W20 – PeopleCap 2018: Capturing and Modeling Human Bodies, Faces and Hands

W20 – PeopleCap 2018: Capturing and Modeling Human Bodies, Faces and Hands

The second edition of PeopleCap brought together expert researchers in the field of capturing and modeling humans from sensor data. This time, PeopleCap took place in conjunction with ECCV, in Munich on the last day of the conference. After five days of exciting papers and events at ECCV, the PeopleCap workshop was very well attended. The workshop consisted of excellent talks given by the invited speakers and stimulating discussions throughout the poster session.

Lourdes Agapito started the workshop with an exciting talk, first describing a deep learning method for inferring 3D pose followed by an optimization based technique to reconstruct 3D face geometry from monocular images. Adrian Hilton presented high end performance capture methods that leverage multiple cameras for both indoor and outdoor settings, some of which are already being used to render novel views during broadcasts of sport events. Franziska Mueller then demonstrated state of the art hand tracking results, showing how to learn to infer 3D hand pose from images of hands synthetically generated using a GAN.

The workshop continued with a poster session that included a selection of invited posters from other portions of the conference, along with posters for the five papers accepted to the workshop, each of which was peer reviewed by three expert reviewers in the field. The session featured a good mix of papers tackling 3D human pose estimation, reconstruction from single images and modeling of cloth, faces and hands.

After the poster session, Yaser Sheikh described methods to model and capture human, emphasiznig the need to capture as much detail as possible because "every detail carries a social signal" that will enable us to study the semantic intentions of body language. Stefanie Wührer finished with an encompassing talk describing methods to learn multilinear models of faces, factoring expression and shape, from noisy 3D face registrations with missing data.

The workshop concluded with the announcement of the best paper award. The award was given to the paper "Can 3D Pose be Learned from 2D Projections Alone?", a paper that demonstrated an interesting weakly supervised approach to estimate 3D pose points, given only 2D pose landmarks.

In summary, the workshop was a great success, allowing researchers in the field to meet and exchange ideas and thoughts, reflecting on the importance of integrating 3D geometry and reasoning in computer vision.

September 2018

Gerard Pons-Moll
Jonathan Taylor