

BMVA News

The Newsletter of the British Machine Vision Association and
Society for Pattern Recognition

Volume 21 Number 1
September 2010

Editor: Professor Roy Davies
Department of Physics
Royal Holloway, University of London
Egham, Surrey, TW20 0EX
Tel: +44(0)1784 443429
Fax: +44(0)1784 472794
email: e.r.davies@rhul.ac.uk

<http://www.bmva.org/>

BMVA News¹ 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 10 December 2010.

Contents

Editorial: <i>Brains, Cameras, Vision</i>	1
Prizes and Awards at BMVC 2010.....	2
Around and About at BMVC.....	3
Announcements at BMVC.....	6
Report on BMVA/EPSRC Summer School 2010.....	6
BMVC Quiz.....	7
IAPR Fellowships Awarded to Majid Mirmehdi and Tim Cootes!.....	8
Recent Promotions in the Area of Machine Vision.....	8
Report on CARS 2010.....	9
Reports on CVPR 2010.....	10
Report on ECCV 2010.....	14
Report on Temporal Pattern Learning and Recognition.....	15
Bringing Together a New Generation of Computer Vision Scientists.....	16
Does Your Student Introduce His/Her Talk Correctly?.....	16
Guidelines when Taking Pictures for BMVA News.....	17
David Marr's <i>Vision</i> – Back in Print!.....	17
Last Views of Aberystwyth.....	18

¹The 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.

Editorial: *Brains, Cameras, Vision ...*

There can be relatively few people in the vision community who have not heard of David Marr and his work on edge detection leading to the primal sketch model of the early mammalian vision system. This work emerged during the late 1970s and has a special place in our culture, even though Canny or other edge detectors have taken precedence for practical vision systems. However, within our vision community there will be relatively few people who know of David Marr as the person who in an earlier existence cracked the wiring code of the cerebellum – the motor system of the human brain.²

Long ago (it must have been around 1971) I went to a thrilling talk by Marr in a packed lecture theatre, to hear about these ideas. Initially, I didn't realize that the person I met at a vision conference in 1979 was that same person, talking about "brain" as he put it (neurophysiologists don't call it "the brain" as most of us do). He was full of life and extremely communicative: while much fêted, I found him non-assuming, interested only in new ideas. But a year or so later, in December 1980, in his review paper at ICPR, Rosenfeld mentioned "the [edge detection] work of the, I'm afraid to say, *late* David Marr": these exact words are indelibly stamped on my mind, as they shocked me to the core with their suddenness and with the tragedy of a youthful person taken away in his prime. Within a year, his book – aptly called *Vision* – came out posthumously, and he acquired a new lease of life, at least in our minds and culture.

But Marr was not merely famous in two independent fields, and did not have two separate existences. His

²David Marr (1969) "A theory of cerebellar cortex." *J. Physiol.*, 202:437–470

train of thought was that of seeing the maths underlying mental function, and inferring what the wiring diagrams of brain are really doing – and to that end he developed a new methodology. It must be remembered that in those days biology was rather non-mathematical (in my own school you had to give up maths to do biology in the 6th form). And, when dealing with such a complex machine as the brain, special methods were necessary. Marr had understood all this when developing his ideas on the cerebellum for his PhD thesis, and the theory he produced at that stage had the virtue (not by chance) of being experimentally verifiable. It was not only by citations of his work but by osmosis within the community that he started to revolutionise the way we think about the brain. So the vision side of his work that we vision scientists see is only the tip of the iceberg.

Thus, it is extremely gratifying that Marr's famous formative work, *Vision*, has recently been republished in paperback with a new foreword by Shimon Ullman and a new afterword by Tomaso Poggio, both of whom he knew well and worked with at MIT. If anyone would like to review this book for BMVA News, please get back to me and I will send for a review copy.

On a separate tack, while putting together this issue of BMVA News, I have been rather overwhelmed not only by the number of articles (two reviews of ICPR have had to be put back until the next issue), but also with the number of photographs I have been sent – well in excess of 400. Selecting the winning photographs has been difficult, but it is extremely gratifying that there is this level of activity and response, as pictures breathe vitality into the many lines of text. It is interesting to see that over time, the sizes of the submitted photographs have grown from ~2 to ~10 Mpixels – which is all to the good, as it means that the printed pictures have the possibility of being excellent quality. However, on the basis that pictures need to be printed at 300–400 dpi, there is little need for much more than 2 Mpixels, or at most 4 Mpixels if I trim them, which I frequently do. Advancing technology and the advertising that accompanies it do not also bring the health warning that all these pixels are of little benefit, as their effects can't be seen without a magnifying glass. For my own part, I have a 7-Mpixel camera, which I bought for its stabilised image capabilities rather than for the number of pixels it possesses. Indeed, I normally use it on the 5-Mpixel setting, which is valuable for saving storage and improving loading time. I guess it's much the same thing as driving a Maserati and going easy on the accelerator pedal ...

