



UNIVERSITY OF  
**SURREY**

IAS

Institute  
of Advanced  
Studies

## Human and Computer Models Workshop

In-person | Wednesday 15<sup>th</sup> May 2024 | University of Surrey

Leggett Building, Manor Park, Daphne Jackson Rd, Guildford GU2 7WG

### Programme

9:30 – 10:00	Registration Tea and Coffee	
10:00 – 11:00	<p><b>Invited Speaker:</b>  <b>“A Systems View of Perceiving Biological Motion: from Features to Brain Circuits”</b>  <b>Professor Frank Pollick</b>  School of Psychology and Neuroscience, University of Glasgow</p> <p><u>Abstract:</u>  This talk will present research results from our lab that have explored the perception of human movement – also known as biological motion perception. In the first part of the talk, I will discuss movement features and how they contribute to the categorisation of different movement styles. Next, I will examine the different frameworks that have been proposed for the neural processing of movement and how these inform evaluation of our research into watching dance. Finally, I will address how individual differences might impact the perception of human movement by reviewing our studies on autism and the effect of CCTV-operator experience in judging harmful intent. Throughout the talk I will highlight ways in which quantitative measures of human movement have been related to perception and underlying brain activity.</p>	
11:00 – 12:20	Oral Presentation Session	
	<p>“The effects of depth of field on attention whilst exploring a virtual environment”.  <i>S. Durant, D. Gulhan, H. Kaur Suri, K. Vekariah, J. Haider-Smith</i></p>	Szonya Durant
	<p>“Spatial Cognition from Egocentric Video: Out of Sight, Not Out of Mind”.  <i>T. Perrett, D. Damen, A. Kanazawa, S. Goel, C. Plizzari</i></p>	Toby Perrett
	<p>“On the incidental and deliberate visual processing of communicative interactions”.  <i>A. Atkinson, Q. Vuong</i></p>	Anthony Atkinson
	<p>“Looking at people in videos: Evidence from human eye movements”.  <i>T. Foulsham</i></p>	Tom Foulsham

12:20 – 12:30	Poster announcements	
	“StableTalk: Advancing Audio-to-Talking Face Generation with Stable Diffusion and Vision Transformer”. <i>F. Nazarieh</i>	Fatemeh Nazarieh
	“Advancements in 3D Plant Phenotyping: Precise Part Segmentation and Trait Measurement Through Video-Derived Point Clouds”. <i>R. Reena</i>	Reena Reena
	“Driving Through Graphs: A Networked Perspective on Scene Representation”. <i>A. Humnabadkar</i>	Aditya Humnabadkar
12:30	Sandwich/buffet lunch in the venue	
13:00	<b>Poster Session 1:</b> Szonya Durant, Toby Perrett, Anthony Atkinson, Tom Foulsham, Fatemeh Nazarieh	
13:30	<b>Poster Session 2:</b> Reena Reena, Aditya Humnabadkar, Arindam Sikdar, Quoc Vuong, Filip Rybansky	
14:00 – 15:00	<b>Invited Speaker:</b> <b>“Multimodal Learning in Video Moment Retrieval”</b> <b>Professor Shaogang Gong</b> Queen Mary University of London, Queen Mary Computer Vision Laboratory  <u>Abstract:</u> Deep learning has revolutionised AI machine learning techniques in computer vision over the past decade largely due to the availability of centralised big data with exhaustive labelling and cheap computing power from Nvidia’s GPUs. However, privacy concerns from data protection and environmental concerns on energy consumption together with an increasing demand for decentralised user-ownership of localised unlabelled data pose fundamental challenges to the established wisdom of deep learning on centralised big data from scratch with exhaustive labelling available for model training. In this talk, I will present challenges and recent progress on exploring multimodal vision-language models for self-supervised learning of fine-grained video-language dynamic details without fine-grained labelling in model training for video moment retrieval.	
15:00 – 15:30	Afternoon coffee	
15:30 – 17:00	Oral Presentation session	
	“Scale-invariant batch-adaptive residual learning for person re-identification” <i>A. Sikdar, A. S. Chowdhury</i>	Arindam Sikdar
	“Measuring temporal adaptation in videos of speech” <i>Q. Vuong, M. Laing, V. Bansal, A. Rees</i>	Quoc Vuong
	“Semantic consistency in identifying human actions” <i>F. Rybansky, S. Rahmani, A. Gilbert, F. Guerin, Q. Vuong</i>	Filip Rybansky
	“Movie-watching reveals that human extrastriate cortex is tiled with somatosensory homunculi” <i>N. Hedger, T. Knapen</i>	Nicholas Hedger
17:00	Close	