12th Vision Researchers Colloquium

Thursday 2 July 2020

ONLINE

Keynote speaker

Professor David Perrett, School of Psychology and Neuroscience, University of St Andrews

In partnership with:

Bristol Vision Institute's

Online Vision Researchers Colloquium 2020

Hosted via Zoom, supported by Slack.
Online access provided once registered via Eventbrite.

Session 1	
09.00	**Welcome by Professor David Bull, Director of Bristol Vision Institute, University of Bristol (to include house-keeping, code of conduct)**
	Chair: Dr Eleanor Caves, Marie Skłodowska-Curie Research Fellow, University of Exeter
Presentations – 15 mins, plus 5 mins Q&A each	
09.10	**Catherine Taylor**, University of Bath, "An egocentric pipeline for transporting physical objects into virtual environments"
09.30	**Doug Sands**, University of Bristol, "Countershading for concealment against multiple backgrounds"
09.50	**Martin Bossard**, Cardiff University, "Is allocentric and egocentric visual information used during ongoing reaching movement?"
Lightning presentations – 5 mins each	
10.10	**Daria Burtan**, University of Bristol, "Visual discomfort, not image statistics, affect gait kinematics"
10.15	**Jing Gao,** University of Bristol, "Rotated bounding box-based CNN for Holstein Friesian cattle detection"
10.20	**Hugo Hammond,** University of Bristol, "Developing continuous multidimensional measurements of audience immersion"
10.25	**Di Ma**, University of Bristol, "BVI-DVC: A training database for deep video compression"
10.30	**Breakout session – all presenters from this session to 'host' a Zoom breakout room. Questions can also be asked at any time in the Slack channel.**
11.00	**15 minutes break**

Session 2	
Presentations – 15 mins, plus 5 mins Q&A each	
11.15	**Welcome back/ Chair: Marek Pedziwiatr, School of Psychology, Cardiff University**
11.20	**Karin Kjernsmo**, University of Bristol, "Do bees like shiny things?"
11.40	**Christian Wallis**, Cardiff University, "Replicability and patterns of individual differences across four popular eye-tracking tasks"
12.00	**Milton Montero**, University of Bristol, "Deep generative models as perceptual front-ends for decision-making"
Lightning presentations – 5 mins each	
12.20	**Angeliki Katsenou**, University of Bristol, "Predicting the rate-quality convex hull across resolutions for adaptive video streaming"
12.25	**Mubaraka Muchhala**, University of Bristol, "The geometry of high-level colour space"
12.30	**Aaron Zhang,** University of Bristol, "Enhancing video compression through deep learning"
12.35	**Breakout session – all presenters from this session to 'host' a Zoom breakout room. Questions can also be asked at any time in the Slack channel.**

13.00	Lunch and break

Session 3	
13.55	**Welcome back/ Chair: Dr Annabelle Redfern, School of Psychological Science, University of Bristol**
Presentations – 15 mins, plus 5 mins Q&A each	
14.00	**Dunia Gonzalez,** University of Exeter, “Colour and motion vision in stomatopod crustaceans”
14.20	**Sandra Winters**, University of Bristol, “Simulating the evolution of inter-specific mating signal diversity in guenons”
14.40	**Jake Deane**, CAMERA - University of Bath, “DogDetector: animal segmentation using synthetic images and 3D generative models”
15.00	**Breakout session – all presenters from this session to ‘host’ a Zoom breakout room. Questions can also be asked at any time in the Slack channel.**
15.15	**10 minutes break**
15.25	**Introduction by Professor Iain Gilchrist, Professor of Neuropsychology, University of Bristol**
15.30	**Keynote: Professor David Perrett**, School of Psychology and Neuroscience, University of St Andrews, “Face and Body Perception”
16:30	**Wrap up – Professor Jonathan Erichsen, Professor of Visual Neuroscience, Cardiff University**
16:35	**Informal networking on Zoom**
17:00	**Event ends**

