Performance Characterisation in Computer Vision (PCCV)

N. Thacker,
Imaging Science and Biomedical Engineering, University of Manchester, UK
email: neil.thacker@man.ac.uk

EU IST Project 1999-14159
Supplimentary 2nd Year Report

This document summarises the presentation made to the commission on 14th April 2003 regarding the progress of the work so far. It also includes our intended work-programme for the remainder of the project as discussed with the external reviewers and project officier at that meeting.

1 Motivation

The motivation for this project comes from several observations regarding published and applied solutions to image processing tasks in the computer vision area. While it is true that image analysis techniques have already proven themselves in many areas of manufacturing (particularly inspection), there are many new application areas emerging due to ongoing changes in the information society. Unfortunately, current research practices and in particular system-building techniques are inadequated to address many of the image interpretation tasks which present themselves. We have identified that one key aspect of this problem is the inability to conduct adequate performance characterisation evaluations of the new technology. In fact such performance characterisation tasks have proven useful in some niche areas, in particular optical character recognition (OCR) and face recognition. These subjects have a sufficiently large research and development community and enough financial drive to have motivated these communities to rationalise the evaluation process. This includes not only the accumulation of sufficiently large data sets but also the agreement on test metrics so that competing approaches can be compared and contrasted. Other areas of computer vision are far more diffuse and do not benefit from the same degree of focus, drive and expertise. The first two of these issues we can do little to remedy. The application of new technology will be adopted only when there is a percieved need. However, we believe that much can be done to try to maximise the potential for adequate performance charaterisation in the future by transferring successful practices from some application areas to other domains and facilitating access to suitable test data sets. Crucially this will also require promotion of the need for performance characterisation in the field, as the percieved need for such work is still seen of marginal value to academics.

2 Partners

The partners involved on this project are, KTH-Stockholm (Henrik Christensen),the University of Essex (Adrian Clarke), and the co-ordinating partner Visual Auomation Ltd.(VAL). VAL is the commercial exploitation company associated with Imaging Science and Biomedical Engineering, University of Mancherster. Originally the project co-ordinator was Patrick Courtney, who left VAL and whos work has been continued by Neil Thacker (a lecturer at the University of Manchester). At the start of the project there was also a steering group comprising leading academics from Italy (Alberto Broggi), Finland (Antti Soini Satakunta) France (Christophe Guizard) and Germany (Siegfried Stiehl), which due to general lack of availability of these high profile academics has been replaced now by an interest group.

3 Aims

The PCCV aims to:

- provide current and future practitioners with the neccesary skills and expertise to facilitate statistical testing, tuning, algorithmic combimation and algorithmic re-use
- improve exploitability of algorithms by improving the understanding of the match between requirements and capabilities of algorithms, including issues of robustness, accuracy and stability
- to extend the accepted application doimain for new scientific and technological developments.

4 Objectives

The objectives are to target current and future practitioners, to encourage and inform RTD-type projects involving image analysis and end users of vision system technology.

5 Outputs

The outputs from the project will be in terms of:

- higher level of skills amongst practitioners
- workshop proceedings describing state of the art
- wider availability and take-up of techniques
- increased awareness in specific sectors
- distribution of reference data
- tutorials and testing services
- standards working groups.

The majority role on these areas will divide up largely as follows, VAL will be responsible for organising tutorials and generating training material and software. KTH will organise workshops and organise publication of proceeedings. The University of Essex will be responsible for all aspects of web-presence.

6 Workplan

The workplan to meet these objective is as follows.

- WP2: tutorials covering training in characterisation techniques.
- WP3: scientific workshops. In an effort to generate a continuing series which provide a forum for researchers to generate cite-able work.

- WP4; web-based dissemination, collating results from the above packages and additional links to specific sectors as they are developed withing the project.
- WP5; sectoral dissemination: to target specific groups including national associations.
- WP6: reference data, a working group on standardisation, and a model for testing services.

The pert diagram as concieved at the outset of the project is shown in Figure 1.

