

BMVA Symposium: Deep Learning in 3D

Wednesday 20th February 2019

Chairs: Chris Holder, Christopher Willcocks and Grégoire Payen de La Garanderie, Durham University

Organiser: Andrew Gilbert, University of Surrey

This technical meeting was held in the BCS headquarters in London. It aimed to explore the key challenges of combining deep learning with 3D vision.

The morning session started with a keynote from **Adrian Hilton** on 3D shape capture across time hence his title: “4D Vision”. Through his work with the entertainment industry, he highlighted specific challenges for 3D vision moving from controlled performance capture environments to real-world unconstrained environments.

In the next talk on 3D Pick & Mix, **Adrián Peñate-Sánchez** reasoned not only about whole objects but also about their parts. This work learns a shared embedding between images and shapes using synthetic images and a realistic renderer. This embedding can be used to retrieve 3D model of furniture using a combination of query images representing the desired parts.

Subsequently, **Oliva Wiles** talked about recovering the 3D shape of sculptures from a single image. The dense depth estimator is trained by extracting sparse point correspondences in a multiview setup — an approach that she called Pointillism. During training, camera pose is estimated using SIFT and assuming an affine camera model.

The next talk by **Andrew Gilbert** delved into volumetric performance capture using a minimal number of viewpoints. The work is based on a convolutional auto-encoder. The network is trained on 8 input views however only 2 inputs are required at test time.

Finally, in last talk of the morning session, **Viswadeep Sarangi** looked at the clinical evaluation of machine learning approaches for 3D gait classification. Interpretability is a key part of the approach due to the medical setting and can also be used to provide insights on which gait features are relevant to classify different gaits.

The afternoon session started with a keynote from **Alex Kendall** who showed that 3D geometry understanding is still important in an era of end-to-end deep learning. For instance, an end-to-end lane-following vehicle can be trained in the real-world — rather than simulation — using reinforcement learning in very few iterations if the solution is adequately constrained using geometry.

In the first talk of the second session, **Michael Edwards** introduced Graph Convolutional Network as a technique to replicate convolutions and pooling on a graph structure rather than the classical 2D or 3D grid representation of an image. This is used to reduce the size of the required network to segment structures in large 3D medical images.

Our last speaker, **Georgi Tinchev** presented his work on relocalising in a map using a laser-based point cloud. The approach allows localisation in difficult environments such as a forest. Moreover, the model is kept small in order to run within the limited computational payload of a drone.

Last but not least, we had four spotlight presentations on a plug-and-train loss for single view 3D reconstruction by **Eduard Ramon Maldonado**, ground classification in 3D point clouds of large-scale heritage sites by **Dimitrios Makris**, fruit detection with 3D deep neural networks by **Justin Le Louëdec** and the recovery of superquadric parameters by **Aleš Jaklič**.

The chairs would like to thanks Adrian Hilton, Alex Kendall, all the speakers and attendees for making this meeting a success.

Report by Grégoire Payen de La Garanderie, Durham University.