

Real-Time Semantic Depth Layer Decomposition for Augmented Reality (in collaboration with SnapChat)

Project Description

We are looking for a motivated candidate, with a background in computer vision/computer graphics/machine learning to work on an exciting new collaborative project with SnapChat.

The increasing popularity of augmented reality/mixed reality (AR/MR) has been driven by the wide range of beneficial applications (e.g. immersive entertainment, communication, collaborative design, medical visualisation) and the prevalence of commodity hardware (e.g. ARkit/ARcore in iPhone and Android smartphones and tablets). These applications require high-speed, accurate tracking of the location of the device as well as real-time decomposition and semantic interpretation of the observed scene.

The former challenge has received substantial research and industrial attention with hardware and software systems that can perform robust, high-frame-rate tracking in real-life environments. The second challenge, however, is still an open research question but it is vital to enable a truly immersive combination of virtual objects into a real scene (e.g. real objects occluding virtual objects, virtual objects casting shadows onto real objects).

In this project, we will develop new models and algorithms to decompose the visual scene into semantically meaningful depth layers to allow insertion of virtual objects before the scene is recombined to provide the augmented/extended experience.

The technical challenge is two-fold. Firstly, the creation of a machine learning approach for tracking and layer decomposition that is sufficiently robust to achieve invariance or equivariance to the range of phenomena that build a scene view (including, e.g., lighting, shape, texture, occlusion). We will build on our previous work on interactive tracking of shape models [Roto++, SIGGRAPH 2016] to train an unsupervised (i.e. without expensive human annotation of images) tracker based on synthetic training data. Such data can be obtained in a generative fashion using game engine technology and real-world examples. In addition, we will advance previous work from Dr Richardt on scene decomposition [Live Intrinsic Video, SIGGRAPH 2016] to break down the video stream into multiple depth layers that may then be combined in novel ways.

The second challenge will be constructing inference algorithms and appropriate representations that ensure these problems can be performed efficiently in real-time; this will particularly build on the expertise of the industrial supervisor who has over fifteen years of experience in real-time computer vision with a long track record including the Koenderink Prize for work that has stood the test of time.

Both of these stages will necessitate the development of a new theoretical model that can combine the more established priors and likelihood functions of geometric computer vision (e.g. our work on the 3D shape of deformable surfaces) with the poorly understood and constrained functions in semantic vision (i.e. recognising the content/objects in a visual image). In recent work, we have demonstrated that it is possible to use our nonparametric priors with popular deep learning methods and we are now exploring techniques to propagate the uncertainty through these networks. This will allow us to combine the probabilistic graphical models from geometric vision with the representation learning from semantic vision.

Progress on this front will represent an important contribution to the state-of-the-art in computer vision beyond the remit of this project and will have impact across the computer vision community. We envisage high-profile papers to be published at top venues in both vision (CVPR/ICCV) and graphics (SIGGRAPH).

Funding Notes

Applicants should hold, or expect to receive, a First Class or good Upper Second Class Honours degree, or the equivalent from an overseas university. A master's level qualification would also be advantageous.

Funding will cover Home/EU tuition fees, a stipend (£14,777 per annum for 2018/19) and a training support fee for 3.5 years. Early application is strongly recommended.

Applicants classed as Overseas for tuition fee purposes are not eligible for funding; however, we welcome all-year-round applications from self-funded candidates and candidates who can source their own funding.

Application

The application deadline for this position is Friday, 27 April 2018.

Applications can be submitted via FindAPhD:

<https://www.findaphd.com/>.

For more general information on studying for a PhD in computer science at Bath, see:

<http://www.bath.ac.uk/science/phd-computer-science/>.

References

W. Li, F. Viola, J. Starck, G.J. Brostow and N.D.F. Campbell, "[Roto++: Accelerating Professional Rotoscoping using Shape Manifolds](#)", in ACM Transactions on Graphics (Proceedings of SIGGRAPH 2016).
A. Meka, M. Zollhöfer, C. Richardt and C. Theobalt, "[Live Intrinsic Video](#)", in ACM Transactions on Graphics (Proceedings of SIGGRAPH 2016).

[Dr Neill Campbell](#)

[Dr Christian Richardt](#)

Department of Computer Science, University of Bath