BMVA News

BMVA News[1] is published every three months. Contributions on any activity related to machine vision or pattern recognition are eagerly sought. These could include reports on technical activities such as conferences, workshops or other meetings. Items of timely or topical interest are also particularly welcome; these might include details of funding initiatives, programmatic reports from ongoing projects and standards activities. Items for the next edition should reach the editor by 14th July 1997.

Contents

[1] The British Machine Vision Association and Society for Pattern Recognition is a Company limited by guarantee, No. 2543446, registered in England and Wales. Registered Office: Granta Lodge, 71 Graham Road, Malvern, WR14 2JS. The Association is a non-profit-making body and is registered as charity No. 1002307. See `http://peipa.essex.ac.uk/bmva/` for more information about the BMVA.

Editorial

A recent posting of the "Vision List" saw the announcement by its moderator Phil Kahn that he would introduce a restriction on conference announcements to be no more than 100 lines. This is to be welcomed as it *is* rather tedious wading (or even just skipping) through the seemingly never-ending programmes of some of the larger conferences. But even more mind boggling than these participants lists is the vast number of conferences that are held nowadays. Browsing the "WorldWideWeb Virtual Library on Conferences" I found 815 conferences (covering over 200 subject areas) listed, implying that every day of the year there are likely to be tens of conferences taking place somewhere or other in the world. Of course, there are rather fewer computer vision conferences; the "Computer Vision Home Page" lists 13 conferences for 1997 but this does not include many more related and/or local ones which would bring up the total.

At any rate, I wonder if others are affected by this plethora (excess?) of conferences in the same way that I am. I remember that ten years ago I used to get quite excited when a new call for papers came out. I would scrutinise each one, and consider whether I might submit a paper. (It must be admitted that it helped that the conferences tended to be held in attractive locations, Barbados rather than Birmingham say! Closer to home, this tendency has hindered us in attracting some international conferences to the UK). Now however, there are so many

venues that few people are capable of visiting them all. Just think of the physical, financial, and mental resources and stamina that would be required! The upshot is that nowadays when I see a conference announcement the only emotion it evokes in me is ennui.

"Watch My Lips"

With audio already a standard feature on PCs and image capture set to follow suit, attention is focusing on algorithms and applications which integrate visual and audio information and exploit the information overlap. A recent IEE colloquium on "Integrated Audio Visual Processing For Recognition, Synthesis and Communication" addressed these issues. There were eight speakers, numerous video demonstrations, and a live link to Glasgow over the cellular network.

Recognition

Three speakers addressed the problem of recognition. The approaches differed both in what parameters were used for the visual recognition, and how the audio and visual results were combined. The recognition task also differed, making it impossible to do direct comparisons of the three approaches.

Christian Benoit from INPG/ENSERG in Grenoble used features extracted from either or both the front and profile views of the face. He found that the front features were the most important, and adding the profile parameters gained very little. Combining the audio and visual results in any simple way tended to degrade the high SNR performance - weighting based on the standard deviation of the results overcame this problem.

Michael Brooke from Bath University used a data-driven approach. He quantised the lip area to just 10 by 6 pixels, and applied PCA to reduce this to just 10 parameters. Co-author Mike Tomlinson (DRA Malvern) argued strongly that combination should occur before classification. A joint HMM framework was used for the audio-visual combining, and best results were obtained with a flexible structure which allowed time slippage between the two.

Ian Matthews from the University of East Anglia also used a data-driven approach, but used a "datasieve" on the lip image to derive histograms of the "granules" in the image. This approach had the advantage that exact positioning of the lips in the image was not necessary. The work is in its early stages and issues of optimisation and integration have yet to be tackled thoroughly, but early results are very promising.

Synthesis

Christian Benoit and Michael Brooke also presented work on visual synthesis. Dr Benoit has carried out extensive work on modelling facial features, and has compared the difference on human perception between real and various combinations of synthetic facial features. A grey jay-bone with transparent skin and red lips certainly looks strange, but seems to convey most of the information! He also showed how facial images can be manipulated to make dead (or living) people speak things they never said! Dr Brooke synthesised images from text by using his HMM as data generators - video demonstrations showed this to be quite effective.

