Newsletter of the
BRITISH MACHINE VISION ASSOCIATION
AND SOCIETY FOR PATTERN RECOGNITION.
Vol 2, No 3, August 1991.

Editor: Dr Edwin Hancock, Dept. of Computer Science, University of York, York. Y01 5DD.
Tel: (0904) 433374, Fax: (0904) 432767, email: erh@uk.ac.york.minster

EDITORS NOTE

The main items this month are a book review by Professor Vicki Bruce of Nottingham University,an article on research into robot heads by Bernard Buxton, a review of the recent BMVA meeting on robot vision for vehicles, also by Bernard Buxton, and, a report on the IAPR Visual Form workshop from Mark Wright of Cambridge University.

There are deadlines for two major conferences approaching. The 11^{th} ICPR requires extended abstracts by 30 October and the second ECCV requires full-form papers by 15 October. Details of the ICPR were given in the last newsletter; details of ECCV are included in this newsletter.

European Heads

It is hard to believe that it is now almost three years ago that Dana Ballard gave his invited talk on "Active Vision" at the Manchester AVC where he described research on the Rochester "head". Since then, several papers on controllable stereo mounts have appeared in leading conferences such as the ICCV. For the most part these have featured other US work although it has been well known on the "grape-vine" that several European groups, including the Turing Institute, the Universities of Oxford and Sheffield in the UK, INRIA in France, Genoa DIST in Italy, Aalborg in Denmark, LIFIA at Grenoble, and the Royal Institute of Technology at Stockholm, have all been interested in building heads. The Turing head, called "Richard", is well known for its commentary on the Robot Olympics held at the Turing Institute last September and will feature in this year's BMVC at Glasgow. The four degree of freedom Sheffield head featured in Mayhew's closing talk at the "Vision for Vehicles" meeting and is described later in this newsletter. I have seen the Sheffield head many times on visits to the AIVRU but now, on a recent trip to Sweden, I have also had the opportunity to see the RIT head. This thirteen degree of freedom head, developed within basic research programmes in computer vision and image processing at RIT supported by the National Swedish Board for Technical Development and in the ESPRIT-BRA project, BR-3038, "Vision as Process", is very different from its predecessors.

In particular, the RIT head was designed to have almost every degree of freedom the vision researcher could wish for and, in a manner similar to human vision, to have independent eye movements **about the lens' centres** with **off-centre** "neck" axes for motion parallax and other exploratory, active visual movements. Thus, of the sixteen possible degrees of freedom that could be on the vision researcher's wish list, the RIT head has neck, pan and tilt axes and, for each camera or eye: vergence and tilt, focus, zoom and iris diaphragm. The thirteenth degree of freedom is a change of baseline, whilst the mechanical rig has provision for extra longitudinal motion of each camera so that the vergence and tilt are always about the lens centre whatever the focus or zoom setting. The three missing degrees of freedom are the neck roll and the cyclotorsion of the eye which is easily implemented by software.

Each of the geometric axes is driven by a stepper motor through a 50:1 ratio harmonic (backlash free) gear giving high precision ($\sim$ 28 seconds of arc) and speeds of the order of $\sim$ 180 degrees per second. The use of stepper motors and the fact that the eyes can be independently rotated without affecting the optical properties of the system (except of course for controlling the direction of the cameras' optical axes and, if the optical axes intersect, the fixation point) should simplify the design of a complete control system but could compromise the precision and stability of the system. For stability and very high precision, analogue control and DC drives are generally superior. In particular, the gearing required to give the stepper motors positional precision may limit their ability to respond to external disturbances.

The head was demonstrated to a colleague and me during our visit to RIT. In particular we saw (and interacted with) it fixating on a point in space, tracking (for example) a human hand as it was slowly moved, stabilizing on a point in space as the two neck degrees of freedom cycled through a pre-programmed sequence of movements and detecting and executing a saccade to a sudden movement in the laboratory such as a person entering a doorway. Although these competences were as yet only implemented separately, all this was achieved by means of simple image processing algorithms (essentially — correlation and image differencing operating on a small region of interest $\sim$ 17x17 pixels) implemented on a single T800 transputer. The system was controlled interactively from a menu of simple buttons on the UNIX host.

