Esprit LTR 21914 CUMULI

Month 24 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). CUMULI 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 the side of a stationary car 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 this year (months 13–24 of CUMULI). A table at the end shows total man months used to date (months 0–24).

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 (0.6mm) have absorbed some of the results.

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 [27]. Another key result is [21], in which all algebraic relations between different tensor components are revealed in a common framework. The connections between algebraic, subspace and shape methods are better understood, and the different approaches have been unified in [25, 7]. This has also resulted in efficient and robust structure and motion algorithms [42, 20, 43]. These shape based factorization methods give estimates of structure and motion that are independent of coordinate systems, ordering of points, and images. In the joint paper [14] (deliverable 1.1/24), different tools for multi-view geometry, like affine shape, multilinear constraints and P-matrices have been brought together.

Grenoble (2mm) have put a substantial effort into developing a robust, efficient, modular library for many types of matching tensor estimation. The library is not yet finished, but preliminary results are reported in [51]. Matching tensors are an implicit, close-to-the-image representation of 2–4 camera geometry, extremely useful for feature correspondence and the intermediate stages of 3D reconstruction. However estimating them accurately is quite complicated as there are complex non-linear constraints, and many possible constraint and feature types, error models and parametrizations, special or degenerate cases, and numerical initialization and optimization methods. Our main idea is to decouple the problem, so that different *(i)* feature types and error models; *(ii)* tensor types and parametrizations; *(iii)* initialization and optimization methods, can be compared quantitatively, and combined to handle new problems.

Grenoble have also refined their factorization based projective reconstruction library, and contributed a discussion of factorization methods to the joint overview paper [14], deliverable 1.1/24.

Sophia-Antipolis (3mm) have completed an analysis of a formalism and equations to allow the detection of special feature configurations under either projective, affine or Euclidean geometry, with a simple uniform representation of points and lines. This front-end processing insures that a robust set of data is input to subsequent motion algorithms, and hence improves their performance.

Sophia has also studied the constraints that tie the coefficients of the so-called trifocal tensor that describes projectively the geometry of three cameras. This has led to a set of algebraic equations that have been used to give a new one-to-one minimal parameterization of this tensor. That parameterization has been used in a new estimation method. A geometric description of these constraints has also been given in [17, 16, 37].

Task 1.2: Euclidean and Affine Multi-camera Geometry

Lund (1mm) 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 were all of the internal parameters are constant was considered in a formulation using the projection of the absolute quadric [26]. Since then, they have developed a theory that enables autocalibration under very weak assumptions on camera calibration, for example, only that the skew is known to be zero [28, 22]. The critical motions for autocalibration under these weak assumptions have also been investigated: an exhaustive description will be given in [29].

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 [34], similar to the case of 5 points in two calibrated images [36].

For the affine camera approximation Lund has developed a factorization method for structure and motion from point, line and conic features [32, 31].

Grenoble (1mm) have reformulated their 'absolute quadric' based autocalibration method in terms of 'direction frames' — triplets of 3D projective points representing a basis for the 3D Euclidean structure — and used this to develop an autocalibration algorithm for a moving projective camera viewing a *planar* scene [50]. Planes are common primitives in built environments, for which feature extraction and matching are relatively easy. But they are singular cases for projective reconstruction, so *(i)* there is more need for calibration, and *(ii)* conventional autocalibration can not be initialized as the projective 3D structure in not available. The new method makes a nonlinear least squares search over all unknown calibration parameters, and also over the unknown Euclidean structure of the 3D plane parametrized by its two circular points or direction frame vectors. Several initialization techniques have been studied. The main disadvantage is that relatively many images are needed (at least 5 if none of the five camera intrinsic parameters are known).

Sophia-Antipolis and Grenoble (1mm) have jointly studied the self-calibration of a 1D projective camera in a 2D world, and described a method for uniquely determining its two internal parameters using the trifocal tensor of three 1D images [15]. The 1D trifocal tensor can be estimated linearly from point correspondences with no approximations, unlike the usual 2D one. Given the tensor, calibration reduces to finding the roots of a cubic equation in one variable. The main interest is that certain common configurations of 2D cameras can be reduced to the 1D model. In particular, we deduce a new method for self-calibrating an ordinary 2D camera undergoing planar motion.

Sophia-Antipolis (3mm) is currently testing methods for self-calibration from uncalibrated image sequences, using minimal parameterizations of the camera and motion parameters for specific types of displacement (fixed axis rotation, pure translation, . . .) or specific sets of camera parameters. The problem of choosing a given displacement of a robot mounted camera in order to achieve a given perpetual task will be addressed next.

Task 1.3: Geometric Features and Uncertainty

Lund (1mm) have continued their theoretical work on the recovery camera motion using only the deformation of apparent contours in images [13, 45, 30]. They have implemented algorithms which calculate both projective and Euclidean motion from multiple images containing silhouettes of unknown general surfaces [30].

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 [8]. As in the point case, the shape based algorithms are independent of coordinate systems and the ordering of the images. In fact, the algorithm also solves the point correspondence problem for the curves.