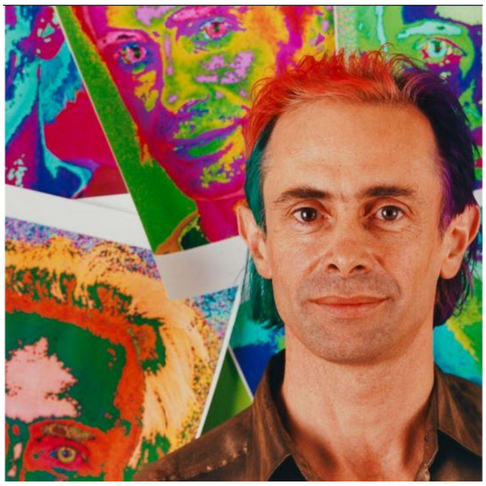
Before, during and after the event, please join in the discussions, share your photos and thoughts by linking to the Bristol Vision Institute Twitter account: @VisionInstitute.

Please contact BVI if you have any questions, on bvi-enquiries@bristol.ac.uk.

Keynote: “Face and Body Perception”

Professor David Perrett, School of Psychology and Neuroscience, University of St Andrews

Abstract



We readily make judgments of others but are unaware of the accuracy or the basis of our opinions. In this keynote, Professor Perrett will focus on the roles of skin tone, face shape and body shape in forming our impressions.

Carotenoid pigments from fruit and vegetables in our diet impart yellowness (or a 'golden glow') to the skin. Carotenoid ornaments are used in many species to signal general health. In humans too, we find that carotenoid skin colour reflects multiple aspects of health, including aerobic fitness, body fat, stress, and sleep duration.

The neural coding of face shape is such that exaggeration along dimensions to which brain cells are tuned increases cell responses. This coding may explain the improvement in speed of recognition when facial cues are caricatured. The coding may also enable the evolution of exaggerated features arising from sexual preferences (e.g., for enhanced sexual dimorphism) or from competition between individuals of the same sex. Both men and women have an exaggerated impression of the body shape desired by the other sex. Such misperception may be driven by competition.

Perceptual judgments about health and attractiveness depend on underlying biological processes: dietary effects on skin colour, hormonal influences on facial growth, and fat and muscle contributions to body shape. Perception is thus based on cues to underlying biology.

Biography

BSc. Psychology, St Andrews (1976), D.Phil., University of Oxford (1981)

Professor Perrett's doctoral studies were on single neurons and highlights included selective responses to faces and to familiarity. He set up a neurophysiology laboratory in St Andrews to develop these studies making discoveries of visual coding for faces, face identity, expression and attention direction, body movements including the biological motion of point light displays, and hand actions (the visual side of ‘mirror neurons’).

With his colleagues, he devised computer graphic programs for automated caricatures and for manipulating facial shape, colour and texture. They used these to study visual aspects of attractiveness, individual recognition, aging, personality, expression of emotion, and most recently health.

Find out more about Professor David Perrett:
Professor Perrett's latest research
Research led by the University of St Andrews: 'Healthy living gives skin a golden glow'
David Perrett at TEDxGhent: 'In your face'

Session 1

Catherine Taylor, University of Bath, *"An egocentric pipeline for transporting physical objects into virtual environments"*

The feeling of immersion in a virtual reality experience is greatly influenced by the way a participant interacts with the computer-generated (CG) environment. Controllers have been the traditional tool for facilitating interaction, however, these offer limited tactile feedback and do not accurately model real-world interactions.

In contrast, we propose an intuitive and immersive interaction mechanism where the behaviour of physical objects - captured using a robust tracking algorithm – are used to control virtual objects. Recent works have shown the power of neural networks for tracking rigid and non-rigid objects [1-4]. However, many state-of-the-art works are restricted to only tracking rigid objects [2,4] or require large amounts of manually labelled training data [1,3], which is time consuming to obtain. In our work, we use a synthetic dataset generation algorithm which creates large, automatically labelled datasets and use this to train a CNN based architecture.

The architecture – known as VRProp-Net+ [5] – predicts rigid and non-rigid model parameters from RGB images and the predicted parameters used to update the behaviour of a corresponding CG model. The RGB images of the physical object are captured from a moving egocentric viewpoint (i.e. a camera attached to a VR HMD). The moving camera allows a dynamic capture volume which is not restricted to a preselected area defined by a motion capture system or multi-camera rig. Previous egocentric tracking approaches have focused on hand pose or have explored the interaction between hands and rigid objects [6-8]. We extend upon this and consider both rigid and non-rigid objects.

