

# BMVA News

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

Volume 20 Number 2  
December 2009

**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** News<sup>1</sup> 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 1 March 2010.

## Contents

Editorial: <i>New Year, New Climate?</i> .....	1
New Year Message from the BMVA Meetings Secretary.....	3
Prizes and awards at BMVC 2009.....	3
BMVA Publicity Officer.....	4
Student Travel Bursaries.....	4
ICPR Tutorials.....	4
BMVC 2010 – Call for Papers.....	5
BMVA Sullivan Thesis Prize – Call for Nominations.....	5
ICCV 2009: <i>Reports by Recipients of BMVA Travel Bursaries</i> .....	6
Sam Johnson.....	6
Jania Aghajanian.....	8
Peng Ren.....	10
Roberto Montagna.....	11
David Ellis.....	12
Alastair Moore.....	13
Patrick Ott.....	14
Report on UK Student Vision Workshop 2009.....	15
Report on ACCV 2009.....	16
Report on Articulated Human Motion.....	16
Robotics and Vision – Call for Participation.....	17
Book Announcements.....	18
Books for Review.....	18

<sup>1</sup>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.

## Editorial: *New Year, New Climate?*

BMVA has been going for a long time, twenty years to be precise. Over this time a lot has happened. Computers have come of age in that even PCs have high speed and huge storage, and above all are no longer news – everybody has several. The World-Wide-Web, which was first talked about twenty years ago, has grown alarmingly, so that everyone has a broadband connection capable of handling hundreds of megabits per second. Surveillance cameras are everywhere, and in London at least, we are each observed by over a hundred on a typical day, Britain being the most watched nation. Machine vision has become easy, to the extent that high speed implementation is no longer the main problem, which is now how to design robust software to handle complex data effectively (I almost said efficiently, but that is now less important). And with all these changes, we can congratulate ourselves that we have over this period done a good job and increasingly matched the needs of the community.

But pride comes before a fall, or can do if we don't look out for ourselves. After all, look what happened to computing. Instead of using computers with difficulty, everyone uses them with ease and scarcely a thought. The subject of computer technology, once central to engineering, has evaporated and computer science is taught in computer science rather than engineering departments. (Do I exaggerate? Yes, but not by a huge amount.) And could machine vision go the same way as computing? Our success is that everyone is using machine vision, and with that familiarity, it is no longer news to anything like the same extent. Already we see the BMVC papers of yesterday, with all their theory of 3D, optical flow and invariants, being replaced by the applications orientated papers of today. The subject is slipping out of our grasp and into that of the general



scientific populace. The best we can do is try to fight a battle of keeping the subject rigorous, but this is by now a rearguard action, as it is almost out of control and is in the process of being taken over by non-cognoscenti. Of course, I'm not aiming to offend anyone – just to say it like it is. What should we do? A common dictum is “If you can't beat 'em, join 'em” – which in this case can be read as, get into collaboration with the doers, the achievers, the engineers, the technologists, the people who are making the money, and then show off your expertise and how machine vision really should be done. For the other way is to bury one's head in the sand and see a lot of papers that reflect no real robustness, no real strategy, and no real provenance deluging our journals. “Well”, I hear the cry, “leave all that to peer review: such papers will never get past the reviewers”. But (a) pukka reviewers will not be bothered to deal with all the non-rigorous papers, so the reviewing time will go through the roof; (b) many of the reviewers will be the non-cognoscenti, who will find the work interesting and let it through. And so machine vision could move into oblivion and even acquire a bad name, as AI did in the 1970s. (There were reasons for that – it arose because of overclaims – but I see that happening to machine vision too, if the new application-style workers don't understand the subject in sufficient depth.)

But there is another far more distressing problem on the horizon – as I look out of my window and see the six inches of snow that fell last night (the worst fall since the early 1980s). Could this be symptomatic of climate change? In spite of the failure of the Copenhagen Conference, it did reveal startling agreement between scientists about the causes of climate change, and about its reality: the failure of the conference was political, which was always going to be a problem. Nevertheless, governments are going to have to move funds into that area, and hence they will have to move them away from other areas, even in science. Whether particle physics (a

huge spender) will survive is unclear; but really radical solutions will be needed, even to the extent that anything that is not in one way or another concerned with energy and the environment will be squeezed to the limit. Oh, there is one other area that will be saved: that of waste management and recycling. If we are concerned about the future of our subject, we should start moving towards these areas at a rate of knots. Unfortunately, I can't see quite how we can deal with the CO<sub>2</sub> problem ourselves, nor with the energy problem; but the waste and recycling problem is an area where we can contribute substantially. However, I wonder if we have yet thought this through. Possibly, Government and the EPSRC will get there first. The consequence of this is that we will miss the boat and be taken by surprise. What we need is a New Year resolution to boldly go where no man has gone before, beyond and behind the frontiers of science, showing how machine vision can lead the way into a more secure future where waste no longer exists, and energy problems are treated innovatively and responsibly. Above all, remember the Chinese curse (for that is what it is) “May you live in interesting times”. I'll leave you to work out whether I'm a pessimist or an optimist, or simply an organ grinder who wants us all to survive somehow.

Finally, I would like to welcome the articles from our triumvirate of action-men, who have thought long and hard about how we will be moving forward in the New Year. They are our Meetings Secretary, Dr Dimitrios Makris, our Publicity Officer, Dr Nicholas Costen, and our Student Bursary Officer, Dr Adrian Clark. Their articles appear on the following two pages.

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

## New Year Message from the BMVA Meetings Secretary



One-day BMVA Technical Meetings have been running for more than 13 years. They are also called symposia, but I personally avoid this term because of the connotations coming from the Greek origin of the word “Symposium” (literally, “drinking together”). In any case, they provide excellent value for money, as researchers can attend high-quality technical presentations and talks at minimum cost. At the same time, presenters have the opportunity to present their research to specialised audiences.

Any BMVA member can propose a topic and chair a relevant meeting, and this is evident by the history. From well-regarded professors to young researchers, alone, in pairs or even in groups, BMVA members take the opportunity to bring existing research communities together or develop fora for new research areas. This volunteering task is not only important for the success of the meetings, but also very beneficial for the chairs themselves, as it provides them with valuable experience and visibility.

The diversity of meetings is also remarkable. They may focus on traditional topics (e.g. “Machine Learning”, “Vision and Robotics”), on specific research problems (e.g. “Human Articulated Motion”, “Facial Analysis and Animation”, “Segmentation of Anatomical Soft Tissue Regions in Medical Data”), on specific approaches (e.g. “Group Theory, Invariance and Symmetry in Vision”), on specific applications (e.g. “Automotive Application”, “Security and Surveillance”), or even go across different disciplines (e.g. “Psychophysics and Vision”).

Fortunately, chairs’ duties are not onerous. Initially, they need to propose a suitable topic to the BMVA Meeting Secretary. Once their proposal is accepted, they have to prepare a “Call for Participation” to invite contributions to the meeting. At the same time they may want personally invited speakers (BMVA has a budget for their expenses). Then, they need to prepare a schedule for a one-day event. Finally, they will chair the meeting on the day. They are also encouraged to edit the papers for a Special Issue of the *BMVA Annals*.

BMVA will normally take care of all the other tasks, such as to advertise the Call for Participation and the Programme via its website and mailing list, book a venue, arrange catering and finally take care of the registration process.

While the majority of the meetings are hosted in Central London, they may take place anywhere in Britain. For instance, last year’s “Facial Analysis and Animation” meeting took place in Edinburgh and proved very successful (that was also attributed to the enthusiasm of the chairs – a group of four young researchers).