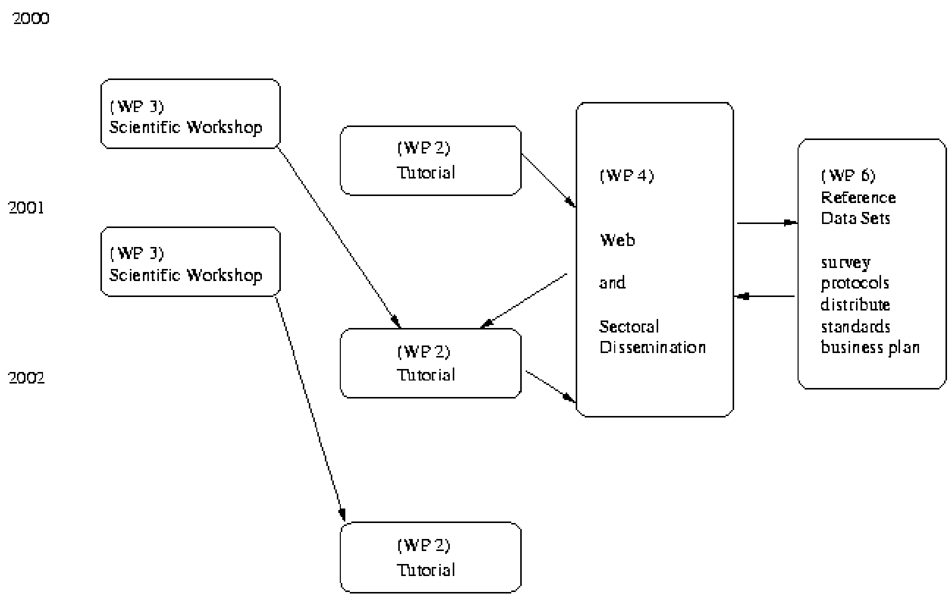

Figure 1: Original Project Pert

Due to various organisational difficulties within the project (including lack of access to funds for the first 9 months and loss of the original project co-ordinator) the actual development of the work has actually proceeded more in line with the interactions shown in Figure 2.

In particular, development of web pages has been delayed and less progress has been made on the development of test data sets due in the main to lack of man-power at the Manchester and Essex sites. This issue has now been addresssed but has required a 1 year extension to the project. Other problems have been beyond our direct control, in particular it proved impossible to convene the orignal steering committee, and though it was always an accepted fact that the steering group might need to be changed during the course of the project, and we have tried to restructure and enlarge this body, ultimately it has proven difficult to elicit susstained committment from highly active but already overstretched individuals. Eventually we took instead the approach of organising a “Vision Engineering Group”. Meetings were organised in July (UK) and August (Ireland) 2002, to promote performance characterisation work as an intrinsic component in vision system engineering. (see minutes distributed at the meeting on the 14th April 2003 in Brussels). We are intending to hold further meeting during the course of 2003.

A key feature of the project plan in both diagrams is the interaction between the research communities and the project through tasks such as workshops and tutorials, and in paticular

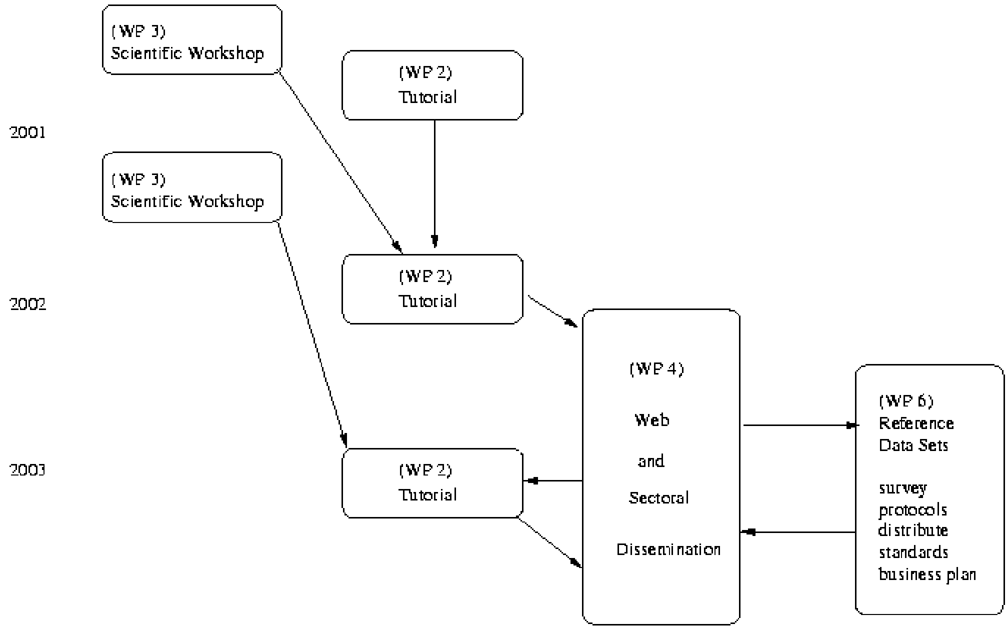

Figure 2: New Project Pert

it has always been the intention to let the community help to focus effort in areas where there is a percieved need. Such interactions will dictate the nature of material ultimately to be made available on the web pages.