The shape based factorization methods for points can be refined using nonlinear 'bundle adjustment' techniques [41]. 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 [9]. Bundle adjustment has also been generalized to combinations of points, lines, conics, curves and patches [1, 3]. This requires the development of techniques for calculating and handling noise characteristics for extracted points, edge curves, correlation patches and other features. This has been possible due to new developments on stochastic models of image acquisition and low level image processing [46, 48, 9] and the work on finding affine correlations [4].

Lund has also continued the work on understanding the statistics of shape, which might potentially be used for tracking and recognition, [6, 5].

Task 1.4: Automatic Estimation of Camera Pose

Lund (1mm) have put much 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 [49].

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 [38] (for 6 or more non-coplanar points) and of Wunderlich [53] (for 4 or more coplanar points) have been implemented, tested and transferred to the industrial partner.

The MATLAB toolbox they have developed [49] 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 (2mm) have extended their MATLAB pose estimation library with several new routines for relative camera orientation, mainly focusing on specialized but industrially common cases involving lines and circular cylinders. The routine for relative orientation from coplanar points has been extended to allow arbitrary combinations of coplanar points and lines, and new routines have been written for orientation from: *(i)* 1 circle and 1 3D point; *(ii)* the sides of two or more 3D circular cylinders (not all aligned). Other circle and cylinder based methods are easily synthesized from the base routines provided.

Grenoble have also supplied a MAPLE implementation of the methods described in [39], for camera pose from ≥ 4 and ≥ 5 3D points and their images.

Imetric (2mm) has evaluated direct solutions for estimating the camera pose and their impact on work procedures in both standard metrology systems and automated systems for machine control and in-process inspection. The direct pose solution is in practical use and has proven to be very reliable. IMETRIC has found that the typical performance is approximately 99.5%. In the remaining cases the solution fails due to erroneous identifications of targets by the IMETRIC software, which are not recovered in later steps. IMETRIC will improve the algorithms in order to make these solutions even more robust. IMETRIC has evaluated the impact of direct solutions for relative orientation on work procedures for the different systems it manufactures. Although IMETRIC has procedures which circumvent the need for relative orientation for a number of automated systems, some applications show major benefits, in particular manual measurement of objects and machine control. In both cases there is no repeat situation, and the placement of an orientation cross is an additional effort in the first case and a major security risk in the second. Due to delays in the development of the new "IMETRIC Metrology Camera System", IMETRIC has not yet had time to implement these algorithms. It now plans to implement them towards the end of '98 or the beginning of '99. The automated systems have been well received in industry as evidenced by customer feedback.

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 near future.

Task 2.1: Incremental Projective Camera Geometry

Grenoble (2mm): the matching tensor library discussed under task 1.1 and the self-calibrating bundle methods discussed below under task 2.2 can also be used for incremental projective structure.

Lund (3mm) 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 third order continuous constraints are needed to reconstruct the scene and the camera motion, *i.e.* optical flow is not sufficient [47].

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 [7, 23]. Recently, they have also begun an investigation of the stability and convergence of factorization methods for recovering projective structure.

Sophia-Antipolis (3mm) have developed a front-end module for systems performing motion analysis of long video sequences (MPEG animation, surveillance recording television programs, etc...), to ease their interactive analysis. First, the image sequence is stabilized in the sense that non-relevant disparity between two frames has been canceled. Then regions of non-negligible disparity (probably corresponding to nearby obstacles, or objects in motion) are segmented, and a projective indicator of the relative location/size of these objects is computed.

Task 2.2: Continuous Constraints and Euclidean Structure

Grenoble (3mm) have continued their work on self-calibrating bundle adjustment for reconstruction of non-static scenes from unsynchronized images. Most current bundle algorithms are designed for static scenes. Motion introduces two problems: *(i)* Self-calibration relies on (projective or Euclidean) *rigidity* in the scene. For arbitrarily moving rigid objects, there are self-calibration constraints from each object but they tend to be weak if the objects are small and have unknown 3D structure. *(ii)* The number of model parameters increases enormously as there is one set for each time instant, rather than a single global set. *E.g.* for the I.V. test sequence around $6 \cdot 2 \cdot 150 \cdot 5 = 9000$ pose parameters are needed to specify the motion of the $3 + 3 = 6$ d.o.f. coordinate frames of two cars in unsynchronized 150 image sequences from 5 cameras (*i.e.* $150 \cdot 5$ distinct times), not counting the parameters for camera calibration and 3D structure within each moving frame. The motion parameters are strongly coupled by a stochastic motion model, so sparse matrix (or equivalently filtering based) techniques can in principle be applied. But since the motion parameters also couple to the 3D structure and camera calibration ones, the sparseness is much harder to manage than in traditional static scene or Kalman filtering type approaches. We are studying iterative and sparse matrix schemes for resolving this large, complex system, and also robust techniques for deciding when the motion of a target becomes non-rigid with respect to its frame. We stress that this complexity is unavoidable if truly optimal structure/motion/calibration estimates are required — many simpler schemes exist in the literature, but give non-optimal solutions.

Lund (2mm) 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 [47]. The work on Euclidean structure from weak assumptions on the camera geometry, mentioned in Task 1.2, is also applicable in this task [26, 28, 22].

They have also developed a method to obtain Euclidean reconstruction from an approximate affine reconstruction. In this method different quasi-tensors are estimated and the affine camera matrices are calculated using the so called closure constraints. This procedure makes it easy to cope with missing data and puts the same emphasis on each image and feature. Finally, the affine reconstruction is upgraded to a Euclidean reconstruction assuming zero skew and unit aspect ratio [24, 31].