Professor Roy Davies
Editor, BMVA News
email: e.r.davies@rhul.ac.uk

Prizes and Awards at BMVC 2010

Best Science Paper Prize

- Ľubor Ladický, Paul Sturgess, Christopher Russell, Sunando Sengupta, Yalin Bastanlar, William Clocksin, Philip Torr – *Joint Optimisation for Object Class Segmentation and Dense Stereo Reconstruction* (Sponsor: BMVA, £750)

Best Industrial Paper Prize

- Alonso Patron, Marcin Marszalek, Andrew Zisserman, Ian Reid – *High Five: Recognising human interactions in TV shows* (Sponsor: Computer Recognition Systems, £750)

Best Student Paper Prize³

- David Tsai, Matthew Flagg, James Rehg – *Motion Coherent Tracking with Multi-label MRF optimization* (Sponsors: EPSRC, £250 and BMVC 2010, £250)

Best Supplementary Material Prize

- Loren Schwarz, Diana Mateus, Victor Castaneda, Nassir Navab – *Manifold Learning for ToF-based Human Body Tracking and Activity Recognition* (Sponsor: RIViC, £250)

Best Poster Prize

- Parthipan Siva, Tao Xiang – *Action Detection in Crowd* (Sponsor: RIViC, £250)

Sullivan Thesis Prize³

- Oliver Oechsle, University of Essex – *Towards the Automatic Construction of Machine Vision Systems using Genetic Programming* (Supervisor: Adrian Clark)

The thesis is available from:

<http://vase.essex.ac.uk/Oechsle-thesis.pdf>
and the BMVA thesis archive.

BMVA Distinguished Fellow 2010

- Professor Bob Fisher (University of Edinburgh)

Dr Fred Labrosse
Aberystwyth University
email: ffl@aber.ac.uk

³Nobody was present at BMVC to take the prize.

Around and About at BMVC



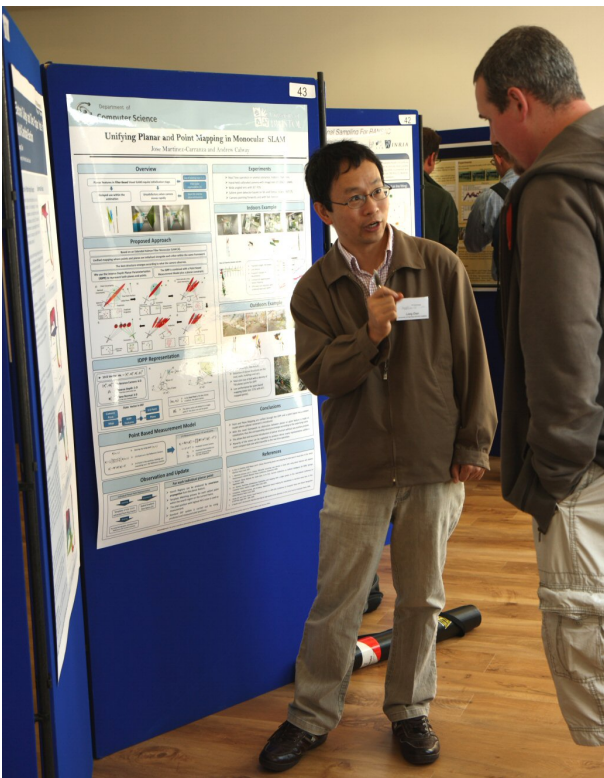
Finding room in the lecture theatre



Activity at the posters



Delegates discuss a poster in the airy presentation area



Liang Chen presenting his poster



Sharing a moment with the cameraman during a tea break:
Paul Rosin and Bob Fisher



Open-air Reception before the Banquet



Phil Torr, Martial Hebert, Joanna Isabelle Olszewska, and Jean Ponce at the Reception



Kristen Grauman, Jean Ponce, Bob Fisher and Martial Hebert at the Banquet



Everyone had a good time!



Music at the Banquet



At the banquet: Phil Torr, Bill Triggs, Ľubor Ladický



Ľubor Ladický receives the Best Science Paper Prize from Fred Labrosse



Alonso Patron receives the Best Industrial Paper Prize



Parthipan Siva receives the Best Poster Prize



Loren Schwartz receives the Best Supplementary Material Prize



Bob Fisher receives the BMVA Distinguished Fellow 2010 Award from Andrew Fitzgibbon



The Venue for the next BMVA was announced as Dundee. Manuel Trucco, next Conference Chair, used the map to prove that Dundee is at the very centre of the British Isles (once enough islands are taken into account)!

The gallery of pictures presented above was taken at BMVC 2010 as the conference unfolded. Its purpose is to act as a reminder for those who were present, as news for people who couldn't be there, and above all as a vote of thanks to the organisers at Aberystwyth University for a highly successful and well-run conference.

Finally, I would like to acknowledge Sandy Spence and Harry Strange who acted as photographers and thus helped me to bring together these views of the life of the conference.

Professor Roy Davies
Editor, BMVA News
email: e.r.davies@rhul.ac.uk

Announcements at BMVC

Results of the BMVA elections

The results of an election to fill 5 vacant places on the Executive Committee of the Association were announced:

- Dr T. Breckon
- Dr A. Clark
- Dr J. Graham
- Dr N. A. Thacker
- Dr I. Reid

In addition, Professor M. Chantler takes an ex-officio place in his capacity as treasurer to the association.

The following were elected as members of the Executive Committee during September 2009, and remain in office for a further year:

- Professor E. R. Davies
- Dr M. Everingham
- Dr A. Fitzgibbon
- Dr D. Makris
- Professor M. Mirmehdi

Location of BMVC 2011

The next BMVC will be held in Dundee and will be organised by Manuel Trucco and his colleagues.