Chairing is an important task for the success of the BMVA technical meetings and is also rewarding for the chairs themselves. If you are interested in chairing one of the future BMVA technical meetings, do not hesitate to contact the BMVA Meetings Secretary (Dr Dimitrios Makris) and discuss your proposal with him, even if it does not exactly fit what has happened in the past. Visit the BMVA meetings webpage for up-to-date information on scheduled meetings, programmes, CfPs and detailed instructions for meeting chairs:

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

Dr Dimitrios Makris  
Kingston University  
[d.makris@kingston.ac.uk](mailto:d.makris@kingston.ac.uk)

## Prizes and awards at BMVC 2009

Unfortunately, the following information wasn’t published with the other prizes and awards in the last issue of BMVA News. For completeness, I’m including it in the present issue – my apologies for the earlier omission.

Sullivan Thesis Prize 2008:

“Combinatorial and Convex Optimization for Probabilistic Models in Computer Vision”  
Pawan Kumar Mudingonda, Oxford Brookes University

BMVA Distinguished Fellowships (2008 and 2009 awarded simultaneously):

2008: Andrew Zisserman  
2009: Bernard Buxton

Professor Roy Davies  
Editor, BMVA News  
email: [e.r.davies@rhul.ac.uk](mailto:e.r.davies@rhul.ac.uk)

## BMVA Publicity Officer

I thought that it might be a good idea to describe the role of BMVA Publicity Officer, which I have recently taken over. The major task is to raise awareness of the Association in the wider academic community: this will take place in two ways. First, we would like a wider, more exciting range of papers submitted to the annual conferences; second, we would like a broader group of people to realize that computer vision can help them with their research, so we can set up collaborations. Obviously, we're all on the lookout for collaborators, but often the people we meet don't fit with our own interests, though they might with those of other people.

To this end we are overhauling the web-site to better display the range of activities members engage in, and I am preparing fliers which can be given to people, pointing them at the web-site. We will require students receiving bursaries to attend conferences to ensure that they acknowledge the support while at the conference, and at the same time to pass out the fliers. In addition, we are particularly keen for anyone going to conferences associated with potentially under-represented user groups (including for example archaeology and cognitive science) to contact me and receive fliers to hand out. In addition, we are seeking to improve the web-site and perhaps create a demonstration video, the basic sources for which are the 'additional materials' uploaded with conference submissions. I will be writing to individuals shortly about using their material in this way.

Dr Nicholas Costen  
Manchester Metropolitan University  
email: n.costen@mmu.ac.uk

## Student Travel Bursaries

If you're an avid, or even a cursory, reader of BMVA News, you'll have noticed quite a few reports of conferences of late. This is because students who receive travel bursaries to attend international conferences have to write such reports as a condition of receiving support. The reports are good reading for readers who were not able to get to those conferences – after all, we can't spend all our time flitting from one conference to another – and are good opportunities for students to hone their communication skills.

I'm pleased to report that the BMVA awarded the following bursaries in 2009:

### ACCV 2009

Mai Xu, Imperial College

AIAA Guidance, Navigation and Control Conference  
Jason Gauci, Cranfield

Computer-assisted radiology and surgery  
Sophia Michopoulou and Irving Dindoyal, UCL  
(shared)

### CVPR 2009

Patrick Buehler, Oxford  
Chen Change Loy, Queen Mary

### ICASSP 2009

Tim Popkin, Queen Mary

### ICCV 2009

Jania Aghajanian, UCL  
David Ellis, Oxford  
Sam Johnson, Leeds  
Zdenek Kalal, Surrey  
Roberto Montagna, UEA  
Alastair Moore, UCL  
Patrick Ott, Leeds  
Peng Ren, York

### ISBI 2009

Dominique Van de Sompel, Oxford

Meeting on the Organization of Human Brain Mapping  
John McGonigle, Bristol

### SIGGRAPH 2009

U Mohammed, UCL

You will see from the list that we have been able to support travel not only to some of the major international vision conferences but also to some events that occupy interesting nooks and crannies of vision research.

The BMVA hopes to be able to support a similar number of travel bursaries in 2010. Details of the application procedure are given on the BMVA website.

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

## ICPR Tutorials

Please note that the deadline for submitting proposals for tutorials at the next ICPR (to be held at Istanbul in 2010) is 15 February 2010.

Denis Laurendeau  
Co-Chair Tutorials, ICPR 2010  
IAPR Secretary  
email: denis.laurendeau@gel.ulaval.ca

## BMVC 2010 – Call for Papers

BMVC 2010 will be held at the Penglais Campus of Aberystwyth University, Aberystwyth, Ceredigion, Wales on 30 August – 2 September 2010.

The British Machine Vision Conference (BMVC) is one of the major international conferences on machine vision and related areas. Organised by the British Machine Vision Association, the 21<sup>st</sup> BMVC will be held in Aberystwyth, Wales.

Authors are invited to submit full-length high-quality papers in image processing and machine vision. Papers covering theory and/or application areas of computer vision are invited for submission. Submitted papers will be refereed on their originality, presentation, empirical results, and quality of evaluation. All papers will be reviewed *doubly blind*, normally by three members of our international programme committee. Please note that BMVC is a single-track meeting with oral and poster presentations, and will include two keynote presentations, to be delivered by Professor Martial Hebert, Robotics Institute, Carnegie Mellon University and Professor Jean Ponce, INRIA, France, with a tutorial to be announced soon.

Topics include, but are not limited to:

- Statistics and machine learning for vision
- Stereo, calibration, geometric modelling and processing
- Person, face and gesture tracking
- Motion, flow and tracking
- Segmentation and feature extraction
- Model-based vision
- Image processing techniques and methods
- Texture, shape and colour
- Video analysis
- Document processing and recognition
- Vision for quality assurance, medical diagnosis, etc.
- Vision for visualisation and graphics

A workshop to be held on 2 September on *Machine Vision Research and Development* by PhD students will also be organised.

Four bursaries will be given to PhD student first authors whose papers have been accepted. These are sponsored by Microsoft Research, EPSRC, Springer, Wiley; the Research Institute of Visual Computing also sponsor several awards: best conference paper, best student paper, best video, best student poster, best workshop paper, etc.

Important dates:

- Paper registration deadline: 23 April 2010, 17:30
- Submission deadline: 30 April 2010, 17:30
- Notification of acceptance: 14 June 2010
- Camera-ready papers: 13 July 2010
- Conference: 30 August – 2 September 2010

General chair: Fred Labrosse, Aberystwyth University  
 Publicity chair: Yonghuai Liu, Aberystwyth University  
 Technical programme chair: Reyer Zwiggelaar, Aberystwyth University  
 Tutorial and workshop chair: Peter Bunting, Aberystwyth University  
 Local organising Committee: Meinir Davis, Alan Woodland, Ran Song.

Further information is available from the conference website:

<http://bmvc10.dcs.aber.ac.uk>

BMVC 2010  
 Aberystwyth University  
 email: [bmvc2010@aber.ac.uk](mailto:bmvc2010@aber.ac.uk)

## BMVA Sullivan Thesis Prize – Call for Nominations

The BMVA annually awards a Best Thesis prize (to commemorate the contribution made by the late Professor Geoff Sullivan) to the best doctoral thesis submitted to a UK University, in the field of computer or natural vision.

Recommendations for the prize are considered by a Selection Panel appointed annually by the BMVA Executive Committee. The decision of the Selection Panel is announced at the end of the following July. When possible, the presentation will be made at the conference dinner of the British Machine Vision Conference, held annually during September.

The BMVA Executive Committee now seeks nominations for the Sullivan Prize for theses examined during the calendar year 2009. Please send any nominations to the BMVA Secretary, Dr Neil Thacker ([secretary@bmva.ac.uk](mailto:secretary@bmva.ac.uk)) by the end of February 2010.

