

V-Lab: Using agent technology in a collaborative environment to bridge the semantic gap in the annotation of video content

Dr David Shotton, Image Bioinformatics Research Group, Oxford University
e-mail: david.shotton@zoo.ox.ac.uk



The problem

- The development of novel search and retrieval procedures for digital video first requires effective means of **semantic content annotation**
- Unfortunately, there exists a '**semantic gap**' between presently achievable automated low-level feature extraction procedures and the high-level semantic descriptions of video content required for 'smart' media retrieval
- As Edward Delp (Purdue University) concluded in his plenary address "*Are low level features too low for indexing?*" at EWIMT 2005, in London on 30 November:
"Rich indexing will always require some form of human or manual work"
- Thus, for the foreseeable future, high-level semantic content analysis of video *will* require **human expertise** to recognise features and events of significance
- However the effort currently required for unaided manual video analysis, approximately ten hours work per hour of video footage, is clearly unfeasible
- We need to **remove this analysis and annotation bottleneck** if we are to enable effective video searching and access to valuable content

IST Directorate D Audio-Visual Search Technologies, Brussels, 19 December 2005

V-Lab: A potential solution



The V-Lab Project aims to reduce the threshold for **interactive video annotation and analysis**, both by enabling **collaborative** video annotation, and by **bridging the gap** between human video content analysis and automated feature extraction, facilitating more cost-effective annotation of valuable video assets

We propose a three-fold approach to advance the state of the art:

- We will create a robust generic framework based on Jane Hunter's *Vannotea* system to enable **synchronous collaborative video analysis and annotation** by geographically separated human experts, as well as by individual analysts
 - V-Lab will provide an efficient user interface and will automate all technical aspects of video presentation and recording of spatio-temporal annotations
- We will use **Semantic Web technologies** (ontologies, RDF, OWL, etc) to provide the semantic framework, based on our BioImage Database, for ontology-based annotation of the content of digital video files
 - This will permit subsequent '**smart**' searches over the stored annotations, and will thus enable asset discovery and knowledge sharing
- We will integrate **computational agents** into this collaborative environment
 - The agents will undertake as much as possible of the **low-level tasks** such as feature recognition and object tracking, *under interactive human control*
 - This will **reduce the cognitive load** on the human analysts, permitting them to concentrate on their primary task of semantic interpretation

IST Directorate D Audio-Visual Search Technologies, Brussels, 19 December 2005

V-Lab: Project scope and consortium strengths

- **Target:** Annotation of scientific and commercial video of high intrinsic worth
- **Development stage:** Pre-competitive research
- **Innovation and risk factor:** Very high **Potential benefits:** Very high
- **Leverage:** Existing Vannotea and BioImage Database systems  

The V-Lab Academic Consortium:

- **David Shotton (PI), Image Bioinformatics Research Group, Oxford University:**
The BioImage Database, an ontology-centred video database; Semantic Web skills; semantic agent integration; provision of scientific video use cases
- **Marina Jirotko, Computing Laboratory, Oxford University:**
Expertise in e-Science usability and requirements analysis
- **Andrew Zisserman, Engineering Dept, Oxford University & Philip Torr, Computer Sciences, Oxford Brookes University:** Automated video analysis algorithms
- **Jane Hunter, University of Queensland, Brisbane, Australia:**
Vannotea collaborative video annotation system
- **Stephen Druce, ACE Centre (a medical charity), Churchill Hospital, Oxford:**
Video use cases for patient evaluation of young people with learning disabilities

We seek consortium partners, particularly from industry!

E-mail me! david.shotton@zoo.ox.ac.uk

IST Directorate D Audio-Visual Search Technologies, Brussels, 19 December 2005