I THE GRANT PROPOSAL

Introduction by the Editors

1 ORGANISATION OF THE BOOK

The research reported in this book is the outcome of a large multi-site industry/academe consortium funded (c.£1.2m) by the U.K.'s Alvey Programme in advanced information technology. The research papers are grouped under the three projects that formed the organisational structure of the consortium. The section on each project has an editorial introduction based mainly on excerpts from the original research proposal. These excerpts have been left largely verbatim to preserve the flavour of the thinking that guided the consortium's efforts.

The editors have attached to their introductory sections reviews entitled ***What Really Happened?*** These allow the reader some insight into the strengths and weaknesses of pursuing fundamental computer vision research within a large multi-site consortium. The editors draw their own conclusions about how things went as a whole in a final ***Summing Up*** at the end of this section.

In this opening section to the book we provide a brief background to the origins of the consortium. We then give an overview of the research grant proposal, and finally we thank the many people who helped the consortium get underway.

2 BACKGROUND

The research proposal for the consortium was entitled *3D Surface Representations and 3D Model Invocation from Stereoscopic Cues.* The proposal arose directly from the 3D IKBS-VISION PROJECT written by Mayhew as part of the *IKBS Architecture Study* (1983) commissioned by the Science & Engineering Research Council (SERC) and the Department of Industry.

Mayhew's 3D IKBS-VISION PROJECT document evolved from lengthy discussions over a period of a year or so between many interested parties, industrial and academic. A meeting held on 30 September 1983 of 34 scientists and engineers from 16 institutions endorsed its main outlines and discussed how a detailed research proposal should be developed. The outcome was *An Overview of the 3D IKBS-VISION PROJECT* written by Frisby. This served as a scene-setting appendix to the final research grant proposal of the consortium submitted to the Alvey IKBS Committee for its approval. An abridged version of this *Overview* is given in the next section by way of introducing the consortium as a whole. It is reproduced largely verbatim, thereby ensuring an authentic introduction to the background leading to the papers contained in this book, and its 1983 date should be borne in mind.

Most sites in the consortium began their work on 1st October 1984, with funding lasting for 3 years.

3 OVERVIEW OF THE RESEARCH GRANT PROPOSAL (Written 1983)

3.1 Research Goals and Target Applications

The proposed research is concerned with solving problems in the design of a general purpose machine vision system capable of delivering useful 3D visual competences to an automated mobile robot system and/or an assembly task work station. Lack of a good visual 'front end' capable of extracting useful 3D information about a robot's immediate environment has been one of the major factors preventing the automation of many industrial tasks, including vehicle guidance and pick-and-place manipulations. The potential value of equipping robots with this competence is widely recognised internationally and various industrial competitors overseas are working on the problems involved. It is important that the UK establishes its ability to compete in this field.

The field of 3D machine vision is a large one, with many different approaches being explored. The guiding strategy adopted in the 3D IKBS-VISION PROJECT has been to select an approach which exploits the existing research strengths of certain UK academics while meeting the expressed interests of the industrial collaborators. The resulting proposals are not therefore presented as an exhaustive account of all that 'could or should' be done in the UK in this major field. Rather, they are offered as a focussed attack on some central problems in image understanding using an approach that this group of academics and their industrial collaborators believes will lead to useful applications in the medium term (3-5 years). The programme will be scientifically rewarding by furthering basic knowledge about some classical issues in AI-oriented image understanding.

3.2 Status as an Alvey IKBS Research Theme

SERC's *IKBS Architecture Study* classified the 3D IKBS-VISION Project as a 'Research Theme'. We are content with that designation because it suits its character, namely a club[1] of cooperating scientists and engineers engaged in basic research who intend to deliver results applicable to a variety of possible products by tackling various fundamental problems in building a general purpose 3D vision system. The view of the club is that it is best to address and solve these general questions prior to launching a demonstrator project.