Sophia-Antipolis (3mm) have finalized the development of their simplified parameterization of the motion analysis problem. This applies to cases in which a robotic system controls the direction of camera translation and stabilizes for rotations, so as to produce an optimal model for the retinal displacement. It has been established that if only the "foveal" part of the visual field is considered, this model is well-defined even when the displacement constraints are only partially verified.

Furthermore, this study has recently been complemented by an analysis of the camera self-calibration problem in this case, demonstrating that it is very easy to recover all of the camera intrinsic parameters required for 3D-reconstruction, while the remaining, unmeasurable calibration parameters are not in fact needed in this case.

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 two crashing cars during a safety test.

I.V. Image Systems (2mm) IV Image Systems have produced test data for WP2. The data consists of film and video images of a side collision between two cars, in total about 900 images from five non-synchronized cameras at a (non-constant) frame rate of about 500 Hz. The cars were marked with both the quad targets commonly used in car crash test today, as well as white disks on a black background. Figure 1 shows several images from the middle of the sequence.

We have tracked a subset of the points through the sequences, both to provide data for the academic partners and to study the problems of tracking in this type of data (many, densely spaced, identical targets). Our current tracking methods are not ideal for this task, requiring a relatively large amount of manual supervision and intervention. In particular, the process of identifying targets between different camera views is entirely manual. We are currently improving our tracking algorithms, as well as the interaction between the tracking software and the operator, to facilitate the tracking of this kind of image sequences. We feel that the techniques being developed in CUMULI for tracking, segmentation and feature identification will be very useful for our application.

We have also been working on integrating CUMULI results on resection and pose estimation in our systems. Compared to standard camera calibration techniques, the CUMULI work on auto-calibration will improve the usability of our systems and minimize the amount of work needed for our customers to perform 3D measurements. This will be true even in the cases where the measurement geometry means that complete auto-calibration is not possible.

Grenoble (4mm) have invested a considerable amount of effort into developing a tracker for the circular targets used in the WP 2 test data. However, the results are not yet acceptable and more work is required. The current implementation does not detect the initial targets as reliably as we would like, and also tends to lose track of them after about 15–30 images. Both of these problems are caused by