Taylor, C., Cosker, D.

[1] A. Kanazawa, M. J. Black, D. W. Jacobs, and J. Malik. End-to-end recovery of human shape and pose. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7122–7131,2018.

[2] Y. Xiang, T. Schmidt, V. Narayanan, and D. Fox. Posecnn: A convolutional neural network for 6d object pose estimation in cluttered scenes. arXiv preprint arXiv:1711.00199, 2017.

[3] S. Zuffi, A. Kanazawa, T. Berger-Wolf, and M. J. Black. Three-d safari: Learning to estimate zebra pose, shape, and texture from images" in the wild". In International Conference on Computer Vision (ICCV), Oct. 2019.

[4] M. Andrychowicz, B. Baker, M. Chociej, R. J´ozefowicz, B. McGrew, J.W. Pachocki, A. Petron, M. Plappert, G. Powell, A. Ray, J. Schneider, S. Sidor, J. Tobin, P. Welinder, L. Weng, and W. Zaremba. Learning dexterous in-hand manipulation. CoRR, abs/1808.00177, 2018.

[5] C. Taylor, R. McNicholas, and D. Cosker, Towards An Egocentric Framework for Rigid and Articulated Object Tracking in Virtual Reality. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (pp. 354-359), 2020

[6] F. Mueller, D. Mehta, O. Sotnychenko, S. Sridhar, D. Casas, and C. Theobalt. Real-time hand tracking under occlusion from an egocentric rgb-d sensor. In Proceedings of the IEEE International Conference on Computer Vision, pp. 1284–1293, 2017.

[7] R. Pandey, P. Pidlypenskyi, S. Yang, and C. Kaeser-Chen. Efficient 6-dof tracking of handheld objects from an egocentric viewpoint. In Proceedings of the European Conference on Computer Vision (ECCV), pp. 416–431, 2018.

[8] B. Tekin, F. Bogo, and M. Pollefeys. H+ o: Unified egocentric recognition of 3d hand-object poses and interactions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4511–4520, 2019

Doug Sands, University of Bristol, *"Countershading for concealment against multiple backgrounds"*

Many animals are darker on their backs than their undersides, but the function of such 'countershading' has received little experimental investigation despite being proposed as a form of camouflage over 100 years ago (Poulton 1890; Thayer 1896). One hypothesis often proposed to explain the countershading of seabirds, and the motivation for colour schemes applied to many aircraft in WWII, is that different backgrounds need to be matched for different viewers: the back is seen from above against dark ground, but the underside is viewed against a light sky.

I used field photography of controlled stimuli, computer-based image analysis and detection experiments to investigate how white belly plumage in seabirds could function as hunting camouflage, and under what environmental conditions would it be most cryptic when viewed from underwater. Under sunny conditions, 79% of images had targets with low visibility against the sky, compared to 0.3% when it was cloudy. More than 90% of images taken when it was sunny and windy had low visibility targets. Given that no back-lit object can be as bright as its background (Penacchio et al. 2015), it was surprising that in 23% of the images, the target was undetectable. Experiments using the photorealistic 3D gaming software Unreal Engine confirmed these findings: direct sunlight causes glare, reducing detectability, and higher wind speed causes increased surface rippling and wave height, affecting refraction at the water surface and leading to distortion of the above-water image.

Sands, D., Scott-Samuel, N. E., Cuthill, I. C.

Penacchio, O., Lovell, P., Cuthill, I., Ruxton, G. & Harris, J. (2015) Three-dimensional camouflage: exploiting photons to conceal form. American Naturalist 186, 553-563. Poulton, E. B. (1890) The Colours of Animals: Their Meaning and Use. Especially Considered in the Case of Insects. Second Edition, London, Kegan Paul, Trench Trübner, & Co. Ltd. Thayer, A. H. (1896) The law which underlies protective coloration. Auk 13, 124-129.

Martin Bossard, Cardiff University, *"Is allocentric and egocentric visual information used during ongoing reaching movement?"*

In everyday life humans and most other animals need to reach targets for various reasons and in different ways. The information used to achieve such tasks is still debated in the literature. While some studies stipulate that individuals only use information from the relationship linking them to their environment (egocentric), others have shown that relationships between objects in the environment (allocentric) are also used when participants have to reach a previously seen target.

