TREC 2007 Video Retrieval Evaluation
Introductions

Paul Over*
George Awad*
Lori Buckland*
Darrin Dimmick*

Wessel Kraaij (TNO ICT)
Alan Smeaton (DCU)

* Retrieval Group
Information Access Division
Information Technology Laboratory
NIST

Goals

- Promote progress in content-based retrieval from large amounts of digital video –
  - combine multiple errorful sources of evidence
  - achieve greater effectiveness, speed, usability
- Model an analyst interested in finding video of certain people, things, places, events, and combinations thereof
  - work on what video is “of” rather than what it is “about”
- Similar to needs of commercial or amateur video producers searching an archive for video to reuse (rather than reshoot)

Goals

- Focus on relatively high-level functionality – near that of an end-user application like interactive search
- Confront systems with unfiltered data and realistic queries
- Measure against human abilities
- Supplement with focus on supporting automatic components
  - Automatic search, High-level feature detection
  - Shot boundary determination
- Integrate and profit from advances in low-level functionality, more narrowly tested
  - face recognition, text extraction, object recognition, advances, etc.

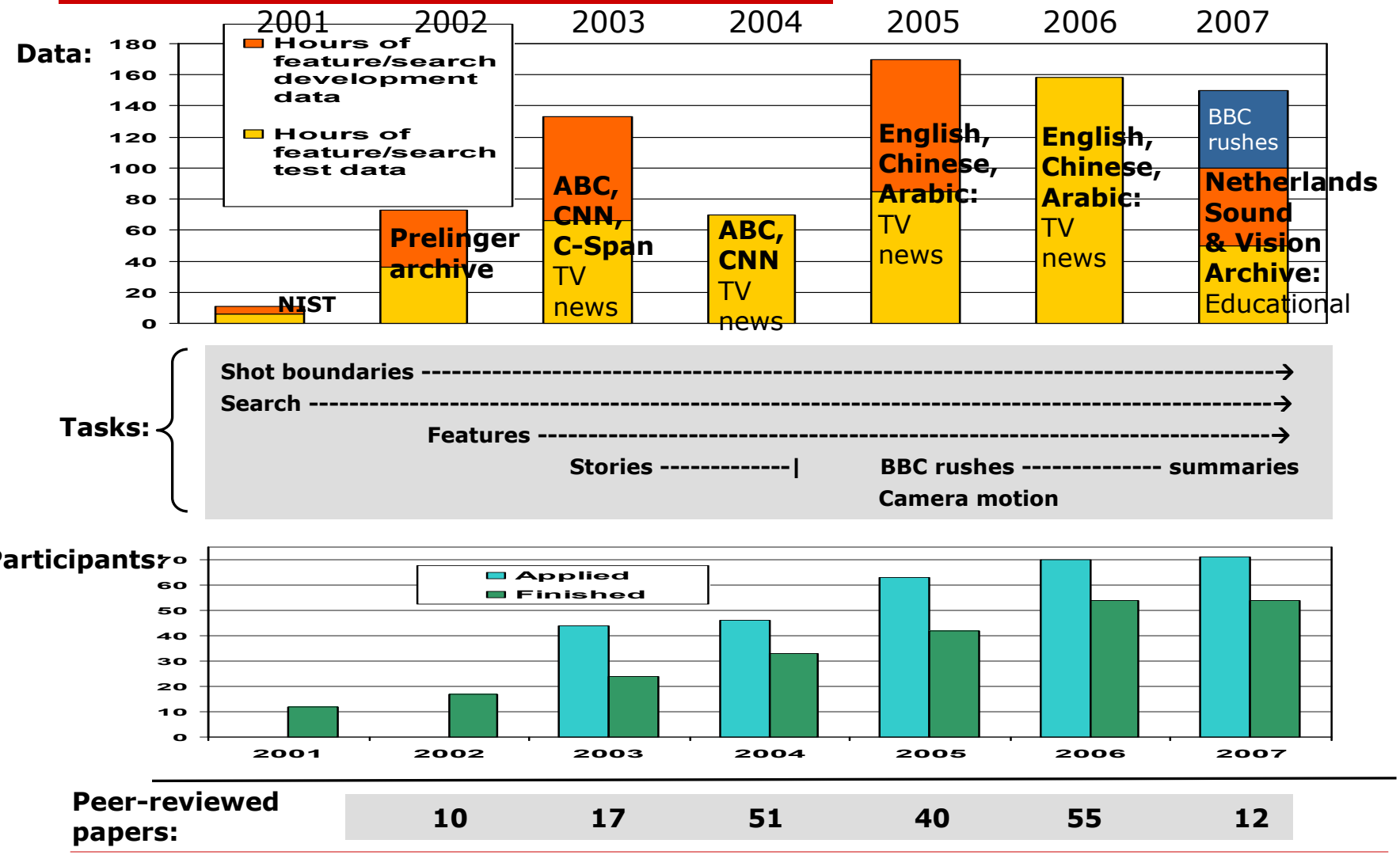
Questions: General

- How can systems achieve such retrieval (in collaboration with a human)?
  - usefulness of generic features
    - which features most useful?
    - how/when to combine?
    - how to learn them efficiently
  - human & system collaboration
    - who does what?
    - what is the optimal interface?
- How can one reliably benchmark such systems?

Questions: Especially this year

- Do features trained on broadcast news work on new sorts of data?
- How can feature detectors trained on broadcast news be used in combination with those trained on new sorts of data?
- Do search approaches effective against broadcast news generalize to the Sound and Vision archive video?
- How are event topics best handled?
- Do denser samples of the video improve effectiveness?
- Etc.

Evolution: data, tasks, participants,...



Evolution: … 2007

- Data:
  - 100 hrs (Netherlands Institute for Sound and Vision)
  - 50 hrs (BBC)

- 4 evaluated tasks
  - S&V news magazine, cultural, educational/entertainment
    1. Shot boundary determination
    2. High-level feature extraction
    3. Search (automatic, manually-assisted, interactive)
       - Base scenario: search of a real, large, diverse, contemporary video archive
  - BBC dramatic rushes