Lip-Tracking

Robert Kaucic from Oxford University presented work on a real-time lip tracking algorithm. This is a hard problem because of the low contrast particularly of the lower lip - AV recognition researchers currently use purple lipstick to highlight the lips, an approach which one suspects may not be acceptable to users! By using known shape constraints and a face tracker which helped predict the angle and orientation of the lips, he showed that lips can be tracked whilst the speaker's head moves about. Colour images are not necessary, but colour information can be used to enhance the tracker.

Transmission

Richard Fryer from Strathclyde University presented a video coding algorithm which he demonstrated over an 8kbit/sec cellular link back to his lab at Strathclyde. He could code a 128×128 image at 10 frames per second using a VQ technique which used four different vector sizes (32,16,8 and 4 pixel squares). Instead of minimising the overall SNR, he applies the smaller vectors where they can most improve the local quality.

Christian Benoit also showed his "Labiophone" system where lip parameters are extracted at 25 frames/sec and transmitted, and the face and lips are synthesised at the receiving end.

Other Topics

Huwida Said from University of Reading gave an overview of person recognition techniques. Druti Shar from University of Strathclyde and Farzin Deravi from Swansea University both presented work on recording audio-visual data-bases. The Swansea database, called DAVID, is very comprehensive, and is intended to provide a resource to the research community so that results can be more easily compared.

The day ended in twenty minutes of discussion, and left one feeling that this could be a very important area, but there is still plenty of work to be done to exploit the full potential of audio-visual integration.

Roger Tucker
HP Labs
Bristol
email: `rcft@hplb.hpl.hp.com`

Building 3D models of objects and environments using computer vision

This one-day BMVA meeting was held on the 11th December at the British Institute of Radiology. The meeting was introduced by John Illingworth of Surrey who noted that there had been a recent growth in interest in this area. This was evidenced by the good attendance at the meeting.

Adrian Hilton (Surrey) began with a survey talk on Volumetric vs Surface based representations. Dave Marshall followed with a survey of results from the Copernicus RECCAD project.

Anthony Ashbrook (Edinburgh) told us that since recognizing rigid objects was "too easy", they had started to investigate articulated objects. He presented a curvature based Hough method for recognizing rigid components of objects given a series of views.

It would be a rare BMVA meeting indeed were PCA's not to get a mention and Heap (Leeds) did the honours with work on 3D PCA based models of a human hand. It seemed as if the algorithm required a lot of manual intervention, but the aim was to eliminate this. Kotcheff (Manchester) continued this theme with interesting work on automatic landmark identification, and gave further background to the difficulties of producing a truly automatic means of eigenshape construction.

After lunch Andrew Fitzgibbon from Oxford kicked off the session with an introduction to the VANGUARD project. After failing to decode the acronym he proceeded with his usual no frills style. The objective of the Oxford work is to produce models without knowing anything. The model is initially recovered up to projective transform, and Euclidean structure may be recovered from this by adding additional information. His demonstration was the recovery of a house model from a series of views all the way around. Progress has been made on the 'structure drift' problem which occurs when you try to join up after a complete circuit.

Jones and Oakley presented work on shape from silhouette. They recovered shape from clay teacups and teapots made specially as example objects. If the day job doesn't pan out they obviously have good prospects in the crockery industry. During question time a common putdown is "Did you know that this work was already done by X in year Y?" where $Y \ll 1996$. Pete Mowforth made an entertaining comment of this kind where Y was 1860. Apparently French sculptors had used a very similar manual silhouette based technique to copy sculptures.

The meeting had a welcome international contribution from W. Niem (Hannover) who flew in to present very impressive models built from a sequence of images. A calibration ring is placed around the object on the floor.

The meeting was concluded with two contributions from system vendors. Pete Mowforth gave an extended commercial for PhotoCD and Boots, followed by some of his shape from stereo results. Applications include the Ontario provincial police who seem to have a lot of homicidal maniacs prowling the snow and leaving footprints. After spraying on textured paint accurate shape can be recovered. Bring back Agatha Christie! Continuing his theme of schlock-horror he showed a VRML model of the chairman available to all with a browser. I am beginning to understand the calls for WWW censorship!

On a more visually appealing note Steve Crampton from 3D Scanners has been building models of a Marilyn Monroe bust with a new hand held range scanner. 3D Scanners has recently won a substantial (cash) prize from the EU for this technology.