Many Alvey planning papers have advised that basic research must be recognised as an essential part of the Alvey endeavour. For example, the *IKBS Architecture Study* recommended that "high quality, speculative research, both

[1] The 'club' referred to here is not the Alvey Vision Club. The latter was founded later and had a much larger membership, joining together participants in Alvey's IKBS and Man-Machine Interface computer vision programmes.

theoretical and practical .... is extremely important for the health of the programme overall .... it must not become the poor relation" (Main Report, 1983, Vol 1, Section 8.3.5). In the field of the present proposal, concentration on particular applications at this early stage could easily stifle progress by generating application-specific strategies. That said, however, it should also be emphasised that the proposal is *not* directed towards the design of a 'completely general purpose visual system', which would of course be a quite premature goal at the present time. Rather, the proposal is solidly based on substantial recent achievements by various members of the club in the domain of general purpose 3D vision from optic flow and stereoscopic cues, and the domain of model-based interpretation. This prior work, which has obvious potential application to vehicle guidance and robot assembly tasks in many industrial settings, augurs well for the 3D IKBS-VISION Project spawning a demonstrator project in 3-5 years time.

3.3 Relationship to Other Research Programmes

Any worthwhile club is composed of like-minded members sharing certain fundamental preconceptions about how to proceed with their cooperative endeavour. The 3D IKBS-VISION proposal is fortunate to have just that basis, while being large enough to avoid the obvious danger of becoming in-grown. In short, the proposal is supported by a coherent and manageable team, spread widely over various organisations but nevertheless able to work well together, as demonstrated by the successful initial preparation of the 3D IKBS-VISION proposal and by its recent endorsement by the members of the club (no mean achievement given their wide distribution, geographically and academically). We strongly urge that this identity be maintained and exploited by the Alvey Directorate.

3.4 The Two Consortia Comprising the IKBS Research Theme

It is proposed that the two consortia will address problems in acquiring depth information from optical flow and stereoscopic disparity cues. The advantage of using these cues is that they are general purpose in character and that internationally accredited expertise exists in the club in both fields. Each of these topics is the subject of well advanced plans for the creation of academic/industrial research consortia, whose scope includes the utilisation of depth information for 3D model invocation and verification.

A basic objective of the optical flow and stereopsis proposals is to compute a type of 3D scene description that has come to be known in the field of image understanding as the '2.5D Sketch'. This is a viewer-centred representation making explicit the disposition, orientations and distances of visible surfaces in the scene (a meaning that is not to be confused with other uses of this term). Thus the 2.5D Sketch is conceived here as an intermediate-level representation serving as the database for subsequent 3D model invocation as well as being a representation able to support in its own right various lower-level tasks such as some forms of trajectory planning, robot guidance, grasp planning etc (see later). The attempt to build a 2.5D Sketch distinguishes the present proposal from those which proceed directly and solely from 2D image descriptions to 3D models.

The desirability for a wide range of tasks (though not all of course) of an intermediate level representation expressing the depth relationships of visible surfaces is widely recognised in AI laboratories in the USA working on 3D machine vision. As the objects of interest are usually volumetric in character, it makes obvious sense in the design of guidance and manipulation systems capable of dealing with them to provide access to a database that describes scene surfaces in terms of their spatial distributions in 3D. Building such a representation using optical flow and stereoscopic cues has demonstrably paid handsome dividends in biological vision systems, greatly facilitating, for example, the identification of figure from ground. Using a 2.5D Sketch based on these cues should enable better exploitation of the strategy of 'segmentation by recognition', because of the opportunities presented for object recognition via the identification of clusters of 3D surface features characteristic of target objects (see later).