Nominated theses should be made available through a web page: the successful author is expected to make



his/her thesis available as a PDF for distribution via the BMVA web-site from September 2010 onwards.

For conditions, please see:

<http://www.bmva.ac.uk/admin/sullivan.html>

Professor Roy Davies  
Editor, BMVA News  
email: [e.r.davies@rhul.ac.uk](mailto:e.r.davies@rhul.ac.uk)

## ICCV 2009: Reports by Recipients of BMVA Travel Bursaries

Japan – home of sushi, sumo and one of the world’s leading countries for electronic and robotic research – a perfect place to host the 12<sup>th</sup> International Conference on Computer Vision (ICCV 2009). The city of Kyoto is no stranger to hosting large international conferences, being well known throughout the world as the location in which the Kyoto Protocol was signed. The Kyoto International Conference Center played host to this event and is where computer vision researchers from across the globe travelled to during September 2009.

Prior to the main conference, two days of tutorials and workshops were held on the Kyoto University campus. Although this was quite far from the hotel, I had been issued with a welcome pack upon arrival so finding the correct combination of subway and buses was a breeze.

The first day consisted of a number of tutorials of which I attended the Local Texture Descriptors tutorial by Matti Pietikäinen and Guoying Zhao in the morning. This mainly focused on the local binary pattern (LBP) descriptor that is becoming widely used due to its descriptive power and computational efficiency. In the afternoon I attended Siwei Lyu and Stefan Roth’s tutorial on “Modelling Natural Image Statistics for Computer Vision”. They gave an excellent introduction to the use of Conditional and Markov Random Fields to model the pixel or wavelet based statistics of images. During the second day I got the chance to present my work on single image articulated human pose estimation at the workshop on *Machine Learning for Vision-based Motion Analysis*. This was an excellent opportunity to share my work and get input from other students working in similar areas. The keynote session by Li Fei-Fei on *Detecting and Recognizing Moving Humans* was particularly insightful and she introduced a number of concepts that were discussed further in my presentation afterwards.



The hotel was situated opposite the main Kyoto station and almost directly below Kyoto tower.

On the Tuesday the main conference began, and with it a change of travel routine. Thankfully, the International Conference Center was situated conveniently at the end of the K-line subway that ran regularly from the city centre. The conference centre itself was an imposing sight looking like a mix of traditional Japanese architecture and something out of Star Wars! Fortunately, the ‘Red HAPPI Helpers’ were on hand to make sure no-one got lost and we were all fully fuelled on coffee, green tea, and Japanese biscuits.



The entrance hall of Kyoto International Conference Center.

Each day of the main conference was packed with events but it was single-track, so everyone had an opportunity to see every oral and poster. The poster sessions were around three hours long after lunch and were sandwiched between morning and afternoon oral sessions. This made for a fairly pleasant day as people could relax and network over lunch and then stroll back to the conference centre for the poster session. Before the first set of talks got underway a number of awards were announced including the Marr Prize for *Best Paper* awarded to Desai, Ramanan and Fowlkes for their work “Discriminative Models for Multi-class Object Layout”. The Computer Vision Significant Researcher Award was presented to Andrew Blake who gave an excellent acceptance speech, offering us all an insight into what he believes lies ahead in the field and the experiences he has had in his distinguished career.



The main hall of Kyoto International Conference Center.

Among the oral presentations given, a number stood out. Probably the most interesting from a purely scientific point of view was the paper entitled “Looking Around the Corner Using Transient Imaging” from the MIT Media Lab which tried to convince us that it is possible to capture images of objects around corners without the use of mirrors or any fancy gadgets other than the camera. The work at this stage was purely proof-of-concept and used lab experiments to show that it is possible to sense images around corners using ‘ultra-high-speed’ femtosecond cameras. These are able to measure the time taken for photons to travel from an emitter on the camera to an object and back via bounces of another surface. This allows a model of the hidden object to be built in a similar way to time-of-flight cameras used to gather depth information.

Another interesting oral was for the paper “A Shape-based Object Class Model for Knowledge Transfer” by Michael Stark, Michael Goesele and Bernt Schiele. They offered a number of contributions such as a novel method of modelling object parts using symmetry axes

and how to transfer knowledge learnt about one object to another. An example shown was that of learning a part model for a giraffe and then using the same learnt structure to rapidly learn a model for a swan. Finally, the Marr Prize winning paper was presented by Deva Ramanan on “Multi-class Object Layout”. The paper showed that learning co-occurrences of object classes and their layout in a scene can improve classification performance. For example, if one was to detect a car then it is very unlikely that a sofa will be present in the same scene. However, if one detects a chair then there is likely to be a human somewhere above it.



The Kyoto University Clock Tower Centennial Hall hosted some of the workshops and tutorials.

After the main conference had ended we returned to the Kyoto University campus for another two days of workshops. The highlight of this final weekend was the PASCAL Visual Object Classes Challenge event which has been growing larger every year. This is an opportunity to see how the current state-of-the-art methods in object detection, image classification and segmentation perform and also to get a chance to question the authors.

Finally the PAMI-TC meeting must be mentioned. This was held on the Wednesday evening and gave an opportunity for everyone present to vote on key issues in the computer vision field such as where the 2013 ICCV was to be held. There were two candidate cities – Sydney and Philadelphia – which provided presentations in an attempt to win over the crowd. The Philadelphia presentation was very impressive but it was Sydney that emerged victorious. This upset many of the Americans present as they believed it was their turn due to the traditional Europe–Asia–America rotation. However, Australia has never had the opportunity to host ICCV and doing so should give a large boost to Computer Vision researchers down-under.

The next ICCV will be the thirteenth and will be hosted in Barcelona, Spain in 2011. Finally I would like to



thank the BMVA for providing a generous travel bursary to allow me to attend this conference.

Sam Johnson  
University of Leeds  
email: mat4saj@leeds.ac.uk

The 12<sup>th</sup> International Conference on Computer Vision (**ICCV 2009**) took place at the Kyoto International Conference Center (Kyoto, Japan), on 27 September – 4 October 2009. The first and last days of the conference were devoted to workshops and tutorials. These days were packed with many interesting tutorials including: MAP Inference in Discrete Models, Local Texture Descriptors in Computer Vision, and Boosting and Random Forests for Visual Recognition. Twenty workshops were held in parallel, which included *Dynamical Vision*, *Embedded Computer Vision* and *3D Digital Imaging and Modeling* as well as some new workshops such as *e-Heritage and Digital Art Preservation*, *Human-Computer Interaction* and *Emergent Issues in Large Amounts of Visual Data*.



Professor Berthold Klaus Paul Horn receiving the *Azriel Rosenfeld Lifetime Achievement Award*.

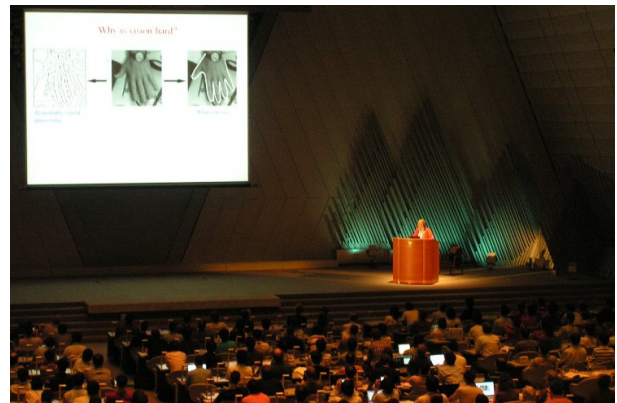
The PASCAL Visual Object Classes Challenge 2009 took place on 3 October. The joint winners of the detection challenge were Felzenszwalb *et al.* from UoC/TTI Chicago using discriminatively trained part based models and Vedaldi *et al.* from Oxford/MSR India using multiple kernels for object detection. The winner of the classification challenge was Gong *et al.* from NEC/UIUC using Gaussian mixture and local coordinate coding. The segmentation prize went to Carreira *et al.* from University of Bonn using ranking figure-ground hypotheses for object segmentation.

