

---

# **TRECVID-2006: Rushes Exploitation Task**

---

Alan Smeaton  
Dublin City University  
&  
Tzveta Ianeva  
NIST

# Rushes Exploitation Task Definition

---

- Goal: research about the feasibility of shifting to work on unproduced video;
- Develop a toolkit for support of exploratory search on highly redundant rushes data
  - Summarize -- remove/hide redundancy of as many kinds as possible
  - Organize – present non-redundant material according to at least 6 not all cinematographic or camera setting features, well motivated from some user/task context point of view
- Develop an evaluation scheme

# Rushes Exploitation Task Definition

---

- Innovation in approaches – no standard keyframes or shot boundaries provided
- Evaluation and presentation of results by participants
- Data: 50 hours of rushes provided by BBC Archive
  - French experience, travel videos and interviews
- Video example

# 2006: Rushes exploitation task participants

---

## Accenture Technology Labs

AT&T Labs - Research

Chinese Academy of Sciences (CAS/MCG)

Curtin U. of Technology

DFKI GmbH

U. Rey Juan Carlos/ Dublin City U.

IBM T. J. Watson Research Center

Institut EURECOM

Joanneum Research Forschungsgesellschaft mbH

Tsinghua U.

U. of Marburg

COST292 ([www.cost292.org](http://www.cost292.org))

## USA

USA

China

Australia

Germany

Spain/Ireland

USA

France

Austria

China

Germany

Fr, Neth, UK,  
Irl, Gr, Turk,  
Serbia & Mont.

Slovakia

Accenture & COST292 have speaker slots to follow

# AT&T Labs - Research

---

- Collaboration w/ Columbia U.;
- Motion-based shot segmentation (cf. Tsinghua);
- Large (374) set of LSCOM HLFs on resulting shots;
- Computed image (KF ?) distances within each video file;
- Annotation into 1 of 15 audio classes, (speech/non, male/female, range of low-level audio features);
- Speaker segmentation;
- Browsing interface built;

# Chinese Academy of Sciences (CAS/MCG)

---

- Apply a range of concept feature identification to audio and visual ... face, interview, person, etc.
- Some SVM classifiers used (LSCOM ?)
- Camera motion infers intention;
- Hierarchical browsing to address redundancy and repetition;
- Interface for filter/browse is built;

# Curtain U of Technology

---

- Not sure what they did, paper outside, no time to read yet !

# DFKI

---

- No paper
- Fast motion-based features w/ spatial aspects
- Clustering



# U. Rey Juan Carlos

---

- A kind of shot bound detection .. detecting substantial differences as “events”;
- User filtering of useless shots ... eg calibration of camera;
- Apply 39 SVMs from DCU feature submission;
- Build interface to filter and browse keyframes;
- Collaboration w/ DCU

# IBM T.J. Watson

---

- No details except an exploration of semantic concept models from B/news to rushes, exploring feature-based and semantic-based clustering;

# Institut Eurecom

---

- Perform SBD then remove inter- and intra-shot redundancy by similarity and then hierarchical clustering;
- User search based on visual dictionary - cluster rushes keyframes into groups, these groups form 'words' in the dictionary;
- Proposed evaluation through simulated user experiments

# Joanneum Research

---

- Pictorial summary, allowing browsing, ordering and annotation of video;
- Used camera motion, motion intensity, global and local visual similarity, audio volume, faces and object re-detection ... all to generate a pictorial summary of video file;
- Performed some user tests and evaluation ... 7 uses, 4 search tasks;

# Tsinghua University

---

- Similar to URJC using motion and SBD to yield KFs, and then ran some hierarchical clustering to remove redundancy, then some HLFs on the remainder.

# Univ. Marburg

---

- Again, unsupervised clustering of shots to eliminate redundant shots;
- Additionally, the following features can be used to “slide and dice” - camera motion, faces, shot lengths, audio information, interviews, ... 13 in total;
- Interactive browsing tool developed and evaluated in-house;

# Observations

---

- A good number of groups can build systems to ingest, analyze, and allow user filtering and summarization;
- Most redundancy detected through clustering;
- Surprising emphasis on audio classification;
- Few groups did actual evaluation ... those that did did classic ad hoc search;