In two experiments, we developed a task in which relying on allocentric information was beneficial for successfully intercepting a moving target that remained visible. With a cursor controlled by their finger on a horizontal table, participants had to intercept the target as it moved across a textured background projected in front of them. During each attempt either the target, cursor, or background were perturbed individually, or all of them were perturbed simultaneously causing no relative changes between these three objects and consequently no change in allocentric information. Relying on allocentric information would therefore not elicit any response to the perturbation, whereas relying on egocentric information would give a response that resembles the combined responses to the three isolated perturbations.

Despite the difference between where and how the finger and the cursor moved, our results show that participants responded in accordance with the individual perturbations. This was even so when the simultaneous perturbation was repeated many times, suggesting that allocentric spatial information cannot be used to control ongoing visually guided actions.

Bossard, M., Crowe, E.M., Brenner, E.

Daria Burtan, University of Bristol, *"Visual discomfort, not image statistics, affect gait kinematics"*

Interacting with urban environments requires more attentional resources and thus cognitive processing load than interacting with nature (Kaplan, 1995), an effect that can even be found whilst looking at images (Berman et al., 2008). We recently showed that this is reflected in a change in gait kinematics: walking towards urban images as compared to nature images decreased walking speed and step length (Joyce & Leonards, 2017).

However, it remains unclear what exactly causes the slowing of gait: differences in aesthetics preference between the two image categories, visual discomfort, or differences in basic image statistics? In line with evidence that nature images contain higher amounts of fractals than urban images (e.g. Ho et al., 2019) and high-fractal content increase perceptual fluency (Joye et al., 2016), we conducted two studies to establish whether the decrease in fractal dimensions increases cognitive load, affecting gait kinematics.

Participants were asked to walk towards abstract images, which were parametrically varied in their fractal content. The task was to rate each image after the walk for visual discomfort (study 1) or for their likability (study 2). In the first study (n=20) visual discomfort, but not fractal dimension predicted walking speed and step length. In the second study (n=19) neither fractal dimension nor aesthetics were predictive of gait changes.

These data suggest that visual discomfort rather than image statistics or aesthetic preferences increase cognitive load, and thus affect gait kinematics.

Burtan, D., Spehar, B., Burn, J. & Leonards, U.

Amboni, M., Barone, P., Hausdorff, J.M. (2013) Cognitive Contributions to Gait and Falls: Evidence and Implications. Movement Disorders, 28(11), 1520–1533. Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. Psychological Science, 19(12), 1207–1212. Ho, S., Mohtadi, A., Daud, K., Leonards, U., Handy, T.C. 2019 Using smartphone accelerometry to assess the relationship between cognitive load and gait dynamics during outdoor walking. Sci Rep, 9(1), 3119. Joyce, K., & Leonards, U. (2017). Attention Restoration Theory in motion: Is gait impacted differently by visual exposure to natural and urban environments? European Conference on Visual Perception, 2017, Berlin. Joye, Y., Steg, L., Unal, A.B., Pals, R. (2016). When complex is easy on the mind: Internal repetition of visual information in complex objects is a source of perceptual fluency. J Exp Psychol Hum Percept Perform 42(1), 103-114. Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework. Journal of environmental psychology, 15(3), 169-182.

Jing Gao, University of Bristol, *"Rotated bounding box-based CNN for Holstein Friesian cattle detection"*

Current Holstein Friesian cattle detection systems using convolutional neural networks (CNN) are based on orthogonal (axis-aligned) bounding box detection [1,2]. In this work, we use rotated bounding boxes instead. Such approaches have previously been implemented in long object detection (e.g. ships [3,4]) to avoid multi-object alignment. The use of Rotated Bounding-boxes reduces the fraction of background pixels considered and consequently, the average precision (AP) of species (cow) detection increases.

This work is based on RetinaNet with a ResNet-50 backbone. We adapted its architecture by adding an angle parameter (0-359 degree) in addition to 4 coordinate parameters. 8 anchors are used per position with a ratio of 1:2.5, and each anchor has an angle of 45 degrees compared with its neighbouring anchors.