7 Workpackages

The following sections discuss in detail goals of each work package. The sections are arranged to reflect the logical order of different project tasks as presented at the review meeting on the 14th April 2003 in Brussels (see Appendix E).

7.1 WP2: Tutorials

The goal here is to provide young researchers / engineers with the neccessary skills to evaluate their own algorithms. This involves dissemination of existing techniques with a particular emphasis on the theoretical relationships between methods. The latter emphasis is necessary because of the sheer range of possible algorithms, which make it rather unlikely that a given evaluation technique can be applied without an understanding of the circumstances in which it is appropriate. The approach taken was to organise half day tutorials at conferences and to provide lecture contributions to summer-schools.

The original intention on the project was to run only 3 tutorials, starting with ECCV 2000 in Dublin. However, late start of the project resulted in the first tutorial being organised and run without PCCV funding. On the other hand, extension of the project beyond the origninal finish date has allowed us to hold far more tutorials than had originally been intended with a further 6 tutorials being held in between September 2000 and July 2002 (Table 1) and a further final tutorial planned for the EPSRC Machine Vision Summerschool in 2003.

Conference	Place and Date	No of attendees
ECCV	Dublin, Ireland;June,2000	10
BMVC	Manchester, UK; Sept, 2001	60+
ICIP	Greece; Oct., 2001	rejected
EPSRC	Surry, UK; June, 2001	40+
EUTIST IMV	Pisa, Italy; Nov.,2001	20
ECCV	Copenhagen, Denmark; May, 2002	10
IV2002	Paris, France; June, 2002	30+
EPSRC	Surry, UK; June, 2002	40+
ICVS	Graz, Austria; 2003	rejected
CVIP	2003	rejected

Table 1: Conferences where the project material was presented

The content of these tutorals has been gradually modified over the period of the project, in line with the original project construction, according to feedback and interaction with specific sectors and the requirements of each conference. In particular, the first tutorials emphasised specific characterisation projects and had only limited theoretical development of methodology. We were able to explain to researchers the difference between scenario evaluation (testing of specific groups of algorithms for a particular task and data set) and technology evaluation (the testing of algorithms for the purpose of characterising performance in a transferable manner so that the algorithm can be used as a module in a larger system). As the tutorials progressed it became clear that there was a need for, and an interest in the statistical methodology underlying algorithm design and testing. Our tutorials have been given a satisfaction rating of 3.6/5 on average, and most of the 210+ people who attended said they would recommend the tutorials to colleagues.

The approach taken in later tutorials can be briefly summarised as making the relationship between algorithm and probability theory explicit, so that assumptions can be tested, and computation of the Cramer/Rao (minimum variance) bound. The reasons for this are that all algorithms need a prediction of performance before they can be used reliably as a module in a larger system, and an algorithm which closely approximates the CR bound in experiments is making optimal use of the data. In presenting such material to the computer vision community we were directly challenging the accepted wisdom that computer vision can be succesful without the need for practitioners to understand statistics. This is potentially a very unpopular shift and one that established workers in the area are unlikely to accept easily. The resulting methodology material has been distributed amongst practioners for comments and feedback.

The tutorial series can be said to have achieved far more than originally planned, with both material describing aspects of good practice and a methodology document being generated by the work. Overall, written and presented materials are not just a prescription of how to apply performance charactersiation to problems already solved but the potential for understanding how to design new evaluation techniques for future applications. This tutoral series was nicely completed by the presentation of a key-note talk at SPIE OPTO-Ireland 2002 which we have included as a supplement to the deliverables (Appendix A) as a good written summary of the content of the early tutorials.

Given the lack of inclusion of ECCV 2000 in this project we request that the deliverables from this workpackage be modified slightly and that the deliverable be replaced with material from the 2002 EPSRC machine vision summer school (Appendix B). The lecture material is rather more restricted than we would have liked due to time limitations (on 1.5 hr lecture) but a copy of a book chapter (Appendix C) was distributed at the lecture in order to cover the bigger picture (summarising largely the material presented in the later tutorials) and is included as part of the deliverable. We request that the tutorial workpackage be accepted as completed.