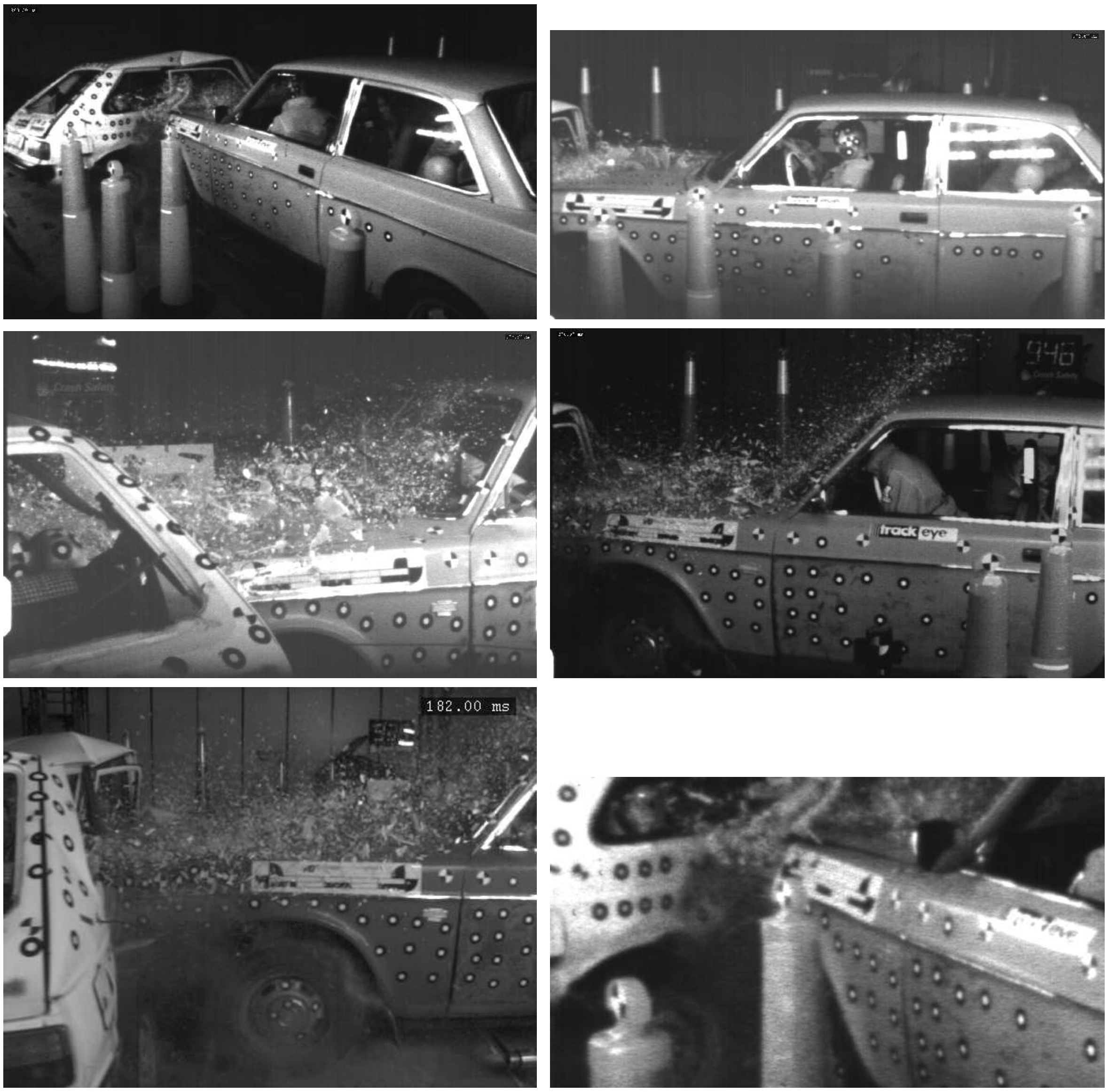

Figure 1: Images from cameras 1–5, and an enlargement from camera 1, from about the middle of the IV crash test sequence.

feature extraction failures, exacerbated by the rather weak signals from the targets used[1]. We tried to work around this by using a contour based target extractor which can also key off the exterior rims of the circular targets, as these black-to-background transitions are very strong for the white car. Unfortunately, the current contour detector is too sensitive to failures in edgel chain extraction. These problems can probably be reduced by rethinking our feature extraction strategy, using a correlation model containing both the internal and external edges of the targets. In any case, we would expect them to be much less severe if whiter targets were used. A further problem with the current tracker is that because nearby targets have very similar images, the tracks occasionally jump between targets. We need to integrate a better motion model and local consistency constraints to reduce this.

The work on bundle adjustment discussed under task 2.2 above is also directly motivated by the needs of the I.V. demonstrator.

Lund (1mm) have made some preliminary experiments on the IV dummy test sequence, where the Euclidean reconstruction of the motion of the dummy's head have been obtained, assuming the same (unknown) focal length of the cameras and that the other intrinsic parameters are known. Recently, investigations on the IV car crash sequence indicate that Euclidean reconstruction of the feature points of the cars can be obtained by assuming known intrinsic parameters except different (and unknown) focal lengths of the cameras.

WP 3: Algebraic Symbolic Reasoning

Task 3.1: System for Geometric Theorem Proving in 2D and 3D

INRIA Sophia (3mm) have continued their study of applications of algebraic geometric reasoning to computer vision. They have developed a variation of these techniques, allowing automatic minimal parameterization of 3D constrained environments given as sets of primitives (currently points, lines and planes). The constraints currently dealt with are incidence relations, parallelism and orthogonality and are transformed into algebraic formulae. The system processes the constraints to separate the initial set of variables (the coordinates of the primitives) into two subsets: a vector of variables $\mathcal{P}$ that can be used as a minimal parameterization and a vector of variables $\mathcal{D}$ that can be deduced from $\mathcal{P}$. The system also provides a set of equations that can be used to compute $\mathcal{D}$ given $\mathcal{P}$. This system has been used in conjunction with a bundle adjustment technique to obtain 3D reconstructions satisfying both image constraints and given 3D constraints. The method has been successfully applied to the "Place des arcades" scene and provides an accurate model in which all coplanarities and orthogonalities are perfectly satisfied [11].

Advances have also been made in automatic theorem generation in 2D [10] but the generalization of this work to 3D is more difficult than expected. This, and the computational cost of these kinds of techniques makes their use in computer vision impractical with current computer technology.

Grenoble (1.5mm) : Grenoble's work on geometric reasoning using Clifford algebra and term-rewriting has been continued [19]. In particular, we have systematically studied general rewriting techniques for our special purposes and developed a sample rewriting package in Objective CAML in combination with the computer algebra system Maple [18]. Moreover, our method and implementation have been extended for geometric reasoning in 3D [12]. A number of geometric theorems have been proved effectively by using our prover in a coordinate-free manner [52]. Application of the method and software to computer vision problems is under investigation.

[1]Unfortunately, the white ink used to screen print the targets was not sufficiently opaque or reflective, so they turn out to be much dimmer than, *e.g.* the white background of the car (see figure).

Lund (2mm): Aiming at the integration of geometrical knowledge into computer vision methods, Lund has concentrated on numerical approaches to geometric reasoning under uncertainty. Their work relies heavily on basic tools from WP1 and WP2, adjusted to take account of the particular requirements imposed by VR and AR on model building techniques. 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 and depth formalism for recovery of projective structure and motion [42, 43], their work has been directed towards the numerical enforcement of affine and Euclidean constraints on the reconstruction. For instance, as has been shown in [43], the formalism enables the efficient handling of a priori information — that certain points are coplanar, or that the line connecting two points is parallel to the line connecting two other points or some plane, *e.g.* the ground plane. The methods allow simultaneous handling of any number of points in any number of images. Another advantage, making it possible to cope with occlusions, is that all points need not be detected in all images. In [43], it is also shown how information from maps of the scene can be taken into account to recover the geometry of the scene more accurately, as well as how known relative placement of the viewpoints can be used to achieve an affine or Euclidean reconstruction, even in the case of uncalibrated cameras.