    4. Video summarization
       - ACM Multimedia 2007 TRECVID Video Summarization Workshop

Sound & Vision data is different!

- still contains people, things, places, events …
- but differs from 2003-2006 news video in style and content
  - Smaller number of training shots – fewer examples
  - More variety of people, objects, locations, events
    - Almost no newsroom and anchor shots
  - Historical (grayscale) video mixed with modern color footage
  - Much less repetition
    - No commercials
    - Fewer previews
    - Little repeated file footage
  - Overall increase in difficulty?
    - Target video segments rarer overall?
    - Diminished usefulness of (near) duplicates?

Additional resources and contributions

- Christian Petersohn (Fraunhofer (Heinrich Hertz) Institute) for the master shot segmentation
- Peter Wilkins at the DCU Centre for Digital Video Processing for formatting the master shot reference
- City U. Hong Kong, U. of Modena and Regio Emilio, and Univ. of Iowa for running the data mirrors
- Roeland Ordelman and Marijn Huijbregts for ASR
- Christof Monz for MT
- Philip Kelly at DCU for running the summarization ground truthing

Additional resources and contributions

- Eric Zavesky at Columbia U.
  - Low-level features, classifier scores for 2007 development data using 374 Columbia models trained on 2005 data
- Yu-Gang Jiang at City U. Hong Kong
  - Keypoint-based and baseline features for 374 LSCOM concept detectors trained on 2006 test data
- Georges Quénot, Stéphane Ayache, and many participants
  - Annotation of large part of the 2007 development data for 36 features
- Sheng Tang and team in the Multimedia Computing Group (CAS)
  - Annotation of 2007 development data for 36 features

International participation

Evaluated tasks: 54 finishers

Asahi Kasei Corporation	SB ** -- --
AT&T Labs	SB -- -- SU
Beijing Jiaotong University (Northern Jiaotong Univ.)	-- -- SE --
Beijing University of Posts and Telecommunications	SB -- -- --
Bilkent University	** FE SE **
Brno University of Technology	SB FE ** SU
Carnegie Mellon University	-- ** ** SU
City University of Hong Kong (CityU)	-- FE SE SU
Columbia University	-- FE ** SU
COST292 Team	SB FE SE SU
Curtin University	** -- -- SU
CWI-CTIT-U.Twente	-- ** SE --
Dublin City University	-- -- SE SU
École Nationale Supérieure des Télécommunications / TSI	-- FE -- --
Etter Solutions Research Group	-- -- SE --
Florida International University, FIU-UM	SB ** -- --
Fraunhofer Institute IAIS and University of Bradford	SB ** -- --
Fudan University	-- FE SE -