Dr Neil Thacker
BMVA Secretary
email: neil.thacker@manchester.ac.uk

Report on BMVA/EPSRC Summer School 2010

The BMVA/EPSRC Summer School on Computer Vision was hosted by Kingston University on 12–16 July 2010. While the main objective of the summer school is to serve the Computer Vision PhD students from British Universities, its reputation seems to grow in industry and internationally, as reflected by the affiliations of the 36 attendants. Students came not only from UK universities and companies, but also from France, Spain, Portugal, Italy and Malta. Special mention is required for the Maltese minority, the largest of the Summer School!



Summer School attendees



Poster session



Students reflecting on the Summer School material during a break



Professor Maria Petrou presenting her lecture on colour



First prize of the Summer School poster competition was awarded to Alexandros Moutzouris ...



... while the second prize went to Siddharth Hegde ...



... and the third prize to John Bartholomew.

As part of the continuous effort of BMVA to maintain an up-to-date high quality programme, a lecture on “Local feature descriptors” by Krystian Mikolajczyk was added in this year’s programme.

Students were enthusiastic about the Summer School experience, as expressed in their feedback: “I think the variety of the topics is very good”, “Most of the lectures were excellent and the quality and presentation was more than what was expected from this course”, “Diverse set of requirements were handled very well. I found it extremely helpful and fun!”

Dimitrios Makris
Kingston University
email: d.makris@kingston.ac.uk

BMVC Quiz

What place in the UK is famous for the following?

- D’Arcy Thompson (who wrote the classic volume *On Growth and Form*)
- 19th Century jute making
- Dennis the Menace (no known relevance to computer vision).

For the answer, and originator of the quiz, see the bottom of page 5.

IAPR Fellowships Awarded to Majid Mirmehdi and Tim Cootes!

The BMVA has been successful in having two more of its members selected as Fellows of the *International Association of Pattern Recognition* (IAPR) in recognition of their ‘scientific/engineering excellence’ and ‘well-known international achievements in the field of pattern recognition’. They have long associations with the BMVA, having organised many conferences, meetings and publications, and both have been on the Executive Committee. The awards took place at the recent ICPR 2010 conference in Istanbul.



Majid Mirmehdi and Tim Cootes relaxing after receiving their IAPR awards at ICPR in Istanbul.

The BMVA would like to present its own congratulations to Professors Mirmehdi and Cootes on having their work recognised internationally with this prestigious award.

For interest, the respective citations to their work are:

- “For contributions to Image Understanding and Computer Vision, and services to IAPR.”
- “For contributions to the development of statistical models of shape and appearance.”

Professor Roy Davies
 Editor, BMVA News
 email: e.r.davies@rhul.ac.uk

Recent Promotions in the Area of Machine Vision

Andrea Cavallaro has been promoted to Professor of Multimedia Signal Processing at Queen Mary University of London. His research is on target tracking, multimodal content analysis and behaviour recognition in multi-sensor networks. He received the PhD in Electrical Engineering from the Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 2002 and the Laurea (Summa cum Laude) in Electrical Engineering from the University of Trieste, Italy, in 1996. He was awarded a Research Fellowship with British Telecommunications in 2004, a Royal Academy of Engineering teaching Prize in 2007; three student paper awards (on target tracking and perceptually sensitive coding) at IEEE ICASSP in 2005, 2007 and 2009; and the best paper award at IEEE AVSS 2009 for “Trajectory association and fusion across partially overlapping cameras”. He published a book on *Video Tracking: Theory and Practice* (co-authored with E. Maggio, Vicon – Wiley, 2010) and a book on *Multi-Camera Networks: Principles and Applications* (co-edited with H. Aghajan, Stanford University – Elsevier, 2009). Professor Cavallaro is Associate Editor for the IEEE Signal Processing Magazine, the IEEE Transactions on Multimedia and the IEEE Transactions on Signal Processing.



Andrea Cavallaro



David Marshall

David Marshall has been promoted to Professor. He has been working in the field of computer vision since 1986. David obtained a BSc degree in mathematics from University College, Cardiff in 1986 and then undertook a PhD in 3D inspection of manufactured objects at Cardiff University, working in collaboration with the Sowerby Research Centre, British Aerospace. In 1989 he joined Cardiff University as lecturer and is now Professor of Computer Vision in the School of Computer Science and Informatics. David’s research interests include articulated modelling of human faces,

models of human motion, statistical modelling, high dimensional subspace analysis, audio/video image processing, and data/sensor fusion. He has published over 130 papers and one book in these research areas and has attracted over £4M in research funding over his academic career. David enjoys close collaboration with national and international external bodies (including British Aerospace, EADS, General Dynamics, QinetiQ, Museum of Natural Sciences, Madrid, Royal Botanic Garden Edinburgh, DSTO Australia, Free University Brussels) and has good inter-disciplinary research links within the University (including the Schools of Psychology, Engineering, Dentistry and Optometry). He is currently director of the Human Factors Technology Centre – an interdisciplinary research centre between the Schools of Engineering, Psychology and Computer Science – and director of teaching and postgraduate student research in the School of Computer Science and Informatics.



Dimitrios Makris



Sergio Velastin

Dr Dimitrios Makris has been promoted to Reader at Kingston University, taking up the post on 1 August. His research interests are mainly in visual surveillance, articulated pose recovery and action recognition. His work on networks of disjoint cameras and semantic models of CCTV scenes opened new directions of research that are followed by a number of researchers internationally. He has been contributing to the BMVA in numerous ways, including being a member of the BMVA Executive Committee, organiser of the BMVA Summer School, and organiser of the BMVA one-day technical meetings. He is also a coordinator of the V&L NET, EPSRC Network on Vision and Language.