IGD (2mm): See task 3.2.

Task 3.2: Application of Geometric Reasoning to Augmented Reality

IGD (3mm): FHG has developed a new calibration and tracking tool for Augmented Reality applications. In the past FHG used only model based algorithms to recover the camera parameters. In many practical applications the access to 3D information on the scene is limited or difficult, so more flexible approaches are needed. Furthermore, prior knowledge or partial information, such as a map of the site, represents useful scene constraints that can be exploited to develop robust and reliable algorithms. The goal of the new tool is twofold:

First, it provides an interface, which enables users to interactively define image features (points and lines) and set constraints on them (metric-input, alignment, coplanarity, orthogonality and parallelism). The tool also contains a special viewer for a map, on which the camera position can be set. A data exchange file describing the user settings has been defined in collaboration with INRIA-Sophia. This file will be exploited by the calibration and reconstruction tools of the other partners.

Secondly, it provides an experimental environment to develop and test algorithms of WP1 and WP2, relevant for Augmented Reality. The tool enables dealing with discrete multiple images and video sequences. The basic functionalities are epipolar geometry and homography computations with linear methods and non-linear optimization. The image features are given interactively and tracked by correlation in case of image sequences. During this first software development part, FHG focused on the case with known internal camera parameters and implemented the structure and motion algorithms for planar and non-planar scenes proposed by Lund University and INRIA Grenoble.

After motion determination, a graphical object can be integrated into the images and moved interactively by the user. It sometimes happens that a real existing object must first be erased from the images before the new virtual one can be inserted. We call this preliminary step "Diminished Reality". FHG achieves this task in the case of planar background partially visible in the different images. Using registration by homography, the pixels of the object in foreground are replaced by the ones in background.

Sophia (4mm) : As identified previously within WP3, the usual calibration techniques sometimes perform poorly when dealing with difficult scenes such as the London sequence. Thus the need for a more flexible calibration method that will take advantage of known properties of the 3D scene. In particular, maps constitute a very rich source of information. Sophia has thus concentrated on the development of such a tool that will recover camera calibration given an image, a map and a set of known 3D geometric properties (collinearity, coplanarity and orthogonality for the time being) over 3D primitives (points up to now). Starting with approximate initial information, the method attempts to refine the calibration of the camera by minimizing both the reprojection error of the 3D points and the fitting error of the 3D geometric model that the 3D points must satisfy. To do so, an iterative method is used which alternatively *(i)* minimizes the 3D geometric fitting error using fixed calibration information and *(ii)* minimizes the reprojection error by modifying the calibration information while fixing the positions of the 3D points.

WP 4: Management

The management of the project is going smoothly. The collaboration is working well and all partners are pulling their weight in the project. There are a few minor changes of workplan discussed below, but nothing that should seriously delay the completion.

Changes of Work Plan

Both the WP1 and WP2 industrial data sets were delayed last year, owing to the reluctance of the industrial partners clients to release data which might contain something commercially sensitive.

I.V. Image Systems managed to negotiate with the Swedish national testing lab VTI to film a test of a side-on crash of two cars specially marked with circular targets. The sequence — containing around 900 images from 5 cameras — was prepared and filmed in November 1997 and processed and delivered by early February 1998. The lateness of the test data and some problems with tracking the fainter-than-expected targets (see the discussion for worktask 2.3) have delayed delivery of the WP2 libraries, but we still expect to finish the WP2 demonstrator on time.

Unfortunately, none of IMETRIC's clients have proved willing to release suitable data. For the moment, we have agreed to continue WP 1.4 without the industrial test data (deliverable 1.4/12). This should not cause any delays in WP1 or materially affect the IMETRIC demonstrator, which will be shown on mock-up data if necessary. The existing IMETRIC data (1.4/6) should be adequate for software development and testing.

As discussed in the month 12 report and review, the workplan of WP3 has been reorganized to better reflect current industrial needs and technological capabilities. IGD's industrial goal and the topic of the WP 3.2 demonstrator have changed from building static 3D models for virtual reality, to camera-scene calibration for augmented reality. WP 3.1, which investigated automated algebraic-geometric reasoning methods for these sorts of visual reconstruction or calibration problems, has now finished. Its main conclusions were that although algebraic methods have advanced very significantly over the last few years, they are still not sufficiently powerful for routine use on large problems with many constraints. Also, it is still very unclear how to handle uncertain data using them. For this reason, the implementation of WP 3.2 will rely more on numerical methods and less on algebraic ones than we had originally intended. To take account of the extra work on numerical methods required for this, we have reassigned 4 man months from WP 3.1 to WP 3.2. The changes will also delay the WP 3.2 demonstrator, which will now be shown in month 36 rather than month 30.

Workshops and Exchanges

The ESPRIT projects VANGUARD, PANORAMA and CUMULI jointly organized a very successful academic workshop "3D Structure from Multiple Images of Large-Scale Environments" (SMILE'98) in Freiberg on 6–7 June 1998, immediately after the 5$^{\text{th}}$ European Conference in Computer Vision ECCV'98. There were 19 papers (including an invited presentation by Richard Hartley and 6 other papers from outside the consortia) and 5 live demos. The proceedings will be published by Springer-Verlag. The workshop was the best attended and most highly rated of the conference, drawing an audience of over 100 people.