It now remains to make this material available on the web under WP4.

7.2 WP5: Sectoral Dissemination

The goals of this part of the project were to make practioners in specific sectors aware of the available techniques and to encourage the larger community in the participation of the larger objectives of the project. The intended approach to this task involved a wide variety of components.

- WA5.1 Collect information from steering group on events.
- WA5.2 Establish and maintain contact with new RTD projects.
- WA5.3 Prepare generic reusable PCCV presentation material.
- WA5.4 Collect information on success stories and best practice.
- WA5.5 Make presentations at industry and trade events.
- WA5.6 Promote activities by articles in newsletter and trade press.
- WA5.7 Produce, publish and distribute annual newsletter.
- WA5.8 Present at relevant trade events (IPOT etc.).
- WA5.9 Organise algorithmic competitions.

In fact all of these packages contain to a lesser or greater degree the need to convince others in the community to support our activities. We have therefore had to view this list as a wish list and attend to specific opportunities as they presented themselves. In particular, we helped to organised and participated in a workshop on Medical Imaging (Dagstuhl, Germany) which falls (loosely) into the category of WA5.9. We recieved several invitations to automation days and had a presence at IPOT 2001 in Birmingham (WA5.8 and WA5.5). Contributions were made to the Intelligent Vehicles newsletter (WA5.6) and we provided information to VIM2001 (Nancy, France). A project synopsis and documents have been circulated to contacts and the tutorials were announced to IST project co-ordinators. The tutorial material is now available for inclusion on our web pages (WA5.3). There have also been some press releases.

Under WA5.2 we have made significant progess in our contact with British Aerospace Systems (BA Sys., Bristol). Following great enthusiasm by BA Sys. for the project we have been able to identify possible examplars for the technology evaluation methodology and are now working with them to produce experiments and publications in the area of object location.

Specifically the work has involved evaluation of an existing system and suggestions for an approach to calculate the covariance on estimated parameters so that this can be compared with performance. We hope that more can be reported regarding this project at the final review meeting.

In addition, in the medical research commnity the Integrated Research Consortium MIAS, funded jointly by the UK govermental research bodies of the EPSRC and MRC, has been encouraged to perform an analysis of Voxel Based Morphometry (VBM). This technique is finding large take up particularly in the research areas of Psychiatry and Medicine. Unfortunately, an evaluation of the technique by algorithmic experts on the IRC has largely concluded that the technique is of little scientific merit. A meeting was held in Oxford in March 2003 where presentations on this subject were made from the University of Manchester (N.Thacker), Kings College London (W.Crumm) and University of Oxford (S.Smith). Unfortunately enthusiasm to publish this critique is lacking at the moment, but we are still hopeful that a solid document will be forthcoming. Intellectual honesty regarding the actual capabilities of poor image processing techniques is just as important as success stories if the field is to be accepted as good science.

WA5.1 has proven to be more difficult than envisaged due to the impossibility to run a consistent steeri ng committe (see comments in WP6). WA5.7 was considered unlikely to be viable due to scarcity of material following the departure of Patrick Courtney.

We see this workpackage as ongoing but have definite plans for only one more activity before the end of the project. That is to re-advertise the project to current IST co-ordinators working in the computer vision area in a effort to identify additional test data sets and to locate examples of good practices within EU projects. Otherwise we request that the commision regards this work package as completed.

Once again, it now remains to make this material available on the web under WP4.

7.3 WP3: Workshops

The goals of the workshops were to provide a forum for researchers to provide information, to document state of the art for dissemination and to reward researchers with cite-able publications. This was to be achieved by high quality peer-reviewed workshops at conferences. As with the tutorials, once again the first planned workshop (ECCV 2000 Dublin) happened too early for the use of PCCV funds. However, it did go ahead (alongside the tutorial) and achieved its aims, producing peer reviewed proceedings. Further workshops were then organised at CVPR USA 2001, covering performance characterisation, PETS and STEREO. Other workshops at CVPR were attended by James Ferryman who wrote articles summarising their content for PCCV. A special issue of Image and Vision Computing covering aspects of the work presented at these workshops is in preparation. This workpackage has resulted in a good means for dissemination of good practice across the machine vision community, each workshop had a typical involvement of 15 researchers with 55 participants. Techniques such as the use of Reciever Operator Curves (ROC) has now become mainstream in robotics partly because of these activities. Appendix F gives details on workshops as presented at review meeting in Brussels on 14th of April, 2003. Although other materials, such as books and proceedings are still to appear, we would like the Commission to accept that this workpackage is completed and would like to thank Henrik Christensen for his hard work.