Sergio A. Velastin has been promoted to Professor of Applied Computer Vision at Kingston University where he also directs the Digital Imaging Research Centre. Sergio has been working for a long time on image processing and computer vision applied to the monitoring of people and vehicles (what later has become known as Visual Surveillance). In particular he

has led research in trying to detect situations of potential danger in crowded environments, especially in public transport networks such as underground trains. He is active in the IET and the IEEE where is currently on the Board of Governors of its Intelligent Transportation Systems Society. His current interests include human action recognition, people tracking in crowded conditions, vehicle tracking and classification, and distributed visual surveillance systems.

Report on CARS 2010

The 24th international congress and exhibition of Computer Assisted Radiology and Surgery 2010 (CARS 2010) took place in Geneva, Switzerland on 23–26 June. CARS aims to provide an environment to link diagnostic imaging and image guided surgery and to provide an international focus to encourage innovative research and the development of these activities. CARS covers several different areas of information technology in radiology such as medical imaging, image processing and visualisation, and image and model guided radiotherapy; it also addresses surgery for clinical applications such as surgical robotics and instrumentation, and surgical navigation and simulation. CARS 2010 included a highly professional programme for the participants as there was a very positive response to the call for papers with approximately 600 submissions from 41 countries.



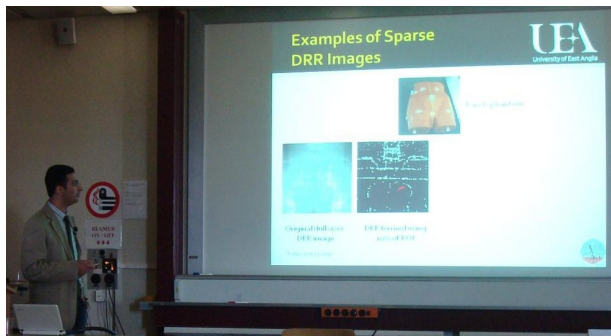
View of the well-known fountain in Geneva

The CARS programme was impressive as it contained different numbers of lecture sessions, tutorials, talks, discussions and a large number of posters all running in parallel. The first day, 23 June, was CARS clinical day with many interesting talks on minimally invasive spinal surgery (MISS), interventional radiology (IR), and interventional oncology. J. Egger from Siemens

Healthcare gave an attractive talk as he proposed a method for automatically calculating and visualising the sealing zones in preoperative computed tomography angiography (CTA) applied to specific diseases like abdominal aortic aneurysms (AAA). At 17:00 the opening session started with an exceptional welcome address from Professor Heinz U. Lemke with a selected theme for CARS 2010 of ‘Outreach’ as it was the 25th anniversary of CARS. This theme highlights the growth and expansion of the scope of CARS meetings since 1985 and the future direction in healthcare, and various approaches were suggested on how this would be achieved.

The second day of the conference moved to the area of computer assisted radiology with three main topics of computed tomography and magnetic resonance (CT-MR), digital angiographic imaging (Angio), and digital radiography. During the day a team of researchers showed a way of accelerating the 3D algebraic CT reconstruction with motion compensation using CUDA-based acceleration, which I found particularly interesting. In another paper, a ray-based approach was applied for boundary estimation of fibre bundles derived from diffusion tensor imaging, and combined with a fibre tracking algorithm allowed estimation of the position and course of fibre tracts in the human brain.

On 25 June the invited speaker E. B. Hug gave a talk on spot scanning based proton radiotherapy, as one of the leading technologies of proton applications. This technology has pioneered at the Paul Scherrer Institute (PSI) and has been in clinical use since 1996.



Osama Dorgham giving his talk

I presented my paper entitled “Fast 2D/3D image registration using accelerated generation of sparsely rendered digitally reconstructed radiographs” in the session on computer assisted radiation therapy (CRAT). This presented our method for speeding up the 2D/3D registration by accelerating the rendering process of DRR images. We address this by only rendering fragments of the DRR images based on regions of interest (ROI) selected automatically using an entropy measure. We also take advantage of the parallel

processing available on modern PCs by mapping the 2D/3D registration of each ROI fragment onto multiple CPU cores.

The last day of CARS finished with different topics on ophthalmology, simulation methods (SM), new approaches to diagnosis and therapy, and image processing and display.

I would like to express my appreciation to the BMVA for giving me the opportunity to attend this conference by awarding me a generous grant. Many thanks, BMVA!

Osama Dorgham
University of East Anglia
email: o.dorgham@uea.ac.uk

Reports on CVPR 2010

The 23rd IEEE conference on Computer Vision and Pattern Recognition (CVPR), held this year in San Francisco, set a new record for this conference with over 2000 registered attendees. The conference was energetic right from the start with workshops on the first day having to move to larger rooms due to unexpectedly large audiences.

There was an extraordinary array of fascinating work on display throughout the conference – far too much to give a meaningful overview in this article – so I will just recount a few personal highlights. Early in the week, Professor David Forsyth gave a talk flirtatiously entitled “Mine is Bigger Than Yours: Big Datasets in Computer Vision”, where he discussed his thoughts on the modern trend in computer vision towards bigger and bigger datasets. His talk was not an objection to collecting and using big datasets, but rather a plea for a more thoughtful, honest, and reasoned approach to data collection. A comment mid-way through his talk summarised his position well: “All good datasets are big, but not all big datasets are good”.

