

Preface

This volume contains the final versions of papers originally given at the workshop *Vision Algorithms: Theory and Practice*, which was held on 21–22 September 1999 during the *Seventh International Conference on Computer Vision* at the Corfu Holiday Palace Hotel in Kanoni, Corfu, Greece.

The subject of the workshop was algorithmic issues in computer vision, and especially in vision geometry: correspondence, tracking, structure and motion, and image synthesis. Both theoretical and practical aspects were considered. A particular goal was to take stock of the ‘new wave’ of geometric and statistical techniques that have been developed over the last few years, and to ask which of these are proving useful in real applications. To encourage discussion, we asked the presenters to stand back from their work and reflect on its context and longer term prospects, and we encouraged the audience to actively contribute questions and comments. The current volume retains some of the flavour of this, as each paper is followed by a brief edited transcript of the discussion that followed its presentation.

The theme was certainly topical, as we had 65 submitted papers for only 15 places (an acceptance rate of only 23%), and around 100 registered participants in all (nearly 1/3 of the ICCV registration). With so many submissions, there were some difficult decisions to make, and our reviewers deserve many thanks for their thoroughness and sound judgment in paper evaluation. As several authors commented, the overall quality of the reviews was exceptionally high. The accepted papers span the full range of algorithms for geometric vision, and we think that their quality will speak for itself.

To complement the submitted papers, we commissioned two invited talks “from the shop floor”, two “expert reviews” on topical technical issues, and a panel session.

The invited talks were by two industry leaders with a great deal of experience in building successful commercial vision systems:

- Keith Hanna of the Sarnoff Corporation described Sarnoff’s real time video alignment and annotation systems, which are used routinely in applications ranging from military reconnaissance to inserting advertisements and annotations on the Super Bowl field. This work is presented in the paper *Annotation of Video by Alignment to Reference Imagery* on page 253).
- Luc Robert of REALViZ S.A. described REALViZ’s MatchMover and ReTimer post-production systems for movie special effects, which are used in a number of large post-production houses. Unfortunately there is no paper for this presentation, but the discussion that followed it is summarized on page 265.

Both presenters tried to give us some of the fruits of their experience in the difficult art of “making it work”, illustrated by examples from their own systems.

The two “expert reviews” were something of an experiment. Each was a focused technical summary prepared jointly by a small team of people that we consider to be domain experts. In each case, the aim was to provide a concise technical update and state

of the art, and then to discuss the advantages of the various implementation choices in a little more depth.

The motivation for these review sessions was as follows. As active members of the vision community and referees of many papers, we continually find that certain basic topics are poorly understood. This applies particularly to areas where a cultural split has occurred, with two or more camps following more or less separate lines of development. There are several such splits in the vision community, and we feel that every effort must be made to heal them. For one thing, it is fruitless for one group to reduplicate the successes and failures of another, or to continue with a line of research that others know to be unprofitable. More positively, intercommunication breeds innovation, and it is often at the boundaries between fields that the most rapid progress is made. The workshop as a whole was intended to take stock of the rapid progress made in vision geometry over the past decade, and hopefully to narrow the gap between “the geometers” and “the rest”. Within this scope, we singled out the following two areas for special treatment: (i) the choice between direct and feature-based correspondence methods; and (ii) bundle adjustment.

Direct versus feature-based correspondence methods: One of the significant splits that has emerged in the vision community over the past 15–20 years is in the analysis of image sequences and multi-view image sets. Two classes of techniques are used:

- **“Feature-based” approaches:** Here, the problem is broken down into three stages: (i) local geometric features are extracted from each image (*e.g.* “points of interest”, linear edges . . .); (ii) these features are used to compute multi-view relations, such as the epipolar geometry, and simultaneously are put into correspondence with one another using a robust search method; (iii) the estimated multi-view relations and correspondences are used for further computations such as refined correspondences, 3D structure recovery, plane recovery and alignment, moving object detection, *etc.*
- **“Direct” approaches:** Here, rather than extracting isolated features, dense spatio-temporal variations of image brightness (or color, texture, or some other dense descriptor) are used directly. Instead of a combinatorial search over feature correspondences, there is a search over the continuous parameters of an image motion model (translation, 2D affine, homographic), that in principle establishes dense correspondences as well as motion parameters. Often, a multi-scale search is used.