** = applied for but did not submit runs

Evaluated tasks: 54 finishers

FX Palo Alto Laboratory Inc.	** ** SE SU
Helsinki University of Technology	** FE SE SU
HuaZhong University of Science and Technology	SB ** ** **
IBM T. J. Watson Research Center	** FE SE **
INESC-Porto	-- ** SE --
Institut EURECOM	-- FE -- SU
JOANNEUM RESEARCH Forschungsgesellschaft mbH	** FE -- SU
KDDI R&D Labs, Inc./ Tokushima U. / Tokyo U.	** FE -- SU
K-Space	-- FE SE --
LIG (Laboratoire d'Informatique de Grenoble)	SB FE ** **
LIP6 - Laboratoire dInformatique de Paris 6	-- FE -- SU
Microsoft Research Asia	-- FE SE **
Multimedia Content Analysis Group (CAS)	-- FE -- --
Multimedia Computing Group (CAS) / Natl. U. of Singapore	-- FE SE **
National Institute of Informatics	-- FE -- SU
National Taiwan University	-- FE ** SU
NHK Science and Technical Research Laboratories	SB ** -- --
Oxford University	-- FE SE --

Evaluated tasks: 54 finishers

Philipps University Marburg	SB FE ** **
The Hong Kong Polytechnic University	-- -- -- SU
Tokyo Institute of Technology	** FE ** **
Tsinghua University / Intel China Research Center	SB FE SE SU
Universidad Autónoma de Madrid	-- ** -- SU
Universidad de Jaén (SINAI)	-- -- SE --
University of Karlsruhe (TH)	SB FE -- --
University of Amsterdam (MediaMill Team)	-- FE SE --
University of California, Berkeley	-- FE ** --
University of California, Santa Barbara	-- FE SE SU
University of Central Florida	-- FE SE **
University of Electro-Communications	-- FE ** --
University of Glasgow	-- -- SE SU
University of Iowa	** FE SE --
University of Louisville	-- FE -- --
University of Modena and Reggio Emilia	SB ** -- **
University of Queensland	-- -- SE --
University of Sheffield	SB -- -- SU

2007: Extended teams

- COST292
  - LABRI, Bordeaux
  - Delft University of Technology, Netherlands
  - Bilkent University
  - Dublin City University
  - National Technical University of Athens
  - Queen Mary, University of London
  - ITI, Thessaloniki, Greece
  - University of Belgrade
  - University of Zilina
  - University of Bristol