Andrew Stoddart
Department of Electrical Engineering
University of Surrey
email: `A.Stoddart@ee.surrey.ac.uk`

Royal Society Discussion Meeting

New Geometric Techniques in Computer Vision

Wednesday 23 July and Thursday 24 July 1997,
The Royal Society, 6 Carlton House Terrace, London
further information available from

> Science Promotion Section
> The Royal Society
> 6 Carlton House Terrace
> London SW1Y 5AG
> http://www-sigproc.eng.cam.ac.uk/ jl/

Organised by R. Cipolla, J. Lasenby, H.C. Longuet-Higgins and A. Zisserman

Currently, geometric methods constitute the most useful tools for computer vision. Supplementing Euclidean geometry with projective geometry has led to improved recognition methods and a better understanding of the geometry of multiple views, particularly in the case of uncalibrated cameras. A combination of differential geometry and motion analysis has produced new algorithms for determining surface shape from the temporal evolution of image curves, and a fusion of geometry and statistics has resulted in powerful methods of tracking image features. The meeting will describe recent developments in such techniques, including their implementation in working systems.

Wednesday 23 July

Session 1: Geometry of Surfaces
Prof. J.J. Koenderink: *Pictorial Relief*
Dr P.J. Giblin: *Apparent contours - an outline*
Dr R. Cipolla - *The visual motion of curves and surfaces*

Session 2: Structure and Motion
Prof. O.D. Faugeras: *Grassmann, Cayley and the stratification of 3D vision*
Prof. T. Kanade: *The factorization methods: theory and applications*
Dr R. Hartley: *Projective reconstruction from multiple perspective views*
Dr A. Zisserman: *Metric reconstruction from multiple views*

Thursday 24 July

Session 3: Model-based Recognition (Grouping and Matching)
Prof. J.L. Mundy: *Object recognition based on geometry: progress over three decades*
Dr S. Carlsson: *Combinatorial geometry for visual recognition*
Prof. L. VanGool: *An invariance based strategy for grouping*
Dr J. Lasenby: *Tensors and 3D-invariants for matching over multiple views*

Session 4: Geometry and Statistics
Prof. A. Blake: *Statistical models of visual shape and motion*
Prof. K. Kanatani: *Statistical optimization and geometric inference in Computer Vision*
Dr P. Torr: *Geometric motion segmentation and model selection*
Prof. J.M. Brady: *Summary and closing remarks*

Forthcoming BMVA Technical Meetings

8 May at BIR
"Optimisation Issues in Computer Vision",
organised by Majid Mirmehdi, University of Surrey.
Note: this meeting will take place on THURSDAY

7-8 July at Oxford University
"Medical Image Understanding and Analysis '97",
organised by Prof. Chris Taylor, Manchester University, Prof. Mike Brady and Alison Noble, Oxford University.

8 October at BIR
"People and Animal Watching",
organised by Tim Cootes and Prof. Chris Taylor, Manchester University.

3 December at BIR
"Document Image Processing",
organised by Tienu Tan,
Reading University.

Scottish Chapter Meeting

There will be a meeting ("Vision in Graphics") of the British Machine Vision Association (Scottish Chapter) on Wednesday May 7, 1997 from 1.30pm to

5.00pm. Talks will be held in the Boyd Orr Building (Lecture Theatre A, BO407), University Avenue, Glasgow.

The draft programme is:

Introduction
Dr Paul Siebert, Turing Institute & Glasgow University

Speculation on Virtual Surfaces and Contours in Machine Vision
Dr Adrian Thomas, Sussex University

Using Complex Shading Functions in Shape-From Shading Calculations
Dr John Patterson & Mr Intaek Kim, Glasgow University

Genetic Algorithms for Animating Articulated Figures
Mr Stephane Etienne, Glasgow University

Facial Reconstruction using Principal Components
Prof Mike Burton & Dr Paul Miller, Glasgow University

The Multiple Surface Fusion Problem
Dr Tim Niblett, Turing Institute

A Live Demonstration of 3D Face Capture using the C3D Model-2020
Mr Colin Urquhart, Turing Institute

> Paul Siebert
> The Turing Institute
> Tel: +44 (0) 141 337 6410
> Fax: +44 (0) 141 339 0796

Book Review

Affine Analysis of Image Sequences
by Larry S. Shapiro

is a PhD thesis published by Cambridge University Press as part of the Distinguished Dissertations in Computer Science series. It gives a detailed description of Shapiro's work on interpreting image sequences. In brief, the approach is to track sets of corners, to group them and to extract both structure and motion from the groups. Each step of the process is described in full, along with results of experiments.