The experts in this session were P. Anandan & Michal Irani, who present the direct approach in the paper *About Direct Methods* on page 267, and Phil Torr & Andrew Zisserman, who present the feature-based approach in the paper *Feature Based Methods for Structure and Motion Estimation* on page 278. In each case, the authors try: (i) to give a brief, clear description of the two classes of methods; (ii) to identify the applications in which each has been most successful; and (iii) to discuss the limitations of each approach. The discussion that followed the session is summarized on page 295.

Bundle adjustment for visual reconstruction: Bundle adjustment is the refinement of visual reconstructions by simultaneous optimization over both structure and camera parameters. It was initially developed in the late 1950’s and 1960’s in the aerial photogrammetry community, where already by 1970 extremely accurate reconstruction of

networks of thousands of images was feasible. The computer vision community is only now starting to consider problems of this size, and is still largely ignorant of the theory and methods of bundle adjustment. In part this is because cultural differences make the photogrammetry literature relatively inaccessible to most vision researchers, so one aim of this session was to present the basic photogrammetric techniques from a computer vision perspective. The issues raised in the session are reported in the survey paper *Bundle Adjustment — A Modern Synthesis* on page 298. This paper is rather long, but we publish it in the hope that it will be useful to the community to have the main elements of the theory collected in one place.

The workshop ended with an open panel session, with Richard Hartley, P. Anandan, Jitendra Malik, Joe Mundy and Olivier Faugeras as panelists. Each panelist selected a topic related to the workshop theme that he felt was important, and gave a short position statement on it followed by questions and discussion. The panel finished with more general discussion. A brief summary of the discussion and the issues raised by the panel is given on page 376.

Finally, we would like to thank the many people who helped to organize the workshop, and without whom it would not have been possible. The scientific helpers are listed on the following pages, but thanks must also go to: John Tsotsos, the chairman of ICCV'99, for his help with the logistics and above all for hosting a great main conference; to Mary-Kate Rada and Maggie Johnson of the IEEE Computer Society, and to Danièle Herzog of INRIA for their efficient organizational support; to the staff of the Corfu Holiday Palace for some memorable catering; and to INRIA Rhône-Alpes and the IEEE Computer Society for agreeing to act as sponsors.

June 2000

Bill Triggs, Andrew Zisserman and Richard Szeliski

Workshop Organizers

Bill Triggs	INRIA Rhône-Alpes, Grenoble, France
Andrew Zisserman	Dept. of Engineering Science, Oxford University
Richard Szeliski	Microsoft Research, Redmond, WA

Sponsors

INRIA Rhône-Alpes
IEEE Computer Society

Program Committee

Michael Black	Michal Irani	Harpreet Sawhney
Stefan Carlsson	Philip McLauchlan	Amnon Shashua
Olivier Faugeras	Steve Maybank	Chris Taylor
Andrew Fitzgibbon	John Oliensis	Phil Torr
Wolfgang Förstner	Shmuel Peleg	Luc Van Gool
Pascal Fua	Jean Ponce	Thierry Viéville
Greg Hager	Long Quan	Zhengyou Zhang
Richard Hartley	Ian Reid	Kalle Åström

Invited Speakers

Keith Hanna	Sarnoff Corporation, Princeton, NJ
Luc Robert	REALViZ S.A., Sophia Antipolis, France

Panelists

P. Anandan	Richard Hartley	Joe Mundy
Olivier Faugeras	Jitendra Malik	

Additional Reviewers

Peter Belhumeur	John Illingworth	Cordelia Schmid
Ingemar Cox	Michael Isard	Steve Seitz
Patrick Gros	Bart Lamiroy	Harry Shum
Radu Horaud	Jitendra Malik	Ramin Zabih
	Peter Meer	

Review Helpers

Yaron Caspi	Yanlin Guo	Garbis Salgian
Athos Georgiades	Steve Hsu	Peter Sturm
Jacob Goldberger	Robert Mandelbaum	Hai Tao
	Bogdan Matei	

Student Helpers and Session Transcripts

Eric Hayman	Antonio Criminisi	Geoff Cross	Joss Knight
-------------	-------------------	-------------	-------------