K. Åström, and A. Heyden. Projective reconstruction of 3d-curves from its 2d-images using error models and bundle adjustments. In *Proc. 1st Scandinavian Conf. on Image Analysis, Lappeenranta, Finland*, pages 581–588, 1997.

[4] Sylvain Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, 1998. to appear.

[5] R. Cipolla, K. Åström, and P. J. Giblin. Motion from the frontier of curved surfaces. In *Proc. 5th Int. Conf. on Computer Vision, MIT, Boston, MA*, pages 269–275. IEEE Computer Society Press, 1995.

[6] Olivier Faugeras and Théodore Papadopoulo. Grassmann-Cayley algebra for modeling systems of cameras and the algebraic equations of the manifold of trifocal tensors. RR 3225, INRIA, July 1997. To appear in the Transaction of the Royal society A.

[7] Olivier Faugeras and Théodore Papadopoulo. A nonlinear method for estimating the projective geometry of three views. RR 3221, INRIA, July 1997.

[8] Olivier Faugeras and Théodore Papadopoulo. A nonlinear method for estimating the projective geometry of three views. In *6^{th} Int. Conf. Computer Vision*, Bombay, India, January 1998.

[9] P Hallberg and A. Vegh. Feature detection, tracking and iterative reconstruction from image sequences. Master's thesis, Dept. of Mathematics, Lund University, okt. 1997. To appear.

[10] A. Heyden. On the consistency of line-drawings, obtained by projections of piecewise planar objects. *Journal of Mathematical Imaging and Vision*, 6(4):393–412, 1996.

[11] A. Heyden. Projective structure and motion from image sequences using subspace methods. In *Proc. 1st Scandinavian Conf. on Image Analysis, Lappeenranta, Finland*, pages 963–968, 1997.

[12] A. Heyden, G. Sparr, and K. Åström. Perception and action using multilinear forms. In *International Workshop on Algebraic Frames for the Perception-Action Cycle, Keil, Germany*, 1997.

[13] A. Heyden and K. Åström. Euclidean reconstruction from constant intrinsic parameters. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, volume 1, pages 339–343. IEEE Computer Society Press, 1996.

[14] A. Heyden and K. Åström. Algebraic properties of multilinear constraints. *Mathematical Methods in the Applied Sciences*, 20:1135–1162, 1997.

[15] A. Heyden and K Åström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 438–443. IEEE Computer Society Press, 1997.

[16] A. Heyden and K Åström. Minimal conditions on intrinsic parameters for euclidean reconstruction. In *Proc. 2nd Asian Conf. on Computer Vision, Hong Kong, China*, 1998. accepted.

[17] F. Kahl and A. Heyden. Using conic correspondences to estimate the epipolar geometry. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, 1998. To appear.

[18] F. Kahl and K. Åström. Motion estimation in image sequences using the deformation of apparent contours. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, 1998. To appear.

[19] S. Maybank. *Theory of Reconstruction from Image Motion*. Springer Verlag, Berlin, Heidelberg, New York, 1993.

[20] S. Mueller. Computer vision based modelling and augmented reality. Presentation of results of REALISE and CUMULI at industrial workshop *From images to knowledge*, SIRA Technology Centre, London, 1–2 October 1997.

[21] L. Quan. Uncalibrated 1d projective camera and 3d affine reconstruction of lines. In *Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

[22] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1997. to appear.

[23] L. Quan and Z. D. Lan. Linear $\geq N$ point pose determination. In *Int. Conf. Computer Vision*, 1998. To appear.

[24] L. Quan and R. Mohr. Uniqueness of 3d affine reconstruction of lines with affine cameras. In *Int. Conf. on Computer Analysis of Images and Patterns*, pages 231–238, Kiel, Germany, 1997. Springer Verlag.

[25] C. C. Slama. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Church, VA, 1980.

[26] G. Sparr. Simultaneous reconstruction of scene structure and camera locations from uncalibrated image sequences. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, volume 1, pages 328–333. IEEE Computer Society Press, 1996.

[27] K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. In R.C̃ipolla B.B̃uxton, editor, *Proc. 4th European Conf. on Computer Vision, Cambridge, UK*, volume 1065 of *Lecture notes in Computer Science*, pages 97–108. Springer-Verlag, 1996.

[28] K. Åström and A. Heyden. Stochastic modelling and analysis of image acquisition and subpixel edge detection. In *Proc. International Conference on Pattern Recognition, Vienna, Austria*, volume 2, pages 86–91. IEEE Computer Society Press, 1996.

[29] K. Åström and A. Heyden. Continuous time matching constraints for image streams. *Int. Journal of Computer Vision*, 1997. accepted for publication.

[30] K. Åström and A. Heyden. Stochastic analysis of image acquisition and scale-space smoothing. In J. Sporring, M. Nielsen, L. Florack, and P. Johansen, editors, *Gaussian Scale-Space Theory*. Kluwer Academic Publishers, 1997.

[31] K. Åström, A. Heyden, F. Kahl, R. Berthilsson, and G. Sparr. A computer vision toolbox. In *Proc. Nordic Matlab Conference*, 1997. To appear.

[32] P. Sturm. Self-calibration of a moving camera by pre-calibration. In *Proc. British Machine Vision Conference*, pages 675–84, Edinburgh, September 1996. British Machine Vision Association.

[33] P. Sturm. Critical motion sequences and conjugacy of ambiguous euclidean reconstructions. In *Scandinavian Conf. on Image Analysis*, volume I, pages 439–46, June 1997.

[34] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. In *Proc. Conf. Computer Vision and Pattern Recognition*, Puerto Rico, 1997.

[35] P. Sturm. Self-calibration of a moving zoom-lens camera by pre-calibration. *Image and Vision Computing*, 15(8):583–90, August 1997.

[36] B. Triggs. A new approach to geometric fitting.

[37] B. Triggs. Linear projective reconstruction from matching tensors. In *Proc. British Machine Vision Conference*, pages 665–74, Edinburgh, September 1996.

[38] B. Triggs. Autocalibration and the absolute quadric. In *Proc. Conf. Computer Vision and Pattern Recognition*, Puerto Rico, 1997.

[39] B. Triggs. Linear projective reconstruction from matching tensors. *Image and Vision Computing*, 15(8):617–26, August 1997.

[40] T. Viéville and D. Lingrand. Using specific displacements to analyze motion without calibration. *Int. Journal of Computer Vision*, 1998. Accepted.

[41] D. Wang. Clifford algebraic calculus for geometric reasoning with application to computer vision. In R. Caferra, L. Fari nas del Cerro, H. Shi, and D. Wang, editors, *Automated Deduction in Geometry*, Berlin Heidelberg, 1997. Springer-Verlag. To appear.