3.5 Constituent Projects of the Consortium Working from Stereoscopic Cues

This consortium [whose work is the subject of this book] proposes three projects, integrated around the theme of building a real-time (<1 sec) stereo processor capable of supporting the creation of 3D surface representations suitable for equipping a robot assembly task workstation with useful 3D visual competences, including the capacity for 3D model invocation and verification. The three projects are as follows:

The PMF Project is based on a computational theory of stereopsis developed over a number of years at Sheffield by Mayhew and Frisby, and recently extended by Stephen Pollard, one of their postgraduates holding a CASE award in collaboration with Margaret McCabe of GEC. This work has produced a simple but highly effective stereo correspondence algorithm called PMF. The underlying principles on which PMF is based are described [in Section II of this book], which also provides illustrations of its performance. The PMF Project is also directed at developing special purpose hardware for implementing PMF.

The 2.5D Sketch Project has as its overall goal the development of ways of using the pointilliste range data delivered by PMF to create descriptions about the nature of the surface boundaries and surface regions that are present in the scene (e.g. methods of labelling depth edges as concave, convex, occluding, extremal, etc. and regions as cylindrical, planar, synclastic, antisynclastic, etc.) This work is central to the objective of building a 2.5D Sketch representation.

The 3D Model-Based Vision Project will be concerned with the utilisation of the 3D scene geometry delivered by the 2.5D Sketch Project for the purposes of 3D model invocation and verification. The distinguishing characteristic of this work will be the development of methods to match 3D structures extracted from the depth map with 3D structures in an object model catalogue (not, as in almost all current object recognition schemes, to match 2D structures extracted from the image with 2D virtual image structures predicted from the object model and imaging geometry).

There will be two main groups involved on the 3D Model-Based Vision Project. The one at Sheffield will be led by Mayhew who proposes to use object models comprised of a 3D relational structure of clusters of 3D surface features that remain stable over a restricted but nevertheless reasonably broad range of viewpoints of a given 3D object. This representation should facilitate the indexing of the model catalogue and the partitioning of the matching task.

Table 1: Participants of the Consortium

GEC Hirst Research Centre
Wembley
Dr Margaret M McCabe* Dr Sheelagh A Lloyd*
Dr Harit Trevedi Dr Jan Wiejak*
Mr Ian Graydon Mr Brendan Ruff
*Dr Bernard Buxton**

IBM UK Ltd Scientific Centre
Winchester
Mr Rodger Hake* Dr John Knapman
Dr Michael Gray Dr Rodney Cuff

AI Vision Research Unit
University of Sheffield
*Prof John E W Mayhew** *Prof John P Frisby**
*Dr Stephen B Pollard** *Dr John Porrill*
*Dr Chris R Brown** *Dr Tony Pridmore*

Department of Artificial Intelligence
University of Edinburgh
Mr Robin Popplestone* Mrs Pat Fothergill
Dr Robert B Fisher Dr Mark J L Orr
Mr Jonathan C Aylett

Department of Computer Science
University of Edinburgh
*Dr Andrew Blake** *Dr Andrew Zisserman*
Dr Gavin Brelstaff

Centre for Research in Perception & Cognition
University of Sussex
*Prof H Christopher Longuet-Higgins**

Individuals denoted with * were either cited in the research grant proposal &/or closely involved in its planning. All others were recruited once the proposal was awarded or deployed to work on it by their companies. Individuals who were on the project throughout its life are shown in italics.

Dr Buxton was closely involved in his role as coordinator of the optic flow consortium in the IKBS Image Interpretation Research Theme. Dr Brelstaff was a CASE student supported by IBM UK Scientific Centre.

The other group is at Edinburgh and will be led by Robin Popplestone[2] who is presently developing the robot language RAPT-2 under a contract from Rutherford Laboratory. One important goal of this work is to equip RAPT-2 with sufficient visual competence for 3D model invocation and verification such that visually acquired data can be used in programming a sequence of actions to cause a robot to unpack, say, a motor from a box and install it in a machined location.

3.6 Participants of the Consortium

See Table 1 above and also the frontispiece photograph.