ICCV 2009 included a special event titled: Colloquium on Fundamental Advances in Computer Vision which had a star-studded list of speakers including: Takeo Kanade, Makoto Nagao, Martial Herbert and Shree Nayar. In this event experienced researchers were invited to provide their perspectives on emerging

problems of modern computer vision such as Internet Vision: Challenges and Opportunities and Robust Computer Vision Techniques and Applications.



Professor Andrew Blake receiving the *Computer Vision Significant Researcher Award*.



Professor Andrew Blake giving his acceptance speech.

The main conference kicked off with the award ceremony held at the astonishing main hall of the Kyoto International Conference Center. The recipient of the *Computer Vision Significant Researcher Award* was Professor Andrew Blake. The *Azriel Rosenfeld Lifetime Achievement Award* was given to Professor Berthold Klaus Paul Horn. The *Best Paper Award* went to Desai *et al.* for the paper titled: “Discriminative Models for Multi-class Object Layout.” Kirmani *et al.* received an *Honorable Mention* for the paper titled: “Looking Around the Corner Using Transient Imaging.”

The conference included 14 oral sessions (48 papers) and four poster sessions (260 papers), covering a wide range of research topics such as segmentation, human detection, learning, activity recognition, video and image sequences, similarity metrics, low-level vision and recognition and matching. I would particularly like to mention a few papers that made their datasets publicly available:

- “Quasi-Periodic Event Analysis for Social Game Retrieval” by Wang *et al.* They made available a



video dataset of annotated social games and social interactions, containing both home movies and laboratory videos. This dataset can be used for action/behaviour recognition.

- “Attribute and Simile Classifiers for Face Verification” by Kumar *et al.* presented an evaluation of human performance on face verification on the LFW data for the first time. They also introduced the PubFig dataset, the largest data set of real-world images (60,000) for face verification (and recognition) publicly available at <http://www.cs.columbia.edu/CAVE/databases/pubfig/>.
- “LabelMe video: Building a Video Database with Human Annotations” by Yuen *et al.* They developed an online video annotation tool soon to be released to the public. This aims to create an open database of videos where users can upload, annotate, and download content efficiently. Currently, the database contains a total of 1903 annotations, 238 object classes, and 70 action classes.



A Museum guide robot at the demo session (Saitama University and Tokyo University of Technology).

In addition to talks and posters, there were four sessions dedicated to demos from universities and exhibitions from industry. These included a museum guide robot (Saitama University and Tokyo University of Technology) which explained exhibits to multiple visitors through gesture recognition. The “On-Site 3D

Video Capture System” from Kyoto University was also among the exciting demos which offered a 3D model of the volunteers to take home as a souvenir! The industry representatives included Omron demonstrating their products on real-time face detection and pose estimation and Toshiba displaying their new range of laptops that enabled automatic login through face recognition.



A group playing traditional music using Japanese instruments at the banquet.



Japanese traditional dance performance at the banquet.



Performance from a group of Geishas at the banquet.

Last but not least, a wonderful banquet was held at the Grand Prince Hotel, where the conference participants enjoyed a variety of great Japanese cuisine as well as traditional Japanese dance and music performances.

I would like to thank BMVA for supporting this trip and giving me the opportunity to meet international researchers and communicate my research interests.

Jania Aghajanian  
University College London  
email: j.aghajanian@cs.ucl.ac.uk

The 12<sup>th</sup> International Conference on Computer Vision (ICCV 2009) was held in Kyoto, Japan on 29 September – 2 October 2009. The city Kyoto, with a population close to 1.5 million, is located in the central part of the island of Honshu, Japan. It is the former imperial capital of Japan. It is now the capital of the Kyoto Prefecture and a major part of the Osaka–Kobe–Kyoto metropolitan area as well.

The conference venue was located at the Kyoto International Conference Center, which is built in a traditional architectural style and enclosed by brooks, woods and mountains. The conference itself was well organised, and everything ran smoothly. Traditional Japanese food was served at the reception and banquet, accompanied by traditional Japanese performances. These traditional elements combined to give participants a flavour of Kyoto as an ancient seat of learning.



Oral Session: Photo courtesy of Shiqi Yu

In keeping with its standing as one of the top computer vision conferences, ICCV this year accepted 308 papers out of 1327 complete submissions, with 48 papers as orals (3.6%) and 260 papers as posters (19.6%). The first session of the main conference was an award ceremony. The paper “Discriminative Models for Multi-class Object Layout” by Chaitanya Desai, Deva Ramanan and Charless Fowlkes won the *David Marr Prize*. Berthold Horn won the *Azriel Rosenfeld Lifetime Achievement Award*, Andrew Blake the *Computer*

*Vision Significant Researcher Award* and Yehezkel Lapidan and Haim Wolfson the award for the *Paper to Have Stood the Test of Time*. The following oral and poster sessions were packed.

Several research trends can be identified at ICCV 2009. The first of these is the increasing interest in exploiting vision research for internet applications. Representative papers include “I Know What You Did Last Summer: Object-level Auto-annotation of Holiday Snaps” by Stephan Gammeter, Lukas Bossard, Till Quack and Luc Van Gool, and “Learning Actions from the Web” by Nazli Ikizler-Cinbis, Ramazan Gokberk Cinbis and Stan Sclaroff.



Poster and Demo Session: Photo courtesy of Shiqi Yu

Another trend is the combination of techniques from vision and graphics, or solving vision problems from a graphics perspective. Representative papers include “Building Rome in a Day” by Sameer Agarwal, Noah Snavely, Ian Simon, Steven M Seitz and Richard Szeliski, and “Reconstructing Building Interiors from Images” by Yasutaka Furukawa, Brian Curless, Steven M Seitz and Richard Szeliski.

Machine learning in vision is still a very active area, and papers on the topic at ICCV 2009 included the *David Marr Best Paper* “Discriminative Models for Multi-class Object Layout” by Chaitanya Desai, Deva Ramanan and Charless Fowlkes.

Finally, there is an increasing amount of work on testing vision techniques in practical situations such as in supermarket checkouts (“Action Detection in Complex Scenes with Spatial and Temporal Ambiguities” by Yuxiao Hu, Liangliang Cao, Fengjun Lv, Shuicheng Yan, Yihong Gong and Thomas S Huang) and in diagnosis (“Quasi-periodic Event Analysis for Social Game Retrieval” by Ping Wang, Gregory D Abowd and James M Rehg).

Peng Ren  
University of York  
email: pengren@cs.york.ac.uk



The 12<sup>th</sup> IEEE International Conference on Computer Vision (**ICCV 2009**) took place in Kyoto, Japan, on 27 September – 4 October 2009. Four days were allocated to workshops, so the main conference actually ran from 29 September to 2 October. The event was divided between two venues: the workshops and the tutorials were held in the campus of Kyoto University, and the main conference took place at the Kyoto International Conference Center. This is a world-class facility, where in 1997 the Kyoto Protocol was devised and signed by the delegates of the United Nations. Such a large and inspiring venue was necessary in order to face the high attendance to this conference: 308 papers, 48 of which were oral presentations, plus 290 workshop papers brought more than 1300 delegates to Kyoto. Moreover, from Wednesday a demo session was running in parallel with the poster sessions. There, different companies and research centres were showing their work practically, although in some cases they were marketing a product rather than presenting a specific technological advancement.