Already the RIT head displays a fascinating range of visual behaviours and is a credit to Kourosh Pahlavan, its principal designer and builder, and to the other members of the Computational Vision and Active Perception Laboratory in Stockholm involved in its development. It will be very interesting to compare this head with the Turing and Sheffield systems and the others now being built in Europe, in particular the high-performance, stabilized head being built by SAGEM, GEC, Oxford and INRIA in ESPRIT project P5390. In addition, the successful development of the RIT head will no doubt spur on the rest of the European competition and we will, I hope, see several papers and perhaps even a few "heads" themselves at the second ECCV to be held in Genoa in May 1992.

Bernard Buxton
GEC-Marconi, Hirst Research Centre

ECCV 92

This conference will take place in Santa Margherita Ligure near Genova between 18 and 23 May 1992. The conference will be a single track event sandwiched between an ESPRIT day on computer vision and a Basic Research Action Workshop. Papers can be submitted as long contributions (25 double spaced pages) or as short contributions (14 double spaced pages). They should reach Professor Giulio Sandini, DIST University of Genova, via Opera Pia 11 A, 16145 Genova, Italy not later than 15 October 1991.

IAPR Workshop on Visual Form

The International Workshop on Visual Form was held from 27th to 30th May 91 on Capri. In his opening address the Chairman, Prof L.Cordella from DIS-Universita' di Napoli, said the aim was to provide a workshop with a specific focus on the area of shape and the central role it plays in computer vision and to provide some indication of the current "state of the art" in this important field. The list of participants included 88 names from 12 countries including 53 Europeans, 22 North Americans and smaller groups from Israel and Japan.

Main themes included 3D object recognition, the extraction of 3D shape from input data and 3D modelling using primitives such as geons and superquadrics. Hierarchical shape description including scale space representations featured prominently as did morphological operations, invariants for shape matching and skeletonisation and symmetry detection using a variety of techniques.

Levine and Rosenfeld both talked about geons as 3D primitives, this was complimented in the talk by Metaxas on deformable superquadrics. R.Haralick showed that morphological techniques could be extended to cope with appreciable levels of noise by introducing the morphological equivalent of a Wiener filter. An interesting and at times amusing talk was given by T.Huang on the modeling of human face motion for model-based video compression. L.Davies discussed massively parallel algorithms for shape analysis on the Connection Machine and M.Leyton expounded his theory on the inference of causal history from shape. After his talk on splines Pavlidis suggested "hacking" can be a useful approach to certain vision problems which have proved intractable to more formal analysis.

Perhaps one of the best received talks was by J.Eklundh. This was entitled "Qualitative Shape: Some Computational Aspects". The main theme of this talk was that the important information which needs to be extracted from an image is qualitative rather than precisely quantitative. To extract this information it was suggested that features must be related to a particular scale within a complete scale space representation and that indeed object descriptions are only meaningful over a certain scale interval. In taking this approach it was stressed that the figure/ground problem can not be assumed to have been solved separately and is implicit in the segmentation of the scale space representation.

J.Toriwaki gave a summary of distance transformation and skeletonisation techniques for shape feature analysis. This talk was the catalyst for an interesting debate from the conference floor. Haralick suggested that few precise models exist for the skeletonisation process but this view was countered by Arcelli and Pavlidis. This lead to constructive discussions in the interval where a number of people agreed to work on the problem afterwards. Haralick has since produced a draft proposal for a framework in which different skeletonisation algorithms can be compared in an objective and quantitative manner.

The workshop provided a useful summary of work on shape and perhaps further evidence of the growing strength of the European research community. The proceedings of the workshop are to be published as a book by Plenum Press.

Mark Wright
Cambridge University Engineering Department

Vision Systems for Robot Vehicles

This one-day meeting held at the GEC-Marconi Hirst Research Centre on Wednesday, 22 May, departed from the style now established for the BMVA short meetings and workshops. First, as for the meeting on "Mammographic Image Analysis" held at the IBM Scientific Centre in January, an industrial participant acted as host and gave a laboratory tour and demonstration at the end of the day. Second, their participation ensured that, in addition to the audience of approximately fifty BMVA members and the ten speakers, there were about ten to fifteen members of the company also present. Finally, four of the speakers came from mainland Europe and considerably enriched the variety of the talks and broadened the scope of the meeting.