During training, we also use angle difference instead of intersection over union to speed up the filtering of positive anchors. We use 3872 and 1070 in-barn images from the University of Bristol's Wyndhurst farm for training and testing respectively. This dataset contains 4950 and 1253 individuals. With a non-maximum suppression (NMS) threshold of 0.28 rather than common 0.4-0.5, the AP is 98.8%.

Jing, G.

[1] Andrew, W., Greatwood, C. and Burghardt, T., 2019. Aerial Animal Biometrics: Individual Friesian Cattle Recovery and Visual Identification via an Autonomous UAV with Onboard Deep Inference. arXiv preprint arXiv:1907.05310. [2] Zin, T.T., Phyo, C.N., Tin, P., Hama, H. and Kobayashi, I., 2018, March. Image technology-based cow identification system using deep learning. In Proceedings of the International MultiConference of Engineers and Computer Scientists (Vol. 1, pp. 236-247). [3] Wang, Y., Zhang, Y., Zhang, Y., Zhao, L., Sun, X. and Guo, Z., 2019. SARD: Towards Scale-Aware Rotated Object Detection in Aerial Imagery. IEEE Access, 7, pp.173855-173865. [4] Xu, Y., Fu, M., Wang, Q., Wang, Y., Chen, K., Xia, G.S. and Bai, X., 2020. Gliding vertex on the horizontal bounding box for multi-oriented object detection. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Hugo Hammond, University of Bristol, *"Developing continuous multidimensional measurements of audience immersion"*

A wide range of media (be it a book, film, computer game, etc.) can result in immersion. It is often assumed that this is a result of the interplay between various cognitive processes. However, most research into immersion relies on post-viewing questionnaires which cannot tap directly into these cognitive processes or dynamically measure how immersion unfolds over the course of the interaction.

We have been developing a methodology for measuring immersion in a continuous, multidimensional way. Our dimensions (attention, emotion, memory) are collected using minimally invasive behavioural and physiological measures.

In this talk I will present data from an exploratory study which demonstrates the utility of these measures and looks at the interactions between them. We show that the individual measures are correlated across participants, suggesting they capture a shared response to the content. We also indicate the different measures converge at common time points in response to the content. However, the different measures are not equally correlated suggesting that they may be measuring different aspects of the immersive experience.

Hammond, H., Bull, D., & Gilchrist, I.

Di Ma, University of Bristol, *"BVI-DVC: A training database for deep video compression"*

Deep learning methods are increasingly being applied in the optimisation of video compression algorithms and can achieve significantly enhanced coding gains, compared to conventional approaches. Such approaches often employ Convolutional Neural Networks (CNNs) which are trained on databases with relatively limited content coverage.

In this work, a new extensive and representative video database, BVI-DVC, is presented for training CNN-based coding tools. BVI-DVC contains 800 sequences at various spatial resolutions from 270p to 2160p and has been evaluated on ten existing network architectures for four different coding tools. Experimental results show that the database produces significant improvements in terms of coding gains over three existing (commonly used) image/video training databases, for all tested CNN architectures under the same training and evaluation configurations.

Ma, D., Zhang, F. and Bull, D.R.

Session 2

Karin Kjernsmo, University of Bristol, *"Do bees like shiny things?"*

Gloss (specular reflection), to many vision scientists, is a nuisance: noise that corrupts the colour reflectance signal; to be avoided by suitable measurement geometry. However, gloss can convey information about surface properties, so should it be added to the visual properties, such as colour and shape, that flowers use to attract the attention of passing pollinators?

Indeed, gloss displayed by flowers is generally often limited to patterns, as opposed to fully covering the petals, which raises another question: if flowers are glossy to attract attention, why not be as glossy as possible?

In this talk, we will use data from experiments involving bumblebees and artificial flowers to show how varying levels of gloss affects shape recognition, choice and learning. Intriguingly, while we found that the bees could indeed learn to identify patterns of gloss, we also found that due to the inherent variability of the visual signal from completely glossy surfaces, it also has the potential to deceive and confuse pollinators. Therefore, floral advertising is a delicate balance between dazzling the customers and being recognisable, and plants seems to have evolved a balance designed to attract their most desired clients.

Kjernsmo, K., Harrap, M., Hall, J.R., Cuthill I.C., Nadia Khuzayim, N., Da Cunha Lobo, M., Doyle, C., Wainwright,B., Tálas, L., Rands, S.A., Scott-Samuel, N.E. & Whitney, H.M.