My area of research is not computer vision, thus I could not appreciate certain sessions. However, this field is so vast that it is difficult to have a clear overview of all its aspects. Recognising this, the organisation announced that since 2009 a new conference had been established, the IEEE International Conference on Computational Photography. The session on *Computational Photography* that used to be part of the program of ICCV has been renamed as *Sensing*, and the papers presented there are less related to photography and more to computer vision.

The oral session that seemed to be the most interesting for my work was *Shading and Color*. Although some interesting papers were presented, none of them was really related to my project. The first one, “Estimating Natural Illumination from a Single Outdoor Image”, by Lalonde *et al.*, described a method to determine the position of the sun from outdoor pictures. The final purpose of this knowledge would be to plausibly illuminate virtual objects in order to insert them in a real photograph. Different cues are taken into account for

this purpose: the shadows, the sky, and the vertical surfaces present in the image. The real novelty of this approach is the use of a prior describing the likelihood of the sun’s position in the sky. To acquire this knowledge, the authors used large numbers of photographs available on the internet, tagged with geographical coordinates and time of the day. The probability distribution of the sun’s position is very different from a uniform distribution, because the places where most pictures are taken are not uniformly distributed on the planet. From this information, the authors obtained a probability distribution that aided the visual cues present in the image. In the same session, Swirski *et al.* presented a paper called “Stereo from flickering caustics”, where the authors proposed a new method to reconstruct 3D information from underwater stereo images. The basic idea does not seem complicated: the waves on the surface of the water refract the light in such a way that below the surface the illumination varies spatially and temporally, causing what is known as underwater caustics. The authors exploit the temporal change of these caustics in order to solve the stereo correspondence problem. They suggest that at each point of the scene the temporal variation in the illumination due to the caustics is unique (or, at least, it is highly unlikely that the same temporal variation happens in different points). Therefore, they use this as a signature to find the stereo correspondence between two images, thus reconstructing a reliable 3D image of the scene.

Another paper worth mentioning, in the *Sensing* session, was by Glasner *et al.*, “Super-resolution from a single image”. Here, the authors suggest a way to recover high frequency details using only the information present in the low resolution one, without the need for any external prior. They exploit the fact that, in an image, certain patches tend to recur within the same scale and across different scales. Therefore, using the redundancy of this information, they reconstruct high frequency information at larger scales.

As my background is computational colour, I attended and presented my paper at the *Color and Reflectance in Imaging and Computer Vision* workshop. The morning keynote, “Photometric Issues in E-heritage”, was given by Professor Katsushi Ikeuchi. He discussed the 3D modelling and colouring of monuments for the purpose of generating an augmented reality experience for tourists. The main issue he considered was the colouring of the models, and how to retrieve the colours of the original monument using colour constancy techniques. Moreover, he delineated the problem of photometric consistency – that is, how the shadows of virtual objects interact with the real environment. The afternoon keynote, “Gradient Field Reintegration for Colour-based Problems”, was presented by Professor Graham

Finlayson. He showed different problems that share the issue of integrating a non-integrable gradient field: removal of shadows from images, data fusion and colourisation. He showed a path-based approach for the first: a shadow can be removed from an image taken with a calibrated camera simply by setting to zero the gradient at the edge of the shadow. In order to avoid setting to zero a material edge, several random paths are used, which enter and exit the shadow from the same point. For data fusion, he suggests obtaining a gradient field from a multichannel image using the structure tensor (or colour tensor) technique. This gradient is non-integrable, so he suggests approximating it with a linear combination of the gradient fields of the channels of the original image. A similar approach is used in colourisation, to solve the optimisation problem described by Levin *et al.* in their paper "Colorization using optimization". Apart from the keynotes, the papers often tackled very specific issues: for example, recognising the growth of certain microorganisms on the walls of monuments (Morimoto and Ikeuchi, "Multispectral imaging for material analysis in an outdoor environment using normalized cuts"), analysing the preservation status of food (Dissing *et al.*, "Temporal reflectance changes in vegetables") and automatically grading the colour and the cut of diamonds (Ren *et al.*, "Diamond color grading based on machine vision").

The next ICCV conference will take place in 2011 in Barcelona, Spain: the most likely venue for 2013 seems to be Sidney, Australia. I would like to thank the BMVA for funding my participation in this event.

Roberto Montagna  
University of East Anglia  
email: r.montagna@uea.ac.uk

The 12<sup>th</sup> International Conference on Computer Vision (ICCV 2009) was held in Kyoto, Japan on 27 September – 4 October 2009, with the main conference being hosted at the Kyoto International Conference Center and associated workshops at Kyoto University. The four days of the main conference were split into a total of 14 oral sessions and four poster sessions, all single track, exposing advances in a broad range of areas of computer vision. Out of an astounding 1317 submissions, 48 papers were accepted for oral presentation and 260 as posters. The core themes for the plenary presentation sessions were segmentation, recognition, human detection, learning, geometry, and sensing.

The headline prize winners were as follows. The *Azriel Rosenfeld Lifetime Achievement Award* was received by

Professor Berthold Horn for his extensive contributions to machine vision, including his landmark work on determination of optical flow. He gave an illuminating speech on how the field has drastically changed since his first work, describing how the now highly competitive nature of computer vision research has both positive and negative consequences. Professor Andrew Blake received the *Computer Vision Significant Researcher Award*. The *Best Paper* prize was awarded to Desai, Ramanan and Fowlkes for "Discriminative Models for Multi-class Object Layout".

While many impressive papers were presented, personal highlights include two related talks which were given by Sameer Agarwal ("Building Rome in a Day") and Yasutaka Furukawa ("Reconstructing Building Interiors from Images") in the geometry session. The former proposed large scale 3D reconstruction of whole cities using Internet photo databases, such as flickr, with a total processing time of the order of a single day. Their key contribution was the development of highly parallelisable matching and reconstruction algorithms that are suitable for distributed computing, allowing 3D reconstruction using 150,000 images. Furukawa's paper focused on reconstruction of indoor environments in a piecewise planar fashion under the assumption that most planes are oriented in just a few dominant directions. Both speakers demonstrated impressive results, and it was refreshing to see these complete systems gaining such exposure despite applications of visual geometry losing popularity in the field.

Other talks where appealing results were shown included those in the shading and colour session, for example in Huang and Chen's "Landmark-based Sparse Color Representations for Color Transfer", colour images are decomposed into an intensity image along with a sparse set of colour 'landmark' pixels which can then be used to transfer the colour properties to a different image. In the *Image and Video Editing* session Pritch *et al.* presented "Shift-map Image Editing" which improves upon Avidan and Shamir's "Seam Carving", and will undoubtedly become a major selling point of image editing software in the near future. Also of note were Collomosse *et al.*, "Storyboard Sketches for Content Based Video Retrieval", where user-provided pictorial 'squiggles', more often seen as guides to segmentation algorithms, are used as a mechanism to perform a search for videos with common visual content.

The poster sessions were inevitably crowded and chaotic with around 65 posters being presented each day. The work presented was of a very high standard, as could be expected from such a large number of submissions, and the sessions served as a useful platform to meet fellow researchers. Of particular personal interest was the session on video and tracking



where much of the work is highly applicable to visual surveillance. Developments in human detection and tracking are still highly active research areas, with tracking in increasingly difficult situations (such as crowds) becoming more reliable. This is just one case where improvements in lower level vision operations could lead to gains in higher level processing such as behaviour analysis in visual surveillance systems.