7.4 WP4: Web Dissemination

The aims of this work package were to promote tutorials and workshops, to make available resulting materials and proceedings and to create an online forum for exchange of information. The approach taken was to extend an existing web server (peipa.essex.ac.uk). This work is the direct responsibility of Adrian Clarke.

Specific workpackages were:

- WA4.1 transfer existing web resource page to server/re-organise.
- WA4.2 purchase and install redundant disk drives.
- WA4.3 add access vcounter to log and analyse hits.
- WA4.4 create an online forum
- WA4.5 add workshops and tutorial materials.
- WA4.6 include pages for success stories and best practice.
- WA4.7 promote web site in newsletter and trade press
- WA4.8 make links to national vision clubs.

The first three packages were delivered very soon on the project. These pages have been in place under PIEPA for about two years now. There are roughly 10-50 hits per month but we have found very little other benchmarking activity outside of the PCCV pages. Adrian's ability to devote time to PCCV has been hampered over the last year by other commitments and medical problems (now cured). We intend to make an appointment from the 1st of May until the end of the project to help with the remaining tasks and those under WP6. This will include, devlopment of an introduction page (pro's and con's), revision of tutorial material to make it more web presentable (it currently exists either as chapters or postrscipt transparencies), and the inclusion of detailed case studies that show benchmarking and design methodology as an integral part of vision system development. The majority of raw source material for this will be provided by the University of Manchester.

7.5 WP6: Reference Data testing Services and Standards

The goals for this workpackage were to encourage the development of common vocabulary and proceedures and to establish services for objective testing and comparison.

Specifically the intention was to attempt the following;

- WA6.1 locate instances of re-usable data sets.
- WA6.2 develop protocol for validation and provision of data sets.
- WA6.3 to enter discussions on provision of data sets for standards bodies.
- WA6.4 to follow up users of data sets to collect success stories.
- WA6.5 to re-establish a standards working group in image processing.
- WA6.6 to develop a business plan for testing services.

- WA6.7 to develop a business plan for self-financing tutorials.

This set of workpackages were probably always ambitious. The research community as a whole is not in a position to begin to define standards, either for test data sets or otherwise. The objectives for this work package thus had to be scaled down in line with something more achievable within the man-power and time-scale on the project, and the expertise within the community.

To a large extent what can be achieved within WP6 are now determined by the success of other work packages within the project. For example, WP6.1 is most likely to develop out of the sectoral contacts developed in WA5.2. As such we are working towards such datasets on collaborative projects with British Aerospace Systems and the MIAS consortium. In addition we intend to provide methodology exemplars, complete with test data, for the areas of object location, co-registration of 3D medical images, and stereo vision. Each is the result of a long historical process which has led us to conclude that this would be worth the use of the limited resource within the PCCV project. Manpower has been allocated at the University of Manchester for the remainder of the project and we will demonstrate the results of this work at the final review meeting. Outcomes of the work will be in the form of: (i) theoretical documents, (ii) evaluation reports and (ultimately) (iii) test data sets. Again, detailed completion of last two tasks requires some contribution from people outside of the project which entirely depends on their good will and availability of time and resources at their end. While we can provide application and methodology document for performance characterisation based on years of expirience in computer vision field, some work packages greatly depend on the feedback and contribution from people outside the project and we request that the commision takes this factor into account in the final review of the project. The theoretical evaluation of TALEOS (object location) for British Aerospace Systems is encluded here as an example (Appendix D).

Development of a protocol for validation and provision of datasets (WA6.2) is something which can only come out of a community and cannot be specified (only encouraged) by ourselves. Currently, our main effort to achieve this has centred around our "Vision Engineering" meetings, and the circulation of the methodology document which was generated out of the tutorials. A methodology document was circulated to all attendees prior to the second meeting with one of the intended aims of the meeting to arrive at a consensus regarding its content and further distribution as a recommended document. The attempt was made to endorse a common document which will contain theoretical background for computer vision algorithm testing and performance characterisation according to all members of interest group. Much of the discussion which followed centered around differences in preconceptions regarding focus. General agreement was not reached at this meeting.