2007: Extended teams

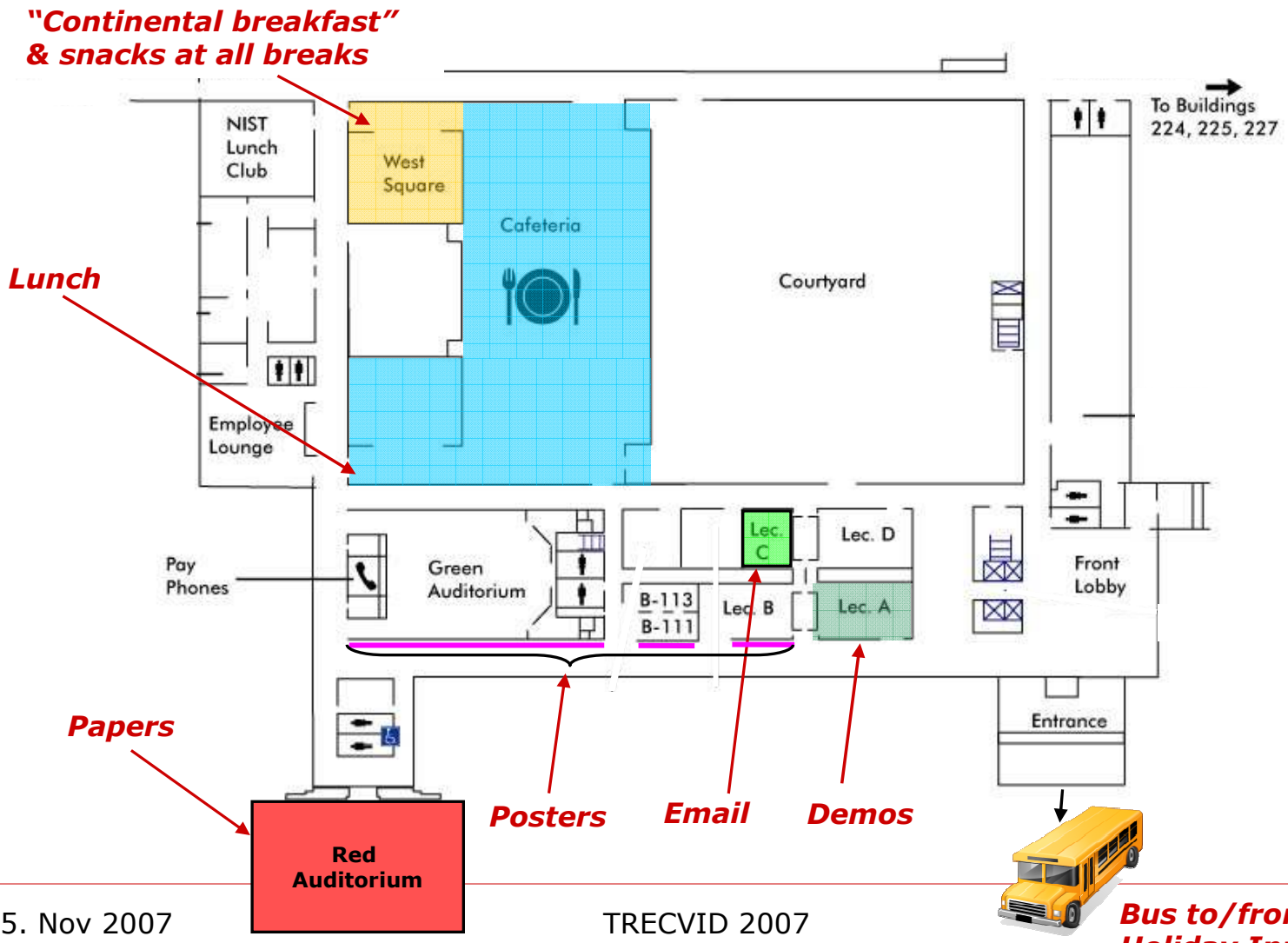
- K-SPACE
  - Queen Mary University of London
  - Koblenz University
  - Joanneum Research Forschungsgesellschaft mbH
  - Informatics and Telematics Institute
  - Dublin City University
  - Centrum voor Wiskunde en Informatica
  - Groupe des Ecoles des Telecommunications
  - Institut National de l'Audiovisuel
  - Institut Eurecom
  - University of Glasgow
  - German Research Centre for Artificial Intelligence (DFKI/LT)
  - Technische University Berlin
  - Ecole Polytechnique Federale de Lausanne
  - University of Economics, Prague

Agenda: Day 1

- Arranged by task
- Time for discussion of approaches & evaluation
- Monday
  - High-level features
  - Lunch
  - Proposal for event detection in surveillance in 2008
  - Video summarization – reports from ACM MM ’07 workshop
  - Demo/Poster previews
  - Demos & Posters
  - Workshop supper

Map: NIST Admin. Building, 1st Floor

(included in the notebook)



Agenda: Day 2

- Tuesday
  - Shot boundaries
  - Search
  - Lunch
  - TRECVID planning
    - More Sound & Vision
      - High-level features
      - Search
    - BBC rushes summarization
    - Event detection in surveillance
    - Other items?

Reminders

- If you are driving to NIST rather than taking the NIST bus, you don't need to stop at the Visitor Center tomorrow.
  - Just show you conference badge and phot ID at the gate as you drive in.
  - Wear your badge at all times while at NIST
- Workshop lunches will be in the main cafeteria.
  - Choose what you want (except bottled or packaged foods)
  - Proceed to the cashier and present your ticket
- The workshop supper is closeby at Smokey Glen in Gaithersburg .This is a very casual restaurant.
  - One ticket is included with your registration
  - You can buy additional tickets at the registration desk
  - If you don't plan to attend, please turn in your ticket at the registration desk so someone else can attend.
- If you are giving a talk, please have your computer connected or presentation loaded before the session begins.

Reminders

- Poster supplies available at the registration desk
- If you are not going to the dinner at Smokey Glen please leave your ticket at the registration desk. If you would like an extra dinner ticket, please check at the registration desk for availability.