3.7 Criteria for Evaluation

A document entitled *An Outline of a Strategy for Pattern Analysis Within the Alvey Programme* written by Julian Ullman proposed various criteria against which proposals of the present type should be evaluated. We believe that the 3D IKBS-VISION proposal successfully meets these criteria, as follows:

[*Warning Note from the Editors.* There is probably no limit to what some people will say to get their proposals funded.]

1 *ENHANCEMENT OF UK COMPETENCE & COMPETITIVENESS*? Yes.

2 *WIN PLACES FOR UK IN WORLD MARKETS?* The proposal is concerned with 'basic research' and hence is not directed at developing a market product in the short term. Nevertheless, industrial collaborators believe there should be medium-term market benefits.

3 *WELL-BALANCED PARTNERSHIP WITH OTHER WORK IN IT?* The proposers believe that the work is timely in terms of other work in Information Technology, such as that concerned with parallel processors, the development of robot object level languages etc.

4 *VIABLE TRAJECTORY TO THE MARKET PLACE?* See comments under 2 above. The key feature here is the keen involvement of large UK firms who look forward to marketing applications in the medium-term.

5 *A GENUINELY COLLABORATIVE CONSORTIUM?* The programme has a clear integrating theme. Evidence that participants really do want to work together is the series of fruitful meetings that has already taken place in drawing up the proposal. Also, certain participants are already involved in collaborative research (e.g. via CASE studentships, shared publications).

6 *EVIDENCE OF TECHNICAL COMPETENCE?* The records of the participants should be used to form an opinion on this.

---

[2] Popplestone left the project about the time that the grant was approved, whereupon Pat Ambler (later to become Pat Fothergill) and Bob Fisher took over his role. Fothergill later left the project on her move to Aberdeen.

7 *COMPLEMENTARY CONSTITUENTS OF THE CONSORTIUM?* The proposal blends together diverse competences in optic flow, stereopsis, hardware, image processing, artificial intelligence, robot objective languages, mathematics, psychophysics, and industrial machine vision applications.

8 *MANAGEMENT & ADMINISTRATION?* McCabe and Frisby have been nominated as Industrial and Academic Coordinators respectively of the consortium as a whole, with Lloyd leading the PMF Project, Blake leading the 2.5D Sketch Project, and Mayhew leading the 3D Model based Vision Project. It is envisaged that each project leader will be responsible for drawing up a detailed research plan in conjunction with the Consortium Coordinators. The latter will between them provide an overall organisational structure to ensure good integration of research at different sites, as well as serving as the interface between the Consortium and the Alvey Directorate.

3.8 Concluding Remarks

The targets of the consortium are substantial but not unrealistic. They are founded on a great deal of prior work at Sheffield, Edinburgh, Sussex, IBM and GEC, and they bring together expertise in mathematics, optimisation techniques, VLSI hardware, robot programming languages, industrial machine vision applications, and computational and psychophysical studies of human vision. A great deal of effort will be necessary to attain the targets specified but we think that effort worthwhile, both in terms of the importance of the fundamental scientific issues and in terms of the potential benefits to UK industry.

4 LINKS TO OTHER IKBS CONSORTIA

The *3D IKBS-VISION PROJECT* in fact spawned two[3] large consortia which together were entitled the *Alvey IKBS Image Interpretation Research Theme*. Its Industrial and Academic Coordinators were respectively Margaret McCabe (GEC) and Frisby.

The consortium dealt with in this book took as its starting point the recovery of useful 3D scene geometry from single pairs of stereoscopic images. Its Coordinator was Mayhew and its participating institutions were the universities of Edinburgh, Sheffield and Sussex, GEC (Hirst Research Centre, Wembley), and the IBM UK Ltd Scientific Centre (Winchester).

The Coordinator of the second consortium was Bernard Buxton (GEC). It had the same general goals but started from an analysis of optic flow. Its participating institutions were Queen Mary College (London University), GEC, BAEd and Plessey. Much of the work carried out by GEC in this consortium is presented in the recent book by Murray and Buxton (1990), also published by MIT Press.