Another highlight was “The Role of Features, Algorithms and Data in Visual Recognition”, by Devi Parikh and Lawrence Zitnick, who attempted to discern exactly where humans out-perform artificial systems in typical computer vision problems. Taking the popular problem of place categorization, in which a system is trained to distinguish between pictures of, say, bathrooms, kitchens, and bedrooms, the authors identified three components present in any classification system: labelled example images to learn from (data), features computed from those images, and a learning algorithm to relate the features to the variable of interest.

The authors then replaced the components in the artificial system one by one with a human (using the popular Mechanical Turk system) and measured the resultant performance increase. The authors drew many conclusions, chief among them being that artificial features are a clear weak point in such classification systems, whereas learning algorithms appear to perform near or above human level – at least for this task.



Alexander Flint enjoying the scenery in Yosemite National Park

This year the conference organisers introduced 90-second “spotlight” presentations for all posters, which took place following the 20-min. orals given in each session. Some of the posters had such long titles that it hardly seemed possible to announce the speaker in the allotted minute-and-a-half, so it was interesting to watch authors work within this new presentation format. Some authors maximized their word count by speaking at breakneck speed, while others gave a tantalizing introduction, inviting the audience to “find out more” by visiting their poster later in the day. One group even performed a 90-second role-play!

Till Quack and Jan Solem gave an inspirational tutorial on computer vision startups in which they discussed their respective experiences in this area. Covering everything from industrial-strength software engineering to marketing and fund-raising, their presentation highlighted the incredible potential for computer vision in the private sector. Drawing on a wealth of personal experience, Till and Jan gave many insightful, and often unconventional, pieces of advice.

Between attempting to absorb as much as possible of the incredible work at the conference, I also had the chance to present my own work at the poster session on shape representation and matching. It was energizing to discuss common interests with colleagues and new acquaintances, and many passers-by had interesting ideas to share.

Following the conference I visited the breath-taking Yosemite national park, which is just hours from San Francisco. Seeing at first hand the spectacular granite cliffs and impossibly high waterfalls made for a great end to a wonderful trip.

Alexander Flint
Oxford University
email: alexf@robots.ox.ac.uk

The Annual IEEE Conference on Computer Vision and Pattern Recognition (CVPR) was this year held in the scenic setting of San Francisco. The conference was held over 6 days, 13–18 June, with workshops and tutorials taking place on 13, 14 and 18 June and the main conference over 15–17 June. It was the largest ever CVPR conference with over 1800 registered attendees. The conference took place in the Hyatt Regency hotel in the financial district of San Francisco. With so many attendees and a surreal staff strike taking place, the conference sometimes felt slightly chaotic. However, overall it ran slickly and the Hyatt Regency made a good venue.



Financial district of San Francisco

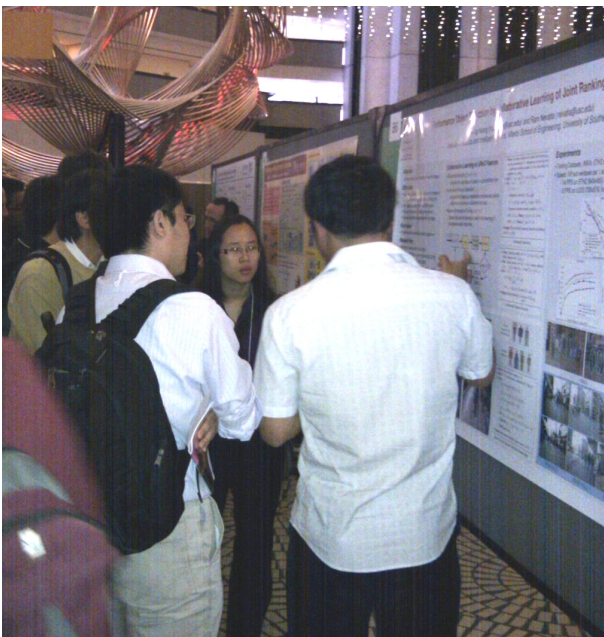
CVPR covers a huge range of topics from event recognition to computational photography. It is a highly ranked conference with a low acceptance rate, so a high calibre of work was on show. I attended the workshop on *Use of Context in Video Processing (UCVP)* on the first day of the conference. It consisted of presentations of five accepted papers, two CVPR preview talks and two invited speakers. Both invited speakers made excellent presentations and of special interest to me was the talk by Fei-Fei Li as I had read much of her work. She made the intriguing point that she believed there was a gap in the research of complex action recognition. At the end of the workshop, there was a group discussion on the subject of context in computer vision. This was a stimulating debate where the value and concept of context was discussed.



Jonathan Letham presenting his paper

My paper in the UCVP workshop was entitled "Contextual smoothing of image segmentation". In this I describe a method of improving region classification of frames in a video by using the context of the scene. The presentation went well and I had a couple of questions at the end. Later, two people approached me with further questions and also suggestions to improve my work.

In another workshop, I attended an entertaining and very interesting talk by Karl Grammer at the Socially Intelligent Surveillance and Monitoring session in the morning. Also that day I saw two good keynote talks given Antonio Torralba and Daphne Koller.