Indeed, the meeting began with talks by two of the visitors from across the channel. Fabrizio Ferrari from the University of Genoa, DIST and Ernst Dickmanns from the Universtät der Bundeswehr, München. In his talk, Ferrari described the architecture of the vehicle and vision system established at DIST in the *VOILA* ESPRIT project. This is a three layer Brooksian architecture in which the lowest level, which was the subject of the talk, is a vision based obstacle detection and avoidance system that works by using stereo vision to detect obstacles lying above or below the vehicle's ground plane. An impressive video of the system in operation with the vehicle manoeuvring down a narrow corridor at DIST was shown and the remainder of the talk dedicated to describing initial results on a quantitative assessment of the performance of the system. In his talk, which followed, Dickmann's again began by addressing the architecture of his 4D approach to dynamic vision. The basic principles underlying this system which is being applied to vehicle guidance problems on the *PROMETHEUS* Eureka project are the use of spatio-temporal models of the world and maximal utilization of knowledge of the vehicle's dynamics, especially its relevant degrees of freedom. Recursive filters are then used to detect the centre or the edges of the road, obstacles, other vehicles etc. and implemented on dedicated hardware. The lower levels of the system thus use only very simple image processing and are able to run at 25 Hz, although the higher, situation assessment, levels which utilize the world model containing the road, other vehicles etc. are much slower and operate on timescales of up to a few seconds. Again, a video of the system was shown and several road driving behaviours were discussed including: lane following, lane changing, stopping, convoying and drawing in to the kerb. A particularly notable feature was that a simulator had been built so that the system could be tested in the laboratory on realistic image data before hardware was constructed and equipment mounted on the road vehicle. Also notable was the fact that this equipment included two movable cameras, a wide angle camera for near parts of the scene and a zoom camera for looking far ahead when the vehicle is moving fast at speeds of 50 km/hour or more.

The remainder of the morning was then taken up with two talks associated with work at the Hirst Research Centre. First, David Castelow described the vision-vehicle system at the HRC which, like the Genoa system, is one of the four experimental platforms developed in the *VOILA* project. Again, a brief video of the system was shown featuring the operation of three modules for object recognition and location, tracking and free space determination. Of these, the object recognition and location and free-space determination both use full frame (512x512x8 bit) stereo images and take 12–15 seconds to run on the multi-transputer system at the HRC. The tracker, however, utilises model based prediction to process sub-regions of the images at about 3 Hz. Of particular interest were some very recent experiments on the accuracy of the vision system which, in some circumstances, can detect errors to a precision of a mm or two. This accuracy can be ten times greater than that of the vehicle's rotating laser scanner and barcode reader. Experiments of this kind on the free-space determination were the main subject of the laboratory demonstrations later in the day for the dozen or so people who had remained behind after the talks. In addition to these experiments, they were also shown the equipment and software facilities used in the HRC vision laboratory which includes the *TINA* stereo vision system and the *MARVIN* multi-transputer architecture for vision, both of which were developed on collaborative projects at the AI Vision Research Unit at the University of Sheffield. *MARVIN* itself was in fact the subject of the last talk before lunch by Mike Rygol of AIVRU. This talk began with a brief review of the original design considerations of *MARVIN*, especially the means of distributing image data to several transputers by means of the Datacube Maxbus and the TMAX (Transputer - Maxbus) cards, the latter developed on the project. A brief description of the C software infrastructure implemented on the machine was also included. This software features a message routing facility similar to that now being built in hardware on the T9000. Rygol also showed a video of the Sheffield *MARVIN* in operation with their vehicle, a converted wheelchair known affectionately as *COMODE*. As in the preceding talk, this included object recognition and location, tracking and free-space determination but also included extensions of these modules for "carrot-following" and docking.

After a buffet lunch, the programme was rearranged slightly, first to enable the chairman to show clips from Malcolm Roberts's video of a commercial caterpillar AGV system operating in a factory in Belgium. This showed AGVs equipped with a GEC rotating laser scanning system (the same as that on the HRC laboratory vehicle) operating both indoors and outdoors in a modern factory and stockyard. This set the tone for one kind of vehicle application in controlled, benign environments. In contrast, the second application talk by Phillippe Lemarquand showed the very rugged forestry and mining scenarios being addressed in the *PANORAMA* ESPRIT project. In this project, it is proposed to use a variety of inertial sensors, global positioning systems, laser range finders and passive vision. If it can provide the performance required, passive vision will be used in conjunction with a laser range finder to replan vehicle trajectories on a timescale of one second or less.