It was intended that the two consortia should be run largely independently, but with coordinators attending the major meetings of each consortium to seize on any possibilities for integration that might arise. This is in fact what happened and there is no doubt that various contacts between the two consortia were helpful in developing the research.

[3]There was a third consortium mentioned in the original proposal, to be led by Andrew Sleigh of RSRE Malvern, but its participants soon withdrew to join with other more related work funded by Alvey's Man-Machine Interface Directorate.

5 ACKNOWLEDGEMENTS

As noted at the outset, the research reported in this book was funded by the U.K.'s Alvey Programme in advanced information technology. The latter was launched in the early 1980's as one response to the perceived threat to UK manufacturing industry of Japan's announcement of its intention to build Fifth Generation Computers. Alvey had four branches: the present consortium fell under the *Intelligent Knowledge Based Systems Directorate* led by Dr David Thomas. His enthusiastic support for the consortium, coupled with his vast expertise in finding ways through research funding bureacracies, had much to do with the research getting underway soon after the Alvey Programme received final government approval. We are very grateful for his help.

We are also grateful for the role played by Brian Oakley, Director of the Alvey Programme, who at all times was most supportive of the consortium.

The *IKBS Architecture Study* (which invited Mayhew to write his 3D IKBS-VISION PROJECT proposal) was chaired by John Taylor, co-chaired by Karen Sparck-Jones, and coordinated by Bill Sharpe and Gareth Williams. All members of the consortium have good reason to thank those individuals, and indeed the many others involved in preparation of the *Study*. The reason is that their work coincided with the early stages of the work of the Alvey Committee. This meant that when the Alvey Programme was approved, the existence of the *Study* enabled the IKBS Directorate of Alvey to get off to a flying start, the main outlines of its mission statement being already in place. This in turn led to the work reported here being started about a year or so ahead of the other computer vision projects funded by Alvey under its Man-Machine Interface Directorate.

GEC and IBM continue to be close collaborators of AIVRU. We are exceedingly appreciative of the medium to long-term perspective they have shown in supporting our research over the past 6 years. In IBM Dr Rodger Hake and Dr John Knapman have been outstanding in their backing. In GEC, we wish to note the first-rate support from Prof Dennis Scotter (Manager of GEC's Long Range Research Laboratory) and Prof Cyril Hilsum (GEC's Director of Research), but most of all the tireless support of Bernard Buxton[4].

We would also like to mention to mention the special role played by Prof Christopher Longuet-Higgins. He agreed early on to act as a consultant for the consortium. This led him to attend all major meetings, during which he injected his usual large number of constructive and insightful remarks. During this period he solved an important problem in structure-from-motion (included here as paper [30]).

Finally, this book would not have emerged without the assistance of Grace Crookes, AIVRU's secretary funded by

[4] Bernard Buxton (of GEC) was seminal figure throughout. Indeed, he and Mayhew recollect making a decision to get the whole project underway in Buxton's kitchen late one night in 1982, though doubt must be cast on the reliability of that recollection given the quantities consumed.

Alvey to assist with coordination of the consortium. We thank her for exceptional competence and patience, both during the Alvey period and subsequently. It has been said that trying to organise academics is rather like trying to herd cats. If anyone is in good position to judge the truth of that remark it is Grace.

6 SUMMING UP

An unusual feature of this book is that each of the subsequent sections starts with the original grant proposal. This is followed by scientific papers recording what emerged. This enables readers to form their own judgements on the success or otherwise of the Consortium in meeting its objectives. (Readers who are not so inclined can of course skip the introductory sections.) We will not therefore attempt an overall review of the scientific content ourselves although we do provide commentaries at the end of each section that highlight certain aspects of the work. Here we simply record some of our general reactions to the Consortium and its work.