There were also four days of workshops and tutorials surrounding the main conference. I attended the 9<sup>th</sup> Workshop on Visual Surveillance to present a poster on “Modelling Pedestrian Trajectories with Gaussian Processes”. The keynote talk was given by Ian Reid, and explored the issues and benefits of combining visual surveillance and visual simultaneous localisation and mapping (SLAM). A highlight was the proposed linking of the SLAM and visual surveillance system’s coordinate frames such that views from static cameras could be used to give additional information to a mobile observer to see around corners, for example. Similar ideas were also prominent in a following talk given by Federico Pernici on tracking people in 3D using a base set of bundle adjusted visual landmarks. The poster session went well, and discussing my work with peers was extremely useful in providing ideas as to future research directions. I also had the opportunity to browse the posters of others, and found that the more informal nature of the workshop session, compared to the main conference, gave more of an opportunity to network with others who held similar interests; and I have made links that will undoubtedly be useful in the future.

Kyoto had many tourist attractions; however, I enjoyed most being able to reside in a working city and absorb the Japanese culture which is so different to anything that I had experienced before. I would like to thank the BMVA for awarding me a travel bursary and thus making the trip possible. ICCV 2011 will be held in Barcelona, and while ICCV can be intimidating due to its size, I would still urge those in their first few years of research such as myself to try to attend.

David Ellis  
University of Oxford  
email: dre@robots.ox.ac.uk

**ICCV 2009:** Kyoto, Japan, 27 September – 4 October 2009.

Thirty hours via Shanghai, without arriving at the destination city, made travelling to this year’s ICCV every bit as arduous as my last international conference (CVPR Alaska). What made it bearable, unlike most travel in the UK, was that Japan felt incredibly clean,

polite and safe. It also transpired that, despite all the warnings to the contrary, a half-hour tutorial with a Japanese colleague before travelling was sufficient to navigate a country where most of the signage was dual language.

Day one began with negotiating the Kyoto rush hour on a bus with the occasional glimpses of temples en route. I started by attending a tutorial on “MAP Inference in Discrete Models”. This was a particularly good overview of the subject for someone writing a literature review. This tutorial continued after lunch but I ducked out half way through to support my colleague, Yun Fu, who was presenting his work on “Spatial support of signal and noise in face recognition” at the *International Workshop on Sub-space Methods*. In the first year of my doctorate I was fortunate enough to attend the previous ICCV and I think it is invaluable for researchers at the very beginning of their careers to get exposure to work at the very highest level.

On day two I went to the *Special Events Colloquium on Fundamental Advances in Computer Vision*. Tutorials and workshops tend to focus on a narrow area of expertise and are, consequently, very technical. Thus, the colloquium made for a nice change as it was interesting to hear more general discussions about directions in Computer Vision research and get an insight into some of the ‘blue sky’ thinking from experts in their field. In particular, I enjoyed the talk on Internet Search by Microsoft’s Harry Shum in which he highlighted interesting directions in user interactivity in search and recognition. I also really enjoyed the talks at *3dRR-09*, including lucid keynote presentations from Derek Hoiem and Luc Van Gool.

The venue for the main conference was the Kyoto International Conference Center. We arrived at the centre via the impeccably clean metro and after a long walk down a tunnel emerged in front of a sprawling concrete structure that resembled something from a George Lucas film. It included satellite structures peeking out of the tree canopy, making you feel this could be the location of Ewok High Council. The Center was used for the 1997 UN Convention on Climate Change – Kyoto Protocol – and it was interesting to experience a little piece of modern history.

The facilities at the Center were really excellent. Until Kyoto, I had never experienced first-hand the effect that good architecture can have on your experience of a conference. The room was well lit, chairs well spaced and the display was clearly visible from all parts of the auditorium. All this made it easy to concentrate on the oral presentations. My vote for the strongest session was *Segmentation II*, with good talks from Vikas Singh, Victor Lempitsky and Andrew Delong.

I think there should be a special mention for “Filter flow” (Seitz and Baker), which tried to capture many different image transformations in a single unified framework – The Matrix! The authors set up the problem as computing space-variant linear filters that transform one image into another and showed some impressive results, including stereo, optical flow, lighting changes, blur and a combination of these effects.

From the posters I thought there was a very nice simple graph construction in “Beyond connecting the dots” (Windheuser *et al.*) for finding minimum paths that pass nearby, rather than through, control points. This would seem to have many general applications. On a personal note, feedback on my poster “Scene shape priors” was very positive – my favourite moment being when a Chinese attendee suggested I read, and cite, one of my previous papers. I’m not sure whether it’s a compliment or not that he read the paper in the first place and thought it worth mentioning, yet couldn’t remember that I had written it!

Unlike previous conferences I was really impressed with the demonstrations. This section of the conference can sometimes be a graveyard of screens with half working bits of kit implementing research that’s a few years old. However, this time there were some great spectacles to write home about including “GelSight: Retrographic Sensing for Touch, Texture, and Shape” (MIT) and a demonstration of new cameras in “Embedded AER Dynamic Vision Sensor for Low-Latency Pole Balancing” (ETH Zurich) that read out individual pixels rather than using a frame buffer. The cameras were so fast that the demonstration consisted of a two-camera setup balancing a pencil on its lead by moving the ground surface backwards and forwards in real time to prevent it falling. Magic!

I also thought the decision to have one long poster session either side of lunch was very successful. It felt like there was more time to take in the posters and that you weren’t bounced around between oral sessions all day. One useful addition would have been the one-page abstract books that were introduced successfully at BMVC this year. The abstracts force authors to distil the most important contributions of a paper into a single page, which is valuable for dissemination and has the added bonus of a ready-made summary for online content.

Other cultural curiosities from Japan included eating limpets and raw scallops, and toilets that play music and have remote controls. Lastly, if you are lucky enough to get back to Kyoto, look for the PVL tag in the A Bar (second cubicle on the left, above your head) – a small contribution from our lab to the rest of the world.

I would like to thank the BMVA for their generous contribution towards my travel, which made it possible for me to attend this very successful and worthwhile conference. The best I have attended yet!

Alastair Moore  
University College London  
email: a.moore@cs.ucl.ac.uk

This year the 12<sup>th</sup> International Conference on Computer Vision (**ICCV 2009**) was held in Kyoto, Japan on 27 September – 4 October 2009. The chosen venue, the Kyoto International Conference Center, was of particular interest as it was there where the leaders of the world signed the Kyoto protocol in 1997. The conference center is nicely embedded into the mountains surrounding Kyoto and a few sights, including a lake and a temple, offered a great distraction from the busy conference schedule.

The first two days after arrival in Kyoto were already very busy as many tutorials and workshops waited to be discovered. I visited the tutorial on “Local Texture Descriptors in Computer Vision” presented by Matti Pietikäinen and Guoying Zhao. The tutorial first introduced the theory behind texture descriptors and then focused around Local Binary Patterns (LBP) while also offering a great number of practical examples. In the afternoon I listened to the tutorial titled “Coloring Visual Search” by Cees Snoek, Theo Gevers and Arnold Smeulders. The speakers first explained the basic principles in color research for computer vision, continued by explaining various forms of color normalisation and what they are used for, while finalising the tutorial talk with a detailed explanation of their approach to the PASCAL visual object challenge for image search.

On the morning of the second day I attended the workshop on *Machine Learning for Vision-based Motion Analysis* and spent the second half of the day listening to the tutorial “Boosting and Random Forests for Visual Recognition” given by Tae Kyun Kim, Jamie Shotton and Bjorn Stenger. All three of them provided the listeners with a strong introduction to boosting and random forests while also covering practical aspects of both algorithms.

The next four days were reserved for the main conference and were rich with many good talks and posters. In the first segmentation session Schoenemann *et al.* presented an approach to segmentation by considering a class of region-based energies for the tasks of image segmentation and inpainting which

combine so-called region integrals with curvature regularity of the boundary region.