FHG initiated and co-organized the first International Workshop on Augmented Reality (IWAR'98) at San Fransisco. The workshop had a strong industrial component represented by Rockwell, Boeing, InterSense and Hughes. It is planned to be a regular event and to take place in Europe in 1999. FHG also intends to present Augmented Reality and the new CUMULI tool at Cebit'99 and/or the 1999 Hannover Industial Fair.

Dongming Wang from Grenoble co-organized the second International Workshop on Automated Deduction in Geometry (ADG'98) in Beijing, which is dedicated to techniques and tools for automated geometric deduction with special attention to the interaction of the subject with the other disciplines and its potential industrial applications in geometric modeling, CAGD/CAD, computer vision, and robotics.

Work from CUMULI was also presented in several talks at a domestic Swedish conference, organized by the Swedish Society of Automated Image Analysis, with 120 participants, of which about one fourth came from industry.

INRIA Grenoble and Sophia have contacted the French industrial transfer organization GRECO-Informatique, who have tentatively agreed to host an industrial workshop on CUMULI related technology in Paris in September 1999. This will be open to all industrialists who wish to attend.

There have also been some local exchanges within the consortium. Fredrik Kahl from Lund is currently visiting INRIA Grenoble for 3 months, working on the WP2 test data. Magnus Olsson from IV has visited Lund, and Lund have also helped IV with some MAPLE calculations needed for CUMULI software development. Didier Stricker from FGH visited Lund in September 1997 to discuss interactions on WP 3.

Deliverables

CUMULI has had the following deliverables in the last 12 months:

The I.V. Image Systems car crash test data set (originally a part of deliverable 2.3/6, but delayed at month 6 as our original schedule was unrealistic) was delivered in February 1998.

The demonstration 3.1/18 of geometric reasoning software will be shown at the review. The system 'straightens' an approximate initial scene reconstruction by enforcing a set of user-supplied geometric constraints such as 3D coplanarity, orthogonality or parallelism. Ritt-Wu characteristic sets (an algebraic deduction method) are used to derive a minimal parametrization for the 3D scene geometry from the constraints, then a constrained bundle adjustment over these parameters optimizes the recovered 3D structure subject to the constraints. The methods developed are described in [11].

For the joint overview article 1.1/24 and 2.1/24 on discrete and continuous multi-image projective reconstruction, we find that we have too much material to fit into a single reasonably-sized paper. We are currently considering how best to present this work. A summary of what we have is included in the appendix.

For the remaining articles which are due we have selected the following papers: [50, 15, 31] for 1.2/24 'Short article: algebraic multi-image Euclidean reconstruction'; [40] for 2.2/24 'Short article: continuous multi-image Euclidean reconstruction'; and [11, 43] for 3.2/24 'Article: automated geometric reasoning for 3D properties'. These papers are included in the appendix to this report. More generally, European Community support for CUMULI over the last 12 months has aided and been acknowledged in (at least) the following papers [1, 2, 3, 4, 5, 6, 17, 16, 15, 21, 22, 24, 30, 32, 34, 31, 33, 40, 43, 50, 51, 35, 44, 12, 19, 18, 52, 11, 10]. This document itself is the month 24 Periodic Progress Report 4.1/24.

The following software deliverables are available for inspection on request:

(1.3/18) 'Initial algorithms for camera pose from lines and conics': various MATLAB pose and relative orientation routines from Grenoble and Lund. (NB: All three industrial partners have now used the point-based camera pose and orientation methods developed in WP1, so these are very successful deliverables).

(1.1/24 + 2.1/24) 'Library for discrete multi-image projective reconstruction' and 'library for incremental multi-image projective reconstruction': various MATLAB projective reconstruction routines from Lund and Grenoble, mainly factorization-based methods. We have merged these two libraries as we find that for factorization methods there is less difference between discrete and incremental methods than we had initially thought.

(1.2/24) 'Algorithms for discrete multi-image Euclidean reconstruction': currently MATLAB bundle adjustment routines from Lund.

(2.2/24) 'Algorithms for continuous multi-image Euclidean reconstruction': C/C++ motion analysis routines from Sophia (point correspondence, segmentation and the detection of specific types of motions for which special-purpose calibration methods apply).

As mentioned above, the tracking library 2.3/24 will be slightly delayed owing partly to the late delivery of the WP2 test sequences, and partly to difficulties caused by the relatively weak signals from the targets used in it. A self-calibrating bundle adjustment routine capable of handling motion estimation in non-static scenes is also being implemented for use in the WP 2 demonstrator.

Summary

At present CUMULI is on basically target and running to budget. The total man-months consumed to date (months 0–24) 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.

References

[1] K. Åström. Using combinations of points, lines and conics to estimate structure and motion. In *Proc. Symposium on Image Analysis, SSAB, Uppsala, Sweden*, pages 61–64, 1998.

[2] K. Åström and A. Heyden. Continuous time matching constraints for image streams. *Int. Journal of Computer Vision*, 28(1):85–96, 1998.