***Amazing That It Happened At All**
Given the current European science climate, in which many large industrial/academic consortia exist funded by ESPRIT, it is easy to forget that the consortium way of financing U.K. academics is relatively novel. Certainly for us, Alvey was a 'first go' at participating in, and indeed coordinating, this sort of large-scale scientific activity. The view we have come to is that Alvey was an heroic effort to get the UK's IT effort into the 20th century before it ended, and one that was remarkably successful. The shame is that the remarkable spirit of collaboration which Alvey created, was not seized upon in a suitable follow-up programme aimed at bringing Alvey's research achievements nearer to the market place. Brian Oakley (Alvey's Director) has lamented this failure in a recent review of Alvey (Oakley and Owen, 1990) and we see no reason to disagree with him.

***Collaboration Is Possible**
Alvey introduced us to the benefits of proper travel funding for inter-site visits. It is easy to forget how difficult it was, indeed usually quite out of the question, for U.K. academics to get on the train to visit colleagues in other institutions. The travel money simply was not there (of course, this can still be a difficult problem for academics on standard non-collaborative grants). Alvey changed all that for us, and as a result we had many and vigorous discussions with consortium members, from which we benefited considerably. Amongst other things, these gave us an insight into the problems experienced by large industrial companies in mounting basic research. It is very clear to us that that endeavour cannot be left to them alone: the role of universities in tackling long-term fundamental issues is crucial. This may seem an obvious point to make but in a decade of university cutbacks and government insistence on academics finding industrial backing for their work, it perhaps needs saying very loudly and very often. Nor do we think industry disagrees; certainly our Alvey industrial sponsors (GEC and IBM) concur. In any event, we found the scientific discussions within the consortium extremely helpful and we hope this is visible in the papers in this book. For us, the inter-relationships between the work conducted at the various sites is evidence that the consortium did achieve that elusive goal of the 'whole being more than the sum of its parts'.

***Collaboration is Difficult**
It came as a bit of a surprise to us to realise in the compilation of this book that the consortium spawned not a single inter-site publication. Presumably this was because, despite the many visits and workshops, the underlying reality was still that at the end of the day individual sites knew they would be judged individually in future rounds of grant getting. Another problem was the rather frequent changes of personnel, especially, as it turned out to our surprise, within the industrial partners. Despite these factors, some code was ported between sites. For example, we benefited greatly from IBM's body modeller WINSOM (see [27]), and we circulated to everyone who wanted it a copy of TINA, AIVRU's stereo-based computer vision environment (see [29]). But it hardly needs pointing out that the character of the consortium was never intended to be that of a closely inter-twined multi-site software house bringing forth a large software product. We were not building a 'demonstrator'. Far from it: the consortium was a club of colleagues who agreed to pursue the goal of enhanced understanding of a set of inter-related research issues, their work genuinely enlarged by inter-consortium debates but in the end conducted separately.

***Medium term Pay-offs For Industry**
It has taken a further two years after the end of our Alvey grant to build a device (MARVIN - paper [10]) that can deliver useful 3D scene geometry from stereo at industrially relevant rates. That achievement fits the time frame we set for ourselves at the outset (3-5 years). On the other hand, it would be foolish to pretend that MARVIN is now an 'off-the-shelf answer'. It offers scope for immediate industrial exploitation but it is better regarded as the precursor of a new device that builds on its performance. Such work is now being planned between AIVRU and GEC. Other sites have had their own follow-up programmes.

***The Alvey Vision Conferences**
These became an annual event, beginning in September 1985. They have covered a wide range of topics but with emphasis on the 'image understanding' approach to recovering useful 3D scene descriptions, rather than on the 'pattern recognition' one with its customary attack on 2D problems. The success of these meetings has been such that they are to continue post-Alvey under the auspices of the newly-formed *British Machine Vision Association* (BMVA). This new organisation combines the Alvey vision community with members of the *British Pattern Recognition Association*. That outcome of Alvey is in itself no mean achievement.

REFERENCES

Oakley, B. and Owen, K. (1990) *Alvey: Britain's Strategic Computing Initiative*. MIT Press