Sameer Agarwal and co-authors presented their work on 3D reconstruction entitled “Building Rome in a Day”. Their goal was to provide a very good reconstruction at the scale of a city within a 24-hour timeframe. At the end of their talk they presented a very impressive demo of a dense reconstruction of the Sistine Chapel.

Schwartz *et al.* gave a presentation on a human detection system which utilises multiple features ranging from edge- over texture- to color-based representations of the input window in combination with a non-linear classifier. The next speaker, Xioayu Wang (*et al.*), presented a paper exploring a similar direction. Their detection system combined edge- and texture-features (HOG&LBP) for human detection and offered a novel approach to handling occlusion for the task of person detection. Both approaches produce very good results in the field of pedestrian detection.

Deva Ramanan was awarded the Marr Prize for his paper entitled “Discriminative models for multi-class object layout”. The paper presented a new way of non-maximum suppression of multiple object detections in an image compensating for and making use of overlapping detections of multiple classes as well as their relationship to each other.

Another piece of work I found particularly interesting was the work of Farhadi *et al.* entitled “A latent model for discriminative aspect”. The paper deals with the problem that a feature representation of an object might be very different under very viewpoints. They propose to use a bilinear model which utilises latent variables to describe the aspect and embed the approach in a max margin learning framework.

Very interesting was also the PAMI TC meeting in which I was allowed to vote where ICCV 2013 is going to be held. The choice was between Philadelphia and Sydney. After a long controversy which led to the final question whether Australia can be considered part of Asia or not, the decision was made to hold ICCV 2013 in Sydney.

Kyoto itself is a wonderful city offering many sights, including an enormous collection of temples. Prior to visiting Kyoto we spent one week in Osaka and Hiroshima and both trips were an outstanding experience.

I am very thankful to the BMVA and the EPSRC project HAML for providing funding so that I could undertake this trip to ICCV.

Patrick Ott  
University of Leeds  
email: ott@comp.leeds.ac.uk

## Report on UK Student Vision Workshop 2009

The UK Student Vision Workshop was held on 11 September at the People’s Palace, Queen Mary University of London. The first three talks were on Video and Graphics. First, Umar Mohammed from UCL presented a method for generating novel facial images, by combining patch-based texture synthesis methods with a global parametric model of the face. Next up was Chuan Li from Bath, who presented a method for extracting the branch structure of a tree from a video sequence by finding areas of the canopy with similar motions. After that was Patrick Buehler from Oxford, with a system for learning British Sign Language signs from subtitled TV clips using upper body modelling and hand positions.

The next set of talks was on Low Level Vision: Oliver Woodward, of Oxford, discussed MAP inference and Markov Random Fields, with noise reduction as an example, and went on to discuss a more general Marginal Probability Field. Nassir Mohammed from Cardiff discussed the non-Gaussian behaviour of image statistics and an optimisation-based method of automatic colourisation of greyscale natural images. Then Alastair Moore from UCL presented a method for super-pixel segmentation of a scene based on a prior distribution of boundaries found in street scenes.

There was a wide range of posters, covering topics from temporal calibration of a robot arm and camera to sketch-based video retrieval and an evolution-based method for video editing.

The last four talks were on object recognition and segmentation. Patrick Ott of Leeds went first, investigating pedestrian detection using implicit segmentation to improve Histogram of Gradient features. Then came Sudeep Sundaram from Bristol, with a framework for describing and learning manipulations of objects, captured using a wearable camera. Chris Russell of Oxford Brookes discussed the use of associative hierarchical CRFs in image class segmentation; finally, Chen Change Loy of QMUL described a system for modelling correlations between activities recorded by a camera network.

The day was very interesting and well presented, giving a good opportunity to look at state of the art developments in British machine vision.

A J Douglass  
Cambridge University  
email: [ajd86@cam.ac.uk](mailto:ajd86@cam.ac.uk)

## Report on ACCV 2009

The 9<sup>th</sup> Asian Conference on Computer Vision (ACCV 2009) was held in Xi'an China on 23–27 September, 2009. ACCV is the main conference of computer vision in the Asian/Pacific region and also around the world. This year ACCV received 670 papers, but only 169 papers from 24 countries were accepted (the acceptance rate was 26%). The conference included three plenary talks, nine oral sessions and three poster sessions.

The first day of the conference started with a *Multiple View and Stereo* session. In this session, I think the most attractive one is the talk on voting approach for estimating robust focal length, presented by Bujnak from the Center for Machine perception, Czech Technical University. The other session in the morning was *Face and Pose Analysis*. Then, an excellent plenary talk by Davis from Maryland was given in the afternoon. In the talk, Davis described the problem of segmenting people under occlusion. His approach mainly utilises AND/OR networks to track multiple people's hands and feet. The speaker demonstrated their approach on the HumanEva dataset, as well as on outdoor sequences in which multiple people interact and partially occlude each other. Following the plenary talk, there was the third oral session entitled *Motion Analysis and Tracking*. The poster session, *Stereo, Motion Analysis and Tracking*, ended the conference of the first day.

On the second day, Quan presented a plenary talk entitled "Image-based Modeling". He showed his group's recent work on using the uncalibrated framework and the quasi-dense approach for 3D reconstructions of outdoor scenes. More detail of the work can be found in his new book called *Image-based Modeling* which will be published by the end of this year. Furthermore, there were three oral sessions with the topics of segmentation, feature extraction, object detection, image enhancement and visual attention. Later, a poster session was held with the same topics.

The last day saw the last plenary talk, "Acquisition of 3D shape from uncalibrated video", presented by Professor Cipolla from Cambridge. The last day consisted of three oral sessions: *Machine Learning*

*Algorithm for Vision, Object Categorization and Face Recognition* and *Biometrics and Surveillance*. In addition, the last poster session was held on the topic of machine learning, recognition, biometrics and surveillance.

In general, this year's ACCV was successful. Some presentations, in particular the plenary talk, are worth studying or discussing in the future. The proceedings of the conference will be published online by Springer. Moreover, the conference was held in a nice place, Xi'an. Xi'an was the capital of 13 dynasties and has 3,100 years of history. The city is also famous for the location of the Terracotta Army from the Qin Dynasty.

Finally, I would like to thank BMVA for awarding me a student bursary, which gave me chance to attend this great conference.

Mai Xu  
Imperial College London  
email: [mai.xu06@imperial.ac.uk](mailto:mai.xu06@imperial.ac.uk)

## Report on Articulated Human Motion

This one-day BMVA Technical Meeting was held on 23 October at the BCS in London and was chaired by Dimitrios Makris (Kingston University) and Aphrodite Galata (University of Manchester). The meeting concentrated on human motion recovery using computer vision techniques. A variety of areas involved in this arduous task were explored, such as detection, tracking and modelling. Thus, in the morning, oral presentations were mainly focused on human detection using low-level features, whereas in the afternoon the session went towards dimensionality reduction methods as the way to model human behaviour, which represents a significant trend in human modelling over the last few years.

The first speaker, Ben Daubney (University of Bristol), described a method for human pose estimation and tracking using sparse and noisy motion features. In his approach a model of human gait allows posture estimation in a monocular scenario using a Pictorial Structure method. Then, Jean-Christophe Nebel (Kingston University) presented a method for estimating human pose from an isolated single frame aiming at initialising articulated tracking algorithms. His approach was based on iterative fitting of a simple articulated model over a spatio-temporal feature clustering using multiple image cues. The last morning speaker was Sam Johnson (University of Leeds) who proposed a



probabilistic algorithm for pose estimation under unconstrained natural images. This method combined HoG descriptor and colour cues in order to detect the limbs, also using a Pictorial Structure Model approach.