WORK TASK	MONTH	Gren	Soph	Lund	I.V.	Imetric	Fraun	TOT.
1 Discrete Images		12/17	10/12	13/17	1/2	7.2/17		43.2/65
1.1 Multi-camera Geometry	0–24	4/4	5/5	4/4	0.7/1	1/1		14.7/15
1.2 Euclidean Structure	0–36	3/4	4/5	3/4		0.6/1		10.6/14
1.3 Features & Uncertainty	0–36	2/4		3/4	0.3/1	0.6/1		5.9/10
1.4 Transfer Imetric	0–36	3/5	1/2	3/5		5/14		12/26
2 Image Streams		13/15	10/11	10/14	4.6/20			37.6/60
2.1 Incremental Geometry	0–24	4/5	5/5	5/5	0.8/1			14.8/16
2.2 Continuous Constraints	6–30	4/5	5/5	3/5	0.3/1			12.3/16
2.3 Transfer I.V.	0–36	5/5	–/1	2/4	3.5/18			10.5/28
3 Reasoning		3/3	16/24	4/7			10/22	33/56
3.1 Geometric Reasoner	0–18	3/3	12/12	4/4			7/7	26/26
3.2 Visual Applications	12–36		4/12	–/3			3/15	7/30
TOTAL RESEARCH		28/35	36/47	27/38	5.6/22	7.2/17	10/22	113.8/181
4 Management	0–36	9/13	2/3	2/3	1.5/3	0.6/1	1.3/2	16.4/25
TOTAL PROJECT		37/48	38/50	29/41	7.1/25	7.8/18	11.3/24	130.2/206

Table 1: Man-months consumed in months 0–24 *vs.* total budgeted man-months for entire task, per work task and per site.

[3] K. Åström, F. Kahl, A. Heyden, and R. Berthilsson. A statistical approach to structure and motion from image features. In *Statistical Techiques in Pattern Recognition, Sydney, Australia*, pages 929–935, 1998.

[4] R. Berthilsson. Affine correlation. In *Proc. Int. Conf. Pattern Recognition, Brisbane, Australia*, pages 1458–1467, 1998.

[5] R. Berthilsson. A statistical theory of shape. In *Statistical Techiques in Pattern Recognition, Sydney, Australia*, pages 677–682, 1998.

[6] R. Berthilsson and A. Heyden. Recognition of planar point configurations using the density of affine shape. In *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, 1998.

[7] 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.

[8] 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.

[9] 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.

[10] D. Bondyfalat, Mourrain. B, and T. Papadopoulo. Automated method for theorem discovering. Workshop ADG'98, China, 1998.

[11] Didier Bondyfalat and Sylvain Bougnoux. Imposing Euclidean constraints during self-calibration processes. In R. Koch and L. Van Gool, editors, *Proceedings of SMILE Workshop on Structure from Multiple Images*, Lecture Notes in Computer Science. Springer Verlag, Lecture Notes in Computer Science, 1998.

[12] Thierry Boy de la Tour, Stéphane Fèvre, and Dongming Wang. Clifford term rewriting for geometric reasoning in 3d. Workshop ADG'98, China, 1998.

[13] 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.

[14] A. Heyden *et al*. Discrete and continuous multi-camera projective reconstruction: A summary of work done within CUMULI. unpublished, 1998.

[15] O. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. In *Proc. 5th European Conference on Computer Vision, Freiburg, Germany*, pages 36–52, June 1998.

[16] Olivier Faugeras and Théodore Papadopoulo. Grassmann-Cayley algebra for modeling systems of cameras and the algebraic equations of the manifold of trifocal tensors. *Transaction of the Royal society A*, May 1998.

[17] Olivier Faugeras and Théodore Papadopoulo. A nonlinear method for estimating the projective geometry of three views. In *Proc. 6^{th} Int. Conf. Computer Vision*, pages 477–484, Bombay, India, January 1998. IEEE Computer Society Press.

[18] Stéphane Fèvre and Dongming Wang. Combining algebraic computing and term-rewriting for geometry theorem proving. In *Proceedings of the 4th International Conference on Artificial Intelligence and Symbolic Computation*, LNAI **1476**. Springer-Verlag, Berlin Heidelberg, 1998. To appear.

[19] Stéphane Fèvre and Dongming Wang. Proving geometric theorems using clifford algebra and rewrite rules. In *Proceedings of the 15th International Conference on Automated Deduction*, LNAI **1421**, pages 17–32. Springer-Verlag, Berlin Heidelberg, 1998.

[20] 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.

[21] A. Heyden. A common framework for multiple-view tensors. In *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, 1998.

[22] A. Heyden and K Åström. Minimal conditions on intrinsic parameters for Euclidean reconstruction. In *Proc. 2nd Asian Conf. on Computer Vision, Hong Kong, China*, pages 169–176, 1998.

[23] A. Heyden, R. Berthilsson, and G. Sparr. Recursive structure and motion from image sequences using shape and depth spaces. *Int. Journal of Computer Vision*, 1998. submitted.

[24] A. Heyden and F. Kahl. Reconstruction from affine cameras using closure constraints. In *Proc. Int. Conf. Pattern Recognition, Brisbane, Australia*, pages 56–65, 1998.

[25] 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.

[26] 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.

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

[28] 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.

[29] F. Kahl. Critical motion sequences. unpublished, 1998.

[30] 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*, pages 939–944, 1998.

