

OPENING KEYNOTE by Andrew Zisserman (*Oxford, UK*)

Towards On-the-fly Large Scale Video Search

Abstract. We would like to be able to find anything in an image or video dataset. The talk will describe our progress on visual search for finding people, specific objects and categories in large scale video datasets. The novelty is that the item of interest can be specified at run time by a text query, and a discriminative classifier for that item is then learnt on-the-fly using images downloaded from Google Image search. We will compare state of the art encoding methods for the problem, and discuss the choices in achieving the best trade-off between three important performance measures for a realtime system of this kind, namely: (i) accuracy, (ii) memory footprint, and (iii) speed. We will also describe steps to achieving `total recall'. There will be demonstrations on a large scale video dataset of BBC broadcasts. This is joint work with Relja Arandjelovic, Ken Chatfield and Omkar Parkhi.

Biography. Professor Andrew Zisserman leads the Visual Geometry Group at the University of Oxford, UK. Andrew's research interests include visual recognition, image retrieval, multi-view geometry, and other aspects of computer vision. Some of Andrew's papers are amongst the most highly cited works in the field. His contributions received multiple awards at the top computer vision conferences including three Marr prizes at the International Conferences on Computer Vision. He has published several books including "Visual Reconstruction" (with Andrew Blake) and "Multiple View Geometry in Computer Vision" (with Richard Hartley). He is a fellow of the Royal Society.