After lunch, Jesus Martinez del Rincon introduced a shape-skeleton tracking approach developed at the University of Zaragoza. The proposal tackles the problem of human pose recovery on video surveillance sequences and defends the advantages of 2D models such as the easy initialisation in uncalibrated environments. It consists of deterministic (Active shape models) and stochastic (Particle filter) techniques in a unique framework with a feedback mechanism between them. The next talk was given by Marco Paladini (Queen Mary) and promoted the use of a framework for recovering 3D shape and motion from uncalibrated 2D image measurements using an iterative factorisation approach. The proposed model is applied to the three different categories in which an object can be classified – rigid, deformable and/or articulated.

Spectral dimensionality reduction methods were the focus of the following two speakers. Fabio Cuzzolin discussed the advantages of using locally linear embedding (LLE) for representing 3D volumes with articulated structure and the soundness of this method against similar methods such as Isomap. Then, Michal Lewandowski (Kingston University) discussed how spectral dimensionality methods are used for modelling human activities. In particular, he showed a new variation of the classical LE algorithm that introduces spatio-temporal neighbour constraints during the creation of the dimensionality reduced space.

The last two talks of the day focused on 3D pose recovery. Aphrodite Galata (University of Manchester) described a tracking framework for calibrated multi-camera scenarios based on learning complex activities. This learning is made over dimensionality reduced spaces that allow the decomposition of complex motion patterns into a vocabulary. Finally, John Darby (Manchester Metropolitan University) described a tracking proposal based on a particle filter, which combines generative and discriminative approaches to address both known and unknown activities.

In conclusion, this meeting has revealed intense and diverse activity in this important field of computer vision. Although convincing results have been demonstrated in specific scenarios, the estimation of articulated human motion remains a very challenging problem for our community. Special attention should be paid to deal with real scenarios, multiple targets and unconstrained motion patterns. There is no doubt that fundamental progress will be achieved in the near future

and will provide exciting material for a new BMVA Technical Meeting on this same topic.

Dr Jesus Martinez del Rincon  
Kingston University  
email: [jesus.martinezdelrincon@kingston.ac.uk](mailto:jesus.martinezdelrincon@kingston.ac.uk)

## Robotics and Vision – Call for Participation

[www.bmva.ac.uk/meetings](http://www.bmva.ac.uk/meetings)

One-day BMVA symposium to be held on 24 March 2010 in London

Chairs: John Illingworth (Surrey University) and Adrian Clark (Essex University)

Amongst the initial inspirations of AI researchers was the challenge of producing autonomous robots and seeing machines. We have since learnt that the problems encountered are far more difficult than could have been imagined. The enormity of the problem has led to specialisation of function and expertise into sub-disciplines, of which robotics and computer-vision are two prime examples. Thus, two related but often distinct communities of researchers have grown up. The purpose of this meeting will be to bring together practitioners from both communities and showcase work where there is overlap of interests, thereby encouraging future multi-disciplinary collaboration. This meeting is the third in a series of biennial meetings on this topic

Papers are sought that relate to the intelligent integration of cameras, computers and actuating machines in challenging environments. Topics that would be of interest include, but are not restricted to:

- Visually guided mobile robots
- Medical and surgical robotics
- Industrial service robots
- Cognitive robots
- Household robots
- Computational and cognitive theories of robots
- Humanoid and toy robots

Please submit an extended summary of about one page A4 (max two pages), PDF preferred, including links or pointers to web-based illustrations, demonstration material or papers giving more details. Papers should be submitted by email attachment to Professor John Illingworth by 17:00 on Monday 8 February 2010.

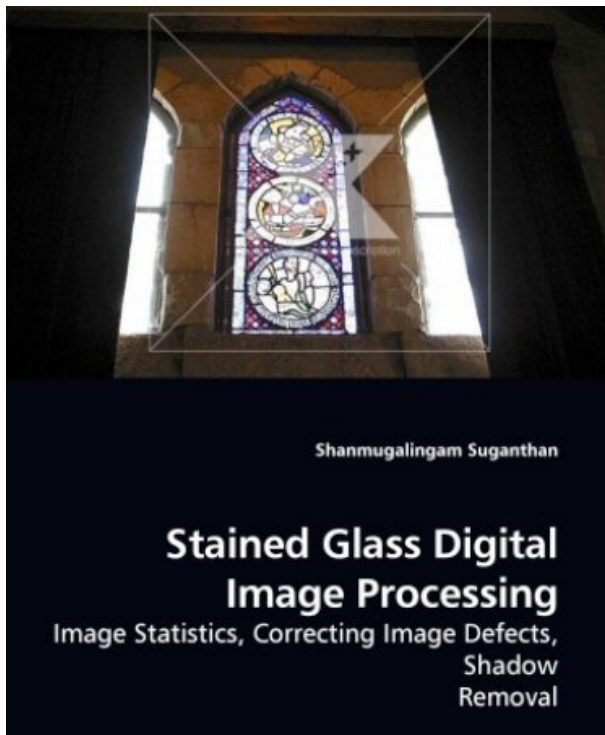
Informal enquiries about the meeting can be directed to John Illingworth by email or phone on 01483 689835.

Professor John Illingworth  
University of Surrey  
j.illingworth@surrey.ac.uk

## Book Announcements

The second edition of the book *Shape Analysis and Classification: Theory and Practice* by Luciano da Fontoura Costa and Roberto M. Cesar Jr. has recently been published by CRC Press. It contains 685 pages; its ISBN numbers are ISBN-10: 0-84937-929-6, ISBN-13: 978-0-84937-929-1. More information is available on the following website:

<http://www.vision.ime.usp.br/~cesar/shape/>



Publisher: VDM Verlag Dr. Müller (26 Aug 2009)  
Paperback: 204 pages, price: £69, ISBN numbers:  
ISBN-10: 3-639-18486-6, ISBN-13: 978-3-639-18486-0

Stained glass has been an important medium for the illustration of European culture and history and is still used today in architecture and the decorative arts. Heritage conservators are interested in using digital

images in the analysis of stained glasses. Stained glass images differ from natural images in many aspects. Their colour is generated by transmitted, rather than reflected light. In many cases the background is partially visible through the glass, typically showing trees, sky, etc. Images of stained glass taken with external illumination very often contain shadows, cast by structures such as protective bars and grilles. The physical structures producing the shadows are often irremovable, because they are difficult to access or constitute structural elements of the window. It is thus necessary to provide a suitable set of image processing tools to remove background and shadows from the digital images. This book introduces systematic procedures to calculate statistics of stained glass on digital images and develops novel algorithms for segmentation, feature extraction, and removal of image defects including non-uniform illumination, shadows of bars and grills and background.

*Note to potential reviewers of this book:* this particular publisher operates by giving the reviewer an e-book rather than a hard copy. Please contact me and I will request the e-book for you from the publisher.

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

## Books for Review

The following books are on offer for review. As always, they will be sent out on a first come–first served basis, so please contact me immediately if you would like to review either of them. (If you are doubtful, go for the *quick view* option, and then return the book to me if you feel you are unable to do a review.)

*Handbook of Texture Analysis*, edited by Majid Mirmehdi, Xianghua Xie and Jasjit Suri, published by World Scientific in 2008. It has 424 pages and its ISBN numbers are ISBN-10: 1-84816-115-8, ISBN-13: 978-1-84816-115-3.

*Statistical Models of Shape: Optimisation and Evaluation*, by Rodri Davies, Carole Twining and Chris Taylor, published by Springer-Verlag in 2008. It has 314 pages and 57 Figures, costs £55, and its ISBN-13 is: 978-1-84800-137-4.

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