At the poster session

On the Tuesday, the main conference began. Two parallel sessions were run in the two main presentation rooms over the three days. After the oral presentations, the poster spotlights took place. The poster sessions took place at the end of each day over a buffet dinner.

Additionally, the poster session at the end of each day was a good opportunity to speak to the authors of all the papers, though these sessions were sometimes crowded. One paper worthy of note was presented in an eye-catching poster spotlight presentation. Shai Bagon presented work that detected a mutually common and non-trivial shape from several vastly varying images and then sketched a binary map of the shape in *Detecting and Sketching the Common*. This was a novel, interesting idea and he gave a good explanation at his poster.

At the conference I also worked as a student volunteer. In doing this I met many other students attending the conference and was a good opportunity hear their experiences and learn about their work. At the end of the week a dinner was put on for the volunteers and the organisers of CVPR so at this I had the chance to meet some of the leaders of the field.



At the UCVP workshop

This was the first conference I had ever attended so the experience in itself was invaluable to me. At the UCVP workshop I presented to experts in the field and also gained valuable feedback on my work. From attending the talks and poster session, I saw cutting-edge work and learnt some stimulating concepts. Additionally, from attending so many talks which varied in standard of presentation, I learnt some important presentation skills. Overall, the CVPR 2010 conference was definitely a worthwhile experience.

Finally, I would like to thank the Royal Academy of Engineering and the BMVA for their generous contribution towards my trip.

Jonathan Letham
Heriot-Watt University
email: jal13@hw.ac.uk

Someone allegedly once said, “The coldest winter I ever spent was a summer in San Francisco.” The fact that the organizers of **The 23rd IEEE Conference on Computer Vision and Pattern Recognition (CVPR)** gave a fleece or a sweatshirt as part of conference goodies didn’t help the city’s reputation (weather-wise). However, access to food from all over the world, beautiful scenery, world-class wineries, more than made up for the strange micro-climates in the bay area. The main conference, tutorials and workshops were held at Hyatt Regency in the Financial District on 13–18 June. Given that the conference overlapped with the FIFA World Cup, we were never far from the action. A huge projector screen and a television in the hotel foyer allowed attendees to catch up with all the action during the breaks.

On the first day I mainly attended the *Use of Context in Video Processing* workshop. One of the highlights of this workshop was the talk by Fei-Fei Li. She demonstrated the importance of context in visual recognition tasks, such as recognizing human activities. Her recent work in this area with B. Yao, which was presented a couple of days later, won the best paper honourable mention award. With many interesting sessions in parallel, I was spoilt for choice on the second day of workshops. Some of the highlights of the day were workshop keynote talks given by Dan Huttenlocher, Daphne Koller and Antonio Torralba. Professor Torralba made an interesting comparison of the money spent for collecting data in computer vision ($\sim \$10^2$ – Amazon Mechanical Turk) v. other fields ($\sim \$10^{10}$ – Large Hadron Collider), and suggested we spend more on it!

One of the first oral sessions – *Object Recognition I: Context* – comprised of various works, which showed how context can be used to discover unknown objects (Lee and Grauman), detect parts (Karlinsky *et al.*), recognize human-object activities (Yao and Fei-Fei). The Computational Photography session contained interesting works, which could potentially be found in high-end cameras in the future. I presented my poster spotlight in the first session after lunch. I received positive feedback on my poster, which proposed a novel method to learn parameters of random field models, from many experts in the field. Some of the attendees from Google were quite keen in our computational efficiency and promising results. With all the excitement of speaking to a constant stream of audience, it was quite difficult to take a few minutes off to get some food, which was served during the poster session.

One of the orals on the second day of the conference addressed the problem of analyzing structural changes of buildings over time in the context of city-scale 3D reconstruction. This work by Schindler and Dellaert

showed some interesting results on a collection of 490 images of Atlanta dating from the 1930s to the 2000s. The PAMI-TC meeting was held during the lunch break. The committee decided to have a one month deadline for conference bids, and disallow any floor bids in the future. Portland, Oregon was chosen as the host city for CVPR 2013. This was the only bid put together by the PAMI-TC Conference Committee. The 2013 team includes Gerard Medioni and Ramin Zabih as general chairs, and Bill Freeman, Greg Hager, Martial Hebert, and Richard Szeliski as program chairs. The TC also strongly supported the poster spotlight sessions introduced this year. One of my favourite moments of the day was in one of the oral sessions. None of the presenters of a paper could make it to the conference, and so one of the authors sent a 15 min. video which was played during his slot. It made the session chair’s job of time-keeping very easy indeed! The day ended with live entertainment by so3music (see <http://so3music.com>) and special guests.

Some of the highlights of the last two days include works by Felzenszwalb and Veksler on scene labelling, Teboul *et al.* on segmenting building facades, a real-time multi-view stereo reconstruction demo by Vogiatzis and Hernandez, and a higher order models in computer vision tutorial presented by Carsten Rother and Sebastian Nowozin.

Overall, it was a great experience to learn and see some of the state-of-the-art work in our field. I also thought having one long poster session on each day was successful, as it provided ample time for discussions, rather than having to rush to an oral session. It appears that the successful poster spotlight sessions will be part of the programme in future conferences as well (e.g. ECCV 2010). More details on CVPR 2010 and the tutorial slides are available at the following website: <http://cvl.umiacs.umd.edu/conferences/cvpr2010>.