Presentation of these two application scenarios was then followed by a most entertaining talk by Patrick Stelmaszyk who introduced himself by asking rhetorically "what a Frenchman with a Polish name was doing addressing an English meeting on vision research in Japan". The answer is that Patrick had recently returned to France after working at Mazda Research in

Japan for approximately one year. He was thus able to present a first-hand account of vision research in Japan and backed it up with very detailed statistics on the latest trends in Japanese R & D. Amongst the most interesting aspects of this talk were Stelmaszyk's own impressions that Japan **was** but **no longer is** a paradise for technology, that the best research was now to be found in the large companies' laboratories and no longer in the government (MITI) laboratories and that, in spite of the fact that more than 150 different commercial vision machines or systems can be bought in Japan, there is **no** government programme of computer vision research there. Stelmaszyk suggested that this was because the Japanese **only** invest in large programmes on topics that are very certain to give an economic return.

Following this entertaining, but serious discussion, the meeting ended with three vision talks. Of these, the first two by Chris Harris and Ian Reid could be described as "regular" vision talks, whilst the last by John Mayhew as usual requires a different description. Chris Harris described the *DROID* system which is similar to that of Dickmann's — it uses a spatio-temporal scene model maintained by Kalman filtering operations — but is designed for unconstrained outdoor scenes and unconstrained camera ego-motion. Again, a video of the system was shown, the most impressive part of which showed a set of driveable regions being obtained as a vehicle was driven round a skid-pan. In fact *DROID* is the basis of yet another of the experimental platforms established in the *VOILA* project, this time at the University of Oxford. With the talk given by Eric Theron at the BMVA meeting on European Vision Projects in March, all four of the systems being developed within this project have now been presented to BMVA members. In addition, the efforts being made to assess the performance of the systems developed in this project, doubtless in part the result of the initiative begun by Patrick van Hove in his ESPRIT workshop at Antibes in April 1990, are notable. The penultimate talk by Ian Reid featured an active vision system, the NEL light striper mounted on a GEC "TURTLE" vehicle at Oxford being employed to recognize classes of objects such as "pallets". In this system which, in distinction to a conventional passive CCD camera, takes up to 10 seconds to collect a 512 column image, it is difficult to detect crease edges reliably, but planar surface patches can be extracted and their surface normals used in a Grimson/Lozano-Pérez, RAF-like system for object recognition.

The final talk by John Mayhew, though less conventional in its presentation, returned to the topic of passive vision but addressed the development of Brooksian layered architectures for visual control. In particular, Mayhew described his latest ideas on developing this kind of control system for the four degree of freedom stereo "head" mounted on the AIVRU vehicle. The four degrees of freedom here include the pan and tile axes for the "neck" and two independent vergence axes for the cameras. In addition to being designed as a layered architecture, this system was also being used to test an application of neural nets to control in that "it knows no more algebra or about adaptive control theory than Mayhew" and therefore has to learn its inverse kinematics and adaptive control from experience. In fact, when he got down to details, Mayhew proposed the neural net as the middle layer of the architecture with a PID controller below and a lattice filter above acting as a predictor. The lattice filter was described as being easier to compute than a Kalman filter and more robust to missing data.

After Mayhew's talk, the meeting closed for tea and the subsequent laboratory tour and demonstrations. Although, as the organizer of the meeting and chairman I may be biased, I felt it was a most enjoyable and successful day and would like to thank all those who helped make it so: the HRC staff for help with the organization, the vision group for the laboratory tour, and GEC for their generosity in hosting the meeting, but most of all the speakers, especially those from abroad who brought the essential elements of science, excitement and fun to the occasion.

Bernard Buxton
GEC-Marconi, Hirst Research Centre

BMVA JFIT Meeting

On 11th of July the BMVA organised a workshop on behalf of the DTI at the JFIT community conference in Manchester. As a PhD student of some six months this was my first opportunity to meet experts in the field of computer vision outside of my own department. So, unusually, the prospect of being made to feel like a poodle in the company of wolves was tinged with some excitement.