Christian Wallis, Cardiff University, *"Replicability and patterns of individual differences across four popular eye-tracking tasks"*

In order to facilitate stable visual perception, the brain must compensate for the effect of eye-movements on the retinal image. Extra-retinal signals which contain information about the eye's position and recent movements play a vital role in this stabilisation, and various eye-tracking tasks have been designed and posited to measure the use of this extra-retinal information.

This study chose four such tasks with the aim of investigating the replicability of earlier results while also exploring patterns of individual differences. Alongside the antisaccade task which has been posited as a possible measure of extra-retinal signals (Thakkar, K.N. and Rolfs, M., 2019. Disrupted corollary discharge in schizophrenia: Evidence from the oculomotor system. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*.), three other tasks were used that explore saccadic displacement, smooth visual pursuit, and saccadic shift.

Results so far show strong correlations between individual differences in the smooth visual pursuit and saccadic displacement tasks, $r(25) = .64$, $p < .001$, but no significant relationships emerge between any of the antisaccade results and the other tasks. Exploratory factor analysis similarly reveals three factors underlying inter-task correlations, around one of which clusters results from the smooth visual pursuit and saccadic displacement tasks. These results support the use of those two tasks in studies measuring extra-retinal signals but would otherwise exclude the antisaccade task as having any meaningful overlap in that role.

Wallis, CH

Milton Montero, University of Bristol, *"Deep generative models as perceptual front-ends for decision-making"*

Evidence integration models such as the Drift-diffusion model (DDM) are extremely successful in accounting for reaction time distributions and error rates in decision making. However, these models do not explain how evidence, represented by the drift, is extracted from the stimuli. Models of low-level vision, such as template-matching models, propose mechanisms by which evidence is generated but do not account for RT distributions.

We propose a model of the perceptual front-end, implemented as Deep Generative Model, that learns to represent visual inputs in a low-dimensional latent space. Evidence in favor of different choices can be gathered by sampling from these latent variables and feeding them to an integration-to-threshold model. Under some weak assumptions this architecture implements an SPRT test. Therefore, it can be used to provide an end-to-end computational account of reaction-time distributions as well as error-rates.

In contrast to DDMs, this model can explain how drift and diffusion rates arise rather than infer them from behavioural data. We show how to generate predictions using this model for perceptual decisions in visual noise and how these depend on different architectural constraints and the learning history. The model thus provides both an explanation of how evidence is generated from any given input and how architectural constraints and learning affect this process. These effects can then be measured through the observed error rates and reaction-time distributions. We expect this approach to allow us to bridge the gap between the complementary, yet rarely interacting literatures of decision-making, visual perceptual learning, and low-level vision/psychophysics.

Montero, M; Costa, R; Bowers, J; Ludwig, C; Malhotra, G

Angeliki Katsenou, University of Bristol, *"Predicting the rate-quality convex hull across resolutions for adaptive video streaming"*

A challenge that many video providers face is the heterogeneity of networks and display devices for streaming, as well as dealing with a wide variety of content with different encoding performance. In the past, a fixed bit rate ladder solution based on a "fitting all" approach has been employed. However, such a content-tailored solution is highly demanding; the computational and financial cost of constructing the convex hull per video by encoding at all resolutions and quantization levels is huge.

In this paper, we propose a content-gnostic approach that exploits machine learning to predict the bit rate ranges for different resolutions. This has the advantage of significantly reducing the number of encodes required. The first results, based on over 100 HEVC-encoded sequences demonstrate the potential, showing an average Bj{\o}ntegaard Delta Rate (BDRate) loss of 0.51\% and an average BDPSNR loss of 0.01 dB compared to the ground truth, while significantly reducing the number of pre-encodes required when compared to two other methods (by 81\%-94\%).

Katsenou, A.; Sole, J.; Bull, D.

Mubaraka Muchhala, University of Bristol, *"The geometry of high-level colour space".*

To communicate about the colours in our environment, humans categorise the continuous colour space, and the categories in our language bias our perception of colours.

We propose that these biases facilitate the perception of noisy photoreceptor signals by directing observers towards the most likely conclusions. The properties of high-level colour space were estimated by testing participants' colour memory across hue and saturation for two delay lengths. We predicted that responses would be biased towards category foci, and that categorical biases would be larger for the longer delay.