[31] F. Kahl and A. Heyden. Robust self-calibration and Euclidean reconstruction via affine approximation. In *Proc. Int. Conf. Pattern Recognition, Brisbane, Australia*, pages 47–55, 1998.

[32] F. Kahl and A. Heyden. Structure and motion from points, lines and conics with affine cameras. In *Proc. 5th European Conf. on Computer Vision, Freiburg, Germany*, 1998.

[33] F. Kahl and A. Heyden. Structure and motion with affine cameras. In *Proc. Symposium on Image Analysis, SSAB, Uppsala, Sweden*, pages 57–60, 1998.

[34] F. Kahl and A. Heyden. Using conic correspondences to estimate the epipolar geometry. In *Proc. 6th Int. Conf. on Computer Vision, Mumbai, India*, pages 761–766, 1998.

[35] G. Klinker, D. Stricker, and D. Reiners. The use of reality models in augmented reality applications. In R. Koch and L. Van Gool, editors, *Proceedings of SMILE Workshop on Structure from Multiple Images*, Lecture Notes in Computer Science. Springer Verlag, Lecture Notes in Computer Science, 1998.

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

[37] Théo Papadopoulo and Olivier Faugeras. A new characterization of the trifocal tensor. In *Proc. 5^{th} European Conf. Computer Vision*, Fribourg, Germany, June 1998.

[38] J. Philip. A non-iterative algorithm for determining all essential matrices corresponding to five point pairs. *Photogrammetric Record*, 15(88):589–599, October 1996.

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

[40] P. Renault, O. Faugeras, and T. Viéville. Continuous multi-image preprocessing for Euclidean reconstruction. RR 3482, INRIA, September 1998.

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

[42] 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.

[43] G. Sparr. Euclidean and affine structure/motion for uncalibrated cameras from affine shape and subsidiary information. In R. Koch and L. Van Gool, editors, *Proceedings of SMILE Workshop on Structure from Multiple Images*, Lecture Notes in Computer Science. Springer Verlag, Lecture Notes in Computer Science, 1998.

[44] D. Stricker, G. Klinker, and D. Reiners. A fast and robust line-based optical tracker for augmented reality applications. In *Proceedings of first International Workshop on Augmented Reality*, 1998.

[45] 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.

[46] 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.

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

[48] 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.

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

[50] B. Triggs. Autocalibration from planar scenes. In *European Conf. Computer Vision*, pages I 89–105, Freiburg, June 1998.

[51] B. Triggs. Optimal estimation of matching constraints. In R. Koch and L. Van Gool, editors, *Workshop on 3D Structure from Multiple Images of Large-scale Environments SMILE'98*, Lecture Notes in Computer Science. Springer Verlag, 1998.

[52] Dongming Wang. Clifford algebraic computing and term-rewriting for geometric theorem proving. In *Proceedings of the 2nd International Theorema Workshop*, 1998. available as RISC-Linz Report no. 98-10, Johannes Kepler University, Austria.

[53] W. Wunderlich. Rechnerische rekonstruktion eines ebenen objekts aus zwei photographien. In *Mittsilungen Geodät. Inst. TU Gras,* Folge 40 (festschrift K. Rimmer sum 70. Geburstag), pages 365–77, 1982.

Appendix: Deliverable Papers

'Overview article: Discrete and Continuous Multi-image Projective Reconstruction' (1.1/24 and 2.1/24): We received so much material for this deliverable that it was impossible to fit it into a single reasonably-sized paper. We are currently considering how best to present this work. The text given here is just a brief summary of what we have.

- A. Heyden *et al*. Discrete and continuous multi-camera projective reconstruction: A summary of work done within CUMULI. unpublished, 1998.

'Short article: algebraic multi-image Euclidean reconstruction' (1.2/24):

- B. Triggs. Autocalibration from planar scenes. In *European Conf. Computer Vision*, pages I 89–105, Freiburg, June 1998.
- O. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. In *Proc. 5th European Conference on Computer Vision, Freiburg, Germany*, pages 36–52, June 1998.
- F. Kahl and A. Heyden. Robust self-calibration and Euclidean reconstruction via affine approximation. In *Proc. Int. Conf. Pattern Recognition, Brisbane, Australia*, pages 47–55, 1998.

'Short article: continuous multi-image Euclidean reconstruction' (2.2/24):

- P. Renault, O. Faugeras, and T. Viéville. Continuous multi-image preprocessing for Euclidean reconstruction. RR 3482, INRIA, September 1998.

'Article: automated geometric reasoning for 3D properties' (3.2/24):

- Didier Bondyfalat and Sylvain Bougnoux. Imposing Euclidean constraints during self-calibration processes. In R. Koch and L. Van Gool, editors, *Proceedings of SMILE Workshop on Structure from Multiple Images*, Lecture Notes in Computer Science. Springer Verlag, Lecture Notes in Computer Science, 1998.
- G. Sparr. Euclidean and affine structure/motion for uncalibrated cameras from affine shape and subsidiary information. In R. Koch and L. Van Gool, editors, *Proceedings of SMILE Workshop on Structure from Multiple Images*, Lecture Notes in Computer Science. Springer Verlag, Lecture Notes in Computer Science, 1998.

The various deliverable libraries and copies of the many other papers published under CUMULI are available for inspection on request.