It wasn't like that though! Perhaps it was because of the time set aside for discussion, or because of the mixture of academics and business people, but I didn't feel out of my depth, or out of sight, instead I found myself taking sides, asking questions, and even expressing doubts!

The day started with coffee, a good opportunity to try to identify the various camps, the representatives of business with their suits and ties, and the academics without. I went and stood with the other postgrads, clearly identifiable by their silence. Following a few mumbled introductions it was time to start.

After a short introduction, the first of eight presentations by different research teams commenced. Whilst some groups presented outline descriptions of integrated vision systems, other talks concentrated on the minutae of particular topics. For computer scientists there was plenty of theory to wet the appetite, improved techniques for simulated annealing, parallel alternatives to the inherently serial region growing schemes, the use of genetic algorithms in medical imaging, and more. For those with commercial concerns, an opportunity to discuss the difficulties faced by small companies doing scientific research, lessons for systems integrators and a chance to lobby for more cash. With little exception the talks were interesting, and whilst wide ranging most of what was said had relevance for anyone working in the field.

To view the workshop as a series of presentations however would be an injustice. What characterised the meeting more than anything was the dialectic interplay between science and business. Through the day academics were softly criticised for being too theoretical, unpragmatic, and oblivious to the needs of business, whilst those engaged in research were able to make a strong empirical case that scientific and engineering research are different. The JFIT meeting provided a forum for discussing this question without animosity. If business and academia are to work in tandem, then meetings of this kind must play an increasingly important role. As a newcomer to scientific research I found this aspect of the workshop highly motivating.

Whilst the BMVA clearly did well to organise this meeting at such short notice, there was one small problem. Double booked with the workshop was another BMVA event, a talk about line detection by Dr. Maria Petrou. At 3.50 I left to run across Manchester, missing the final session. Stereotypically the heavens opened as I left the building.

Jon Guiton
School of Computer Studies, University of Leeds

Book Review

" Visual Agnosia: Disorders of object recognition and what they tell us about normal vision". Martha J Farah. MIT Press, 1990. ISBN 0-262-06135-X. Price £22.50.

One of the formative influences on the development of David Marr's theory of vision was the dissociations between different aspects of object perception reported by Elizabeth Warrington "..here was this young woman calmly telling us not only that her patients could convey to her that they had grasped the shape of things that she shown them, even though they could not name the objects or say how they were used, but also that they could happily continue to do so even if she made the task extremely difficult visually... It seemed clear that the intuitions of computer vision people were completely wrong and that even in difficult circumstances shapes could be determined by vision alone." (Marr, 1982, p.36). Here is a classic example of the use of abnormalities in normal visual object recognition to help us understand processes of normal vision and a nice pointer from Marr to the way that studying such deficits may complement insights about natural vision systems gained through the study of computer vision.

In the years since Marr drafted his book, the discipline of cognitive neuropsychology has developed rapidly. Cognitive neuropsychology is characterised by the careful investigation of the patterns of impairment in single cases of brain damage and the attempt to use such patterns as one source of evidence for cognitive modelling. Patients may be grouped at a later stage in terms of their cognitive deficits but (in contrast to more traditional neuropsychological studies) they are not grouped at study on the basis of site or nature of their neurological damage. This type of cognitive approach was developed earliest in the study of reading disorders but is now developing rapidly to supplement experimental and computational studies in the area of visual recognition. Farah's book, possibly the first monograph devoted to a review of the visual agnosias, is timely. The visual agnosias encompass a group of object recognition impairments in which patients have difficulties recognising objects visually but can be shown by other means to have intact non-visual representations of the objects. For example, visually agnosic patients are usually able to give a verbal definition of an object and describe its uses, even though they are unable to recognise the object visually. The patients are not blind-and some indeed may have apparently normal vision as assessed by their abilities to copy drawings that they are unable to recognise. Some (e.g. patient H.J.A. studied by Humphreys and Riddoch) may also be able to produce detailed drawings from memory of objects that they cannot recognise. Such counterintuitive dissociations between the ability to describe an item in words or in a picture and the ability to recognise the item, force us to think in detail about the functional architecture of the visual recognition system. Different agnosic patients have differing patterns of impairment and patients have typically been classified into different groupings as a result. For example, one type of visual agnosia is termed "prosopagnosia" where there is particular difficulty in recognising faces.