Muchhala, M., Scott-Samuel, N., Baddeley, R.

Aaron Zhang, University of Bristol, *"Enhancing video compression through deep learning"*

This presentation introduces a new Convolutional Neural Network (CNN) based post-processing approach for video compression, which is applied at the decoder to improve the reconstruction quality. This method has been integrated with the Versatile Video Coding Test Model (VTM) 4.0.1 and evaluated using the Random Access (RA) configuration using the Joint Video Exploration Team (JVET) Common Test Conditions (CTC).

The results show coding gains on all tested sequences at various spatial resolutions over different quantisation parameter ranges, with average bit rate savings of 3.90% and 4.13%, when PSNR and VMAF are used as quality metrics respectively. The computational complexities of different CNN architecture variants have also been investigated.

Zhang, F., Feng, C. and Bull, D. R.

Zhang, F.; Feng, C. & Bull, D. R. ENHANCING VVC THROUGH CNN-BASED POST-PROCESSING Proc. IEEE Int Conf. on Multimedia & Expo, 2020

Session 3

Dunia Gonzalez, University of Exeter, *"Colour and motion vision in stomatopod crustaceans"*

Stomatopod crustaceans have evolved complex visual systems. Their eyes are separated into hemispheres and bisected by a midband, which contains photoreceptors that have evolved up to twelve spectral sensitivities in the visible and UV range. However, their visual system does not facilitate enhanced spectral discrimination.

Recently, it was proposed that stomatopods detect colours as distinct excitation patterns that are processed without undergoing spectral comparisons or input from hemisphere photoreceptors involved in luminance vision. Such a system would allow for enhanced colour constancy and processing speeds, resulting in rapid responses to colour regardless of the lighting in the environment.

Motion vision was examined using the stabilising (optokinetic) response to motion. The optokinetic action spectrum of two shallow-water stomatopod species, Gonodactylus chiragra and Pseudosquilla ciliata, were recorded to determine the spectral sensitivities of their motion vision. The spectral sensitivity curves of both species matched those of their hemisphere photoreceptors, suggesting that motion detection is dependent on intensity, rather than colour vision. Additionally, the innate response to colour when selecting a burrow refuge was examined. Stomatopods may avoid entering burrows occupied by intraspecific competitors as their brightly coloured carapaces and appendages could play important roles in signalling their presence to conspecifics. P. ciliata, but not G. chiragra, avoided colours reflected from the carapaces of conspecifics.

Additionally, colour preferences when selecting a burrow occurred independently of intensity. The findings of both the optokinetic and burrow choice experiments demonstrate clear functional partitioning of the stomatopod visual system across different eye regions and for different tasks.

Gonzales, D.

Cheroske, AG., Cronin, T.W., Durham, M.F., Caldwell, R.L.(2009). Adaptive signalling behavior in stomatpods under varying light conditions. Mar. Freshwater Behav. Physiol. 42, 219-232. doi: 10.1080/10236240903169222. Gagnon, Y.L., Templin, R.M., How, M.J., Marshall, N.J. (2015). Circularly polarized light as a communication signal in Mantis Shrimps. Curr. Biol. 25, 3074-3078. doi: 10.1016/j.cub.2015.10.047 Land, M.F. (1999). Motion and vision: why animals move their eyes. J. Comp. Physiol. A. 185, 341-352. doi: 10.1007/s003590050393 Marshall, N.J., Jones, J.P. and Cronin, T.W. (1996). Behavioural evidence for colour vision in stomatopod crustaceans. J. Comp. Physiol. A. 179, 473-481. doi: 10.1007/bf00192314 Schaerer, S. and Neumeyer, C. (1996). Motion detection in goldfish investigated with the optomotor response is “color blind”. Vision Res., 36, 4025-4034. doi: 10.1016/s0042-6989(96)00149-6 Stojcev, M., Radtke, N., D’Amaro, D., Dyer, A.G., Neumeyer, C. (2011). General principles in motion vision: color blindness of object depends on pattern velocity in honeybee and goldfish. Vis. Neurosci. 28, 361-370. doi: 10.1017/S0952523811000101 Thoen, H.H., How, M.J., Chiou, T.H., Marshall, J. (2014). A different form of color vision in mantis shrimp. Science 343, 411-413. doi: 10.1126/science.1245824. Thoen, H.H., Sayre, M.E., Marshall, J., Strausfeld, N.J. (2018). Representation of the stomatopod’s retinal midband in the optic lobes: putative neural substrates for integrating chromatic, achromatic and polarization information. J. Comp. Neurol. 526, doi: 10.1002/cne.24398