I would like to thank BMVA and PASCAL2 for their generous support, which made it possible for me to attend this conference.

Karteek Alahari
Oxford Brookes University
email: karteek.alahari@brookes.ac.uk

Report on ECCV 2010

The 11th European Conference on Computer Vision (ECCV 2010) was hosted by the Foundation for Research and Technology – Hellas (FORTH), Crete, Greece on 5–11 September 2010. The picturesque Creta Maris Hotel played host to this 7-day Computer Vision event.



A view of the Creta Maris Hotel

Prior to the main conference, there was a day of tutorials. There were 8 sessions each hosted by some of the most prominent names in the field of Computer Vision. Workshops on several areas of computer vision were held on the last two days of the conference.

The main conference ran for four days. The ECCV conference received 1174 submissions and the acceptance rate was 27%. Each day had two oral sessions and two spotlight and poster sessions. Each oral session consisted of 20-min talks. Each poster session had a duration of two hours and was proceeded by a one-min introduction to each poster. Various sessions were categorised to focus on the following research areas: Computational Imaging; Faces and Humans; Geometry; Image Features, Matching and Motion; Object and Scene Recognition; Segmentation; and Visual Learning. Each day also had 25 demonstrations on applications of computer vision techniques to various real world problems. Olivier Faugeras (INRIA) gave his invited talk on: “The geometry of low-level vision: mathematics, biological and machine vision”.

The paper on “Visual recognition with humans in the loop” particularly caught my attention. The paper could form the basis of significant future research in the area of Object and Scene Recognition. The paper talked about incorporating human feedback along with computer vision to solve difficult recognition tasks. The paper focussed on identifying various species of birds, aided by human feedback. The authors show the effect human feedback can have in improving the performance

of vision systems. Although, their work is in a nascent stage, the authors foresee a lot of emphasis on this avenue of research in the coming decade.

The conference gala dinner served as the stage for the various award presentations. It was a night to sample some local delicacies and enjoy local Cretan dance and music. Ľubor Ladický et al. bagged the best paper prize for their paper on “Graph cut based inference with co-occurrence statistics”. The paper titled “Ambrosio-Tortorelli segmentation of stochastic images” won the award for the best student paper. The Koenderink Prize, awarded for ECCV papers from 10 years ago that have had high impact in the field, was shared by two papers: “Stochastic tracking of 3D human figures using 2D image motion” and “Unsupervised learning of models for recognition”. Additionally, there were 10 awards for exceptional reviewers and the efforts of all the organisers were acknowledged. Finally, it was announced that ECCV 2014 will be held in Zurich, Switzerland.



Local Cretan Dance Performance at the Gala Dinner

Two special sessions were highlights of the conference:

- The session on “Vision and Industry”. This session aimed to bring in closer contact between academic researchers and leading industrial labs and firms that translate computer vision into profitable businesses. There were talks from managers of IBM TJ Watson Center, GE Global Research Center, Siemens Corporation, Lockheed Martin, Adobe Systems, Inc and DxO.
- The session on “Research Funding for Vision”. This session discussed the research funding strategies followed by major funding agencies, namely NSF and the EU Research Programme. Managers from both agencies gave brief talks about their respective programmes, budgets and deadlines for applications. At the end of the session there was a panel discussion.

In summary, the conference strengthened ECCV’s position as one of the top conferences in the field. The

baton has now passed to Florence, Italy for the next edition in 2012.

Finally, on a personal note, I would like to thank the BMVA for the travel award and in general for this initiative of awarding PhD students an opportunity to travel and present their research.

Ankur Patel
University of York
email: ankur@cs.york.ac.uk

Report on Temporal Pattern Learning and Recognition

This one-day meeting was held on 22 September at the BCS in London and was chaired by Andrew Gilbert and Eng Jon Ong, both from the University of Surrey. The meeting concentrated on the field of learning and recognising temporal based patterns. This is a popular field and the talks showed the large variation of approaches and applications, ranging from tennis action detection using bag of word based approaches to analysing the approach bridge design engineers used to solve a problem through hierarchical HMMs.

The first speaker Dumebi Okwechime (University of Surrey) described the analysis of human conversation through observing social signals such as nodding, laughing, talking and the activity of the body. He used *a priori* data mining to mine the social signal instances to derive frequently reoccurring rules. With this he was able to predict how a person would react to another person's action and also animate an avatar. Next was Ioannis Kaloskampis (University of Cardiff), he was part of a multidisciplinary project involving Artificial Intelligence, Design Engineering and Cognitive Psychology. The approach monitored and tracked a bridge design engineer action in a paper and pen design setup. He extracted primitive actions through qualitative spatio-temporal relations between the objects including their motion trajectories. Then these activities are analysed for errors with the aid of a hierarchical HMM. The final talk before lunch was given by Michal Lewandowski (Kingston University); he applied Temporal Laplacian Eigenmaps to view dependant videos, and combined these view dependant manifolds to form a unified representation that models each action independent of style and viewpoint.