In her book, Farah argues that although cognitive neuropsychologists study single cases rather than groups based on areas of brain damage, their interpretations and resultant groupings of these patients into particular types is strongly influenced and ultimately limited by their prevailing theoretical models. She thus thinks it is essential to re-examine the visual agnosias purely empirically, in order to describe carefully the different perceptual and/or memorial deficits that may be observed to underly patients' failures to recognise objects. Her monograph presents a careful re-examination of the deficits affecting object perception and recognition, and she separates this review of perceptual symptoms from chapters in which she attempts fresh interpretations. In her book she challenges some traditional divisions, for example, that between the "apperceptive" versus the "associative" agnosias, traditionally distinguished according to whether the patients have perceptual problems or not. Farah suggested that ALL agnosics can be found to have perceptual problems, though arising from different levels in vision.She also challenges the idea that there may be agnosias specific to particular categories of objects (e.g. faces). Instead, she suggests that there may be two rather different computational tasks involved in object recognition-the representation as a series of separable parts (e.g. the letters of a word) or the representation of the parts themselves as whole units (e.g. the representation of a face). Brain damage may affect one or the other of these representational abilities selectively to produce apparently selective types of agnosia. Finally, Farah suggests that available evidence supports a parallel distributed processing architecture underlying these representational abilities and though she herself does not offer a well-specified new framework cast in such terms, she points the way towards one. Some of Farah's suggestions appear to be sensible reformulations of older ideas but others are provocative (e.g. the suggestion that all agnosias involve perceptual deficits) and it is particularly in these provoctive areas that the book will act as a stimulus to further research in this field.

This is a scholary text, which is essential reading for students and researchers in visual perception and cognition. Researchers in machine vision who wish to understand more about natural vision will find much of interest in this book, though it might be found a little hard going by those with no background at all in cognitive neuroscience as it does assume some familiarity with basic technology and methodology. Overall, I think the book makes an important contribution to a fascinating and rapidly developing field and though I disagree with some of Farah's analyses, I was delighted to have the excuse to read the book carefully enough to discover these contentious areas.

Vicki Bruce
Department of Psychology, University of Nottingham

Teaching Company Schemes

BMVA members may be interested to know of a funding scheme called TCS, which has as its aim the transfer of technology from academia to companies. The scheme has been in operation for over fifteen years and is currently funding just under four hundred programmes, of which two involve machine vision systems. It would seem likely that other members of BMVA would qualify for such funds.

In a TCS programme, a graduate (or group of graduates) work in a company, supervised by someone in the company together with an academic(s). It is via the academic supervision that knowledge and expertise becomes transferred to the company, which is an essential requirement of the scheme. In contrast with many sources of funding, the work has to be 'near market'. The work should be important for the company, or for a section if the company is large. TCS funding covers between 50 percent and 70 percent of the costs of a programme, and the company covers the rest. TCS money is paid via the academic institution in the form of an SERC grant. Some of the costs in the budget provide for support for the academic to release his or her time from other duties.

Anyone interested in learning more about the scheme should contact TCS consultant Dr N B Cryer tel 0483 740423, fax 0483 747181.

Future BMVA Programme.

The following is a list of BMVA meetings currently planned

- 23-26 September 1991 *BMVC91*
- 23 October 1991 *"Natural and Machine Vision"*
- 11 December 1991 *"Rigorous Neural Nets"*

The majority of meetings will be in Central London and are scheduled as one-day events. No registration fee is payable by BMVA members although a charge of £5 will be levied for non-members (unless they are members of a co-sponsoring organisation). Fuller details including final programme will be mailed to members nearer the time of each meeting.

DIARY

13-16 August 1991 7th Scandinavian Conference on Image Analysis, Aalborg, Denmark
4-6 September 1991 6th Int Conf on Image Analysis and Processing, Como, Italy
23-26 September 1991 2nd British Machine Vision Conference, Glasgow
18-23 May 1992 2nd European Conference on Computer Vision, Italy
August 30 - October 3 1992 11^{th} International Conference on Pattern Recognition, The Hague, Netherlands