Sandra Winters, University of Bristol, *“Simulating the evolution of inter-specific mating signal diversity in guenons”*

When closely related species overlap, selection for discrimination between con- and heterospecific mates can lead to the evolution of species-typical mating signals and associated mating biases. Such processes can be important drivers of animal diversity. We simulate this process using a machine learning approach in which mating signals, modeled as points in phenotype space, evolve across evolutionary time using a genetic algorithm and selection for various types of mate choice.

Our simulations are based on mating signals in guenons, a recent primate radiation in which species commonly form mixed-species groups and exhibit colorful and diverse face patterns hypothesized to function in the maintenance of reproductive isolation. We show how diversification in isolation and mate choice on secondary contact can induce rapid phenotypic diversification in guenons, resulting in face patterns that are distinctive between species but stereotyped within species, similar to those observed in guenons today. Strong selection against hybrids is key to diversification, with even low levels of hybrid reproduction typically lead to the merging of populations on secondary contact.

Our results show how phenotypic evolution plays out across ecological scenarios in guenons, and ultimately support a role for reinforcement in the evolution of guenon face patterns. This research highlights how simulated evolution grounded in realistic species biology can generate insights into the evolution of phenotypic diversity.

Winters, S; Higham, JP

Jake Deane, CAMERA - University of Bath, *“DogDetector: animal segmentation using synthetic images and 3D generative models”*

Segmentation is one of the classic research areas in computer vision. Recent advances have made use of deep learning models to achieve state of the art performance, taking advantage of greater computational power. However, these methods typically require large sets of training data, which for some object classes – such as animals – are not available.

In recent years however synthetic data has been used to train machine learning models that are able to compete with their peers trained on real data (Varol et al. 2017; Dunn et al. 2019; Mu et al. 2019). In this paper, we present a generative 3D canine model which we use to create synthetic data for refining segmentation models via a generative adversarial frame-work (Goodfellow et al. 2014), using real and synthetic canine images and segmentation maps where the synthetic data was generated by the aforementioned generative 3D canine model.

We performed experiments on three datasets with 9 permutations of generator and discriminator to investigate the effect of synthetic data on our segmentation refinement experiments compared to real data. We found that it is possible to refine” off the shelf” baseline semantic segmentation models (Long, Shelhamer and Darrell 2014; Chen et al. 2017; He et al. 2018) to deliver superior performance for a single asset class using synthetic data to augment existing datasets.

Deane, J, Kearny, S, Cosker, D, Kim, K

Chen, Liang-Chieh et al. (May 2017). “DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs”. In:arXiv:1606.00915 [cs]. arXiv:1606.00915 [cs].

Dunn, Kenneth W. et al. (Dec. 2019). “DeepSynth: Three-Dimensional Nuclear Segmentation of Biological Images Using Neural Networks Trained with Syn-thetic Data”. en. In:Scientific Reports9.1, pp. 1–15.issn: 2045-2322.doi:10.1038/s41598-019-54244-5.

Goodfellow, Ian et al. (2014). “Generative Adversarial Nets”. en. In:Advances in neural information processing systems, pp. 2672–2680. He, Kaiming et al. (Jan. 2018). “Mask R-CNN”. en. In:arXiv:1703.06870 [cs].arXiv:1703.06870 [cs].

Long, Jonathan, Evan Shelhamer, and Trevor Darrell (2014). “Fully Convolu-tional Networks for Semantic Segmentation”. en. In:CoRRabs/1411.4038. Mu, Jiteng et al. (2019). “Learning from Synthetic Animals”. en. In:ArXivabs/1912.08265.

Varol, G ̈ul et al. (July 2017). “Learning from Synthetic Humans”. en. In:2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),pp. 4627–4635.doi:10.1109/CVPR.2017.492. arXiv:1701.01370.