After lunch, Teo de Campos and Mark Barnard (both University of Surrey) compared the use of either space time shapes (STS) and bag of words (BoW) for action

recognition for a number of datasets including their tennis dataset, and three others. They argued that, for a number of datasets, the use of STS formed from a single feature vector based on HOG3D descriptor for the foreground object was more effective at describing the action than the more popular BoW approach. Following this, Sudeep Sundaram (University of Bristol) presented a talk on visual recognition of actions and events using a wearable camera stabilised using an affine transformation. Subsequently, the image was divided into spatio-temporal blocks of HOGs which are concatenated in time into action cells. Actions were then represented as graphs of action cells which are matched using MRFs. He also reported that performance improved when knowledge of the user's location was used to bias potential actions. Next, Parthipan Siva (Queen Mary, University of London) talked about action recognition in crowds. His work was unique in that he was attempting to recognise actions from surveillance data. This is challenging since the subject of interest may not be in the foreground. This was achieved by starting with a single example of an action represented by SIFT trajectories provided by the user. Additional examples were automatically extracted from the remaining training data and SVM classifiers were trained to recognise the corresponding action. The results presented were promising considering the challenging nature of the data.

Finally, Yuxuan Lan (University of East Anglia) presented work on improving visual features for lip reading, which is also highly temporal in nature. In particular, she presented a new feature called HiLDA that was a concatenation of a set of past, present and future frames before being dimensionally reduced using LDA. She provided experimental results showing that the HiLDA features lead to better discrimination between different visemes used in lip reading in comparison with traditional features such as AAMs. Thus, it was found that HMMs trained using HiLDA resulted in the best person-independent lip reading performance.

In conclusion, this meeting has revealed a wide variety of interesting approaches and applications that deal with temporal patterns. However, the problems in this area remain challenging and as such, there is little doubt that fundamental progress will continue to be made in the near future. This will in turn provide exciting novel research material for a follow-up BMVA technical meeting on this topic.

Dr Andrew Gilbert
University of Surrey
email: a.gilbert@surrey.ac.uk

- I'm going to show you new approach B, which gets better results A'.
- Slide 3 shows X, B and A'.
- So, I'll first tell you more about problem Y and technique Z.
- Then I'll ...

You get the point: this introduces the content of the talk, not the structure of the talk. The quick summary also helps people remember the content, just in case you want people to use your method and cite your paper.

Now that I have your attention, let me also comment on the last paragraph in your student's paper, where s/he introduces the rest of the paper. How many times have you read a paper where the last paragraph in the introduction starts: Section 2 gives the background of this paper, Section 3 gives the theory, Section 4 gives the experiments and Section 5 gives the conclusion?

Doesn't that tell you a lot? I bet you certainly remember the structure of the paper. Now, if instead the paragraph gave the 'take-home' message with pointers to the paper, then someone might remember the message and cite the paper. For example: "In this paper, we claim that technique X (see Section 3) works better (see Section 4 for evidence of this) than previous technique Y when applied to datasets like A, B, C (see Section 2)."

Professor Bob Fisher
University of Edinburgh
email: rbf@inf.ed.ac.uk

Guidelines when Taking Pictures for BMVA News

For best quality, **pictures should be at least 2 Mpixels in size**, but there is no point in going above 4–5 Mpixels, as the additional resolution won't be visible using our current methods of reproduction. Note that going above 6 Mpixels results in handling difficulties.

The basic rule for high quality publishing is that picture resolution should be 300–400 pixels per inch (~140 pixels per cm). Taking the column width as 8 cm, this means that pictures need to be ~1 Mpixel in size, but this takes no account of the fact that I frequently trim them, so 2 Mpixels is a better guide to submitted size.

Professor Roy Davies
Editor, BMVA News
email: e.r.davies@rhul.ac.uk

David Marr's *Vision* – Back in Print!⁴

Vision – A Computational Investigation into the Human Representation and Processing of Visual Information, the classic monograph by David Marr, has just been reprinted with a new foreword by Shimon Ullman and a new afterword by Tomaso Poggio.

David Marr's posthumously published *Vision* (1982) influenced a generation of brain and cognitive scientists, inspiring many to enter the field. In *Vision*, Marr describes a general framework for understanding visual perception and touches on broader questions about how the brain and its functions can be studied and understood. Researchers from a range of brain and cognitive sciences have long valued Marr's creativity, intellectual power, and ability to integrate insights and data from neuroscience, psychology, and computation. This MIT Press edition makes Marr's influential work available to a new generation of students and scientists.

In Marr's framework, the process of vision constructs a set of representations, starting from a description of the input image and culminating with a description of 3D objects in the surrounding environment. A central theme, and one that has had far-reaching influence in both neuroscience and cognitive science, is the notion of different levels of analysis – in Marr's framework, the computational level, the algorithmic level, and the hardware implementation level.

Now, thirty years later, the main problems that occupied Marr remain fundamental open problems in the study of perception. *Vision* provides inspiration for the continuing efforts to integrate knowledge from cognition and computation to understand vision and the brain.

David Courtnay Marr (1945–1980), one of the originators of the field of computational neuroscience, was Professor of Psychology at MIT. Shimon Ullman is Samy and Ruth Cohn Professor of Computer Science at the Weizmann Institute of Science, Rehovot, Israel. Tomaso Poggio is Eugene McDermott Professor in the Department of Brain and Cognitive Sciences at MIT. Both Ullman and Poggio both worked with David Marr at MIT.

July • 7 x 9 • 432 pp. • 150 illustrations
£22.95 • paper • 978-0-262-51462-0

Carrie Stewart
The MIT Press
email: cstewart@HUP-MITpress.co.uk

⁴Please ask me if you would like a review copy of this book – Ed.

Last Views of Aberystwyth ...