The approach to recovering scene structure and motion relies upon the assumption of an affine camera.

Affine structure and affine epipolar geometry are described in detail. Shapiro uses an approach based on Tomasi and Kanade's factorisation method to obtain structure from all available points, and stresses the improvements this gives over choosing four points to define a local reference frame.

The important subject of robustness and outlier rejection is discussed. For a group of points to be exhibiting rigid motion, they must satisfy the constraints of affine epipolar geometry and must lie on a hyperplane in 4D space. Shapiro shows how outliers, points far from the hyperplane, can be detected and thus rejected in an orthogonal regression framework, and gives efficient algorithms for achieving this. He demonstrates how this improves the fit to the remaining data.

There is a chapter on computing the rigid motion of observed objects, again using affine epipolar geometry.

Finally, Shapiro describes affine transfer, where the positions of points only visible in a subset of available images of an object are calculated from those points that are visible in all.

The book is well written, and there are plenty of diagrams and examples to elucidate the equations.

This work will be of interest to those working on tracking and shape reconstruction, and would be a good introduction to some of the issues in that field. The sections on affine geometry and outlier detection in a regression framework may be of interest to a wider readership.

> Tim Cootes
> Manchester University
> email: `bim@sv1.smb.man.ac.uk`

"Machine Vision: Theory, Algorithms, Practicalities"

E.R. Davies

published by: Academic Press, London
2nd edition (October 1996)
paperback, pounds 32-50
ISBN 0-12-206092-X
xxxi + 750 pp.

This book covers the subject of machine/computer vision, including particularly image processing,

image analysis and automated visual inspection. It emphasises the theory underpinning practicalities such as lighting, real-time hardware implementation, and tradeoffs between important parameters, such as speed, implementation cost, accuracy and robustness - as a result giving fair weight to topics such as Hough transforms.

The second edition incorporates many recent advances in the theory and practice of machine vision, including:

- 3-D and motion interpretation
- Invariants
- Camera calibration
- Artificial neural networks
- X-ray inspection and foreign object detection
- Mathematical morphology
- Robust statistics
- An updated and very extensive list of references

For further details, contact the author:

Professor E.R. Davies
Royal Holloway,
University of London
email: `E.R.Davies@rhbnc.ac.uk`

Searchable Database of BMVA Members' Research Interests

The British Machine Vision Association (BMVA) is pleased to announce a database of its members' interests on the World-Wide Web. The database may be searched under a number of headings either by application area or by research interest via the URL

`http://peipa.essex.ac.uk/bmva/bmva-ri.html`

The results of searches may be returned as a table or in a printable format.

The database provides a valuable starting-point for end-users and for industry in making contact with people working in specific areas. It forms an example of the BMVA's role in transferring vision technology

from the research laboratory into industry and the user community.

Inclusion in the database is available to all current BMVA members: a form is available via the above-mentioned Web page.

Adrian Clark
University of Essex
email: `alien@essex.ac.uk`

EPSRC Studentships

Four EPSRC Studentships are available from the Centre for Vision, Speech and Signal Processing Research (CVSSP) at the University of Surrey in Guildford. Principal areas of interest include, but are not restricted to:

- Image Processing
- Image Analysis
- Range Data Analysis
- Industrial Inspection
- Feature and Shape Extraction
- Robotics
- Remote Sensing
- Medical Image Processing
- Fundamental Pattern Recognition Methodology
- Shape Recognition
- Image Coding
- MultiMedia and Digital Broadcasting
- Image Databases

Further information on the group can be found at:

`http://www.ee.surrey.ac.uk/Research/VSSP/`

Initial expressions of interest can be via one of several routes. Get in touch with either:

- Centre Director
  Professor Josef Kittler
  Tel: 01483 259294
  email: `J.Kittler@surrey.ac.uk`
- School Postgraduate Tutor
  Dr John Illingworth
  Tel: 01483 259835
  email: `J.Illingworth@surrey.ac.uk`