It follows that it will be extremly difficult to adopt an unified strategy within the community which has diverse background, training and views. WA6.3 is unrealistic due to the general lack of standards bodies to approach, but we will continue to collate test data sets and make them available to others in the field. It is not known at this time if the original co-ordinator on the project had something specific in mind for this work package when the proposal was written. We feel that WA6.5 is a task which needs to carry on beyond the duration of this project due to difficulties we encountered in organising interest group and unavailability of experts in the field. We think that the best way forward in order to address packages WA6.2 and WA6.3 is to hold further meetings, redistribute written methodology document to wider community and also encourage others to write their own best-practice documents. We will continue to develop our work further and provide test data sets and examples to validate presented techniques. Through this project we hope to rase awareness for need for a common methodology for computer vision systems performance characterisation and this work represents a good basis for the future standardisation in this field.

WA6.4 requires both the existence and identification of test data sets, agreement to distribute them on web pages (WP 4) and finally, other groups outside the project making use of the data. To summarise, we feel that WA6.3, WA6.4 and WA6.5 deliverables are very unlikely to materialise fully within the short lifetime of the PCCV project and in order to address such ambitious proposals the ongoing support from leading experts in Europe is required, as well as adequate resources. Should the oportunity arise for contributions to these work packages we will include it in the project. We hope that our limitations will be understood by the reviewers of our project and the commision, and the requirement for substantive delivery on these aspects of the project be waived.

Although a commercial venture offering unspecifieable testing procedures is a non-starter, the provision of expandable testing services is something that we can attempt as an extension to our existing web activities. Adrian Clarke intends to set up an evaluation server based around his HATE system for performance characterisation. This attends to both WA6.2 and WA6.6, though in a manner probably not intended (or expected) when the proposal was funded. The advantage of this approach is that it can be sustained and extended at a sustainable cost beyond the life of the PCCV project. Manpower at Essex will be used to support this activity over the rest of the project. We request that the commision will take these activities as completion of WA6.2 and WA6.6 at the end of the project.

WA6.7 would probably have made rather more sense if VAL had continued as a viable company after the departure of Patrick Courtney. The company now exists only on paper and the directors have no intention of setting up a tutorial business. Approaches were made to SIRA, which is a company in Britain which specialises in such courses for industrialists, and we were told that tutorials on performance characterisation were probably premature given the current state of the art in industry. As it stands, it is intended that Neil Thacker will continue to organise tutorials in his capacity as independant academic. Clearly, materials generated within the lifetime of the project will have visibility beyond that due to our own web pages. We also intend to make links with CV-Online, which is maintained by Rob Fisher in Edinburgh.

8 Finances

8.1 General Finances (all figures in euro)

	Essex	**KTH**	**VAL**
Travel Budget	6000	5500	8000
Travel Spend	2505.2	0	6787.88
Travel Balance	**3494.8**	**5500**	**1212.12**
Consumables Budget	29000	11000	1656
Consumables Spend	4273.5	0	62.58
Consumables Balance	**24726.5**	**11000**	**1593.42**
Equipment Budget	4500	0	0
Equipment Spend	4500	0	0
Equipment Balance	**0**	**0**	0

8.2 Finances (manpower months)

	Essex	**KTH**	**VAL**
Staff Budget	6 mm	4 mm	15.1
Staff Spend	2 mm	4 mm	5.5
Staff Balance	**4 mm**	**0**	**9.6**
Sub-contracting Budget	0	0	3000 euros
Sub-contracting Spend	0	0	0
Sub-contracting Balance	**0**	**0**	**3000 euros**

A Optoelectronics, Photonics and Imaging 2002 (OPTO 2002)

Keynote Presentation: Using Quantitative Statistics for Construction of Machine Vision Systems

Optical Metrology Imaging/ Machine Vision Conference
5-6 September 2002, Ireland

B EPSRC Summer School 2002 Transparencies

Part 1: Statistics and Error Propagation
Part 2: Image Processing Stability
Part 3: Evaluating Representation
Part 4: Pattern Recognition and Neural Networks

17-21 June 2002, Surrey, UK

C EPSRC Summer School 2002 Hand Out

**Performance Characterisation in Computer Vision:
The Role of Statistics in Testing and Design**

17-21 June 2002, Surrey, UK

D Theoretical Evaluation for BA Systems

Computing Covariances in TALEOS for Quantitative Vision Application

18 July 2002

E Review Meeting Transparencies I

PCCV Project Overview

14 April 2003, Brussels, Belgium

F Review Meeting Transparencies II

PCCV Project Workshops

14 April 2003, Brussels, Belgium