Multimedia Information Retrieval at ORL

Dr Kenneth R Wood
ORL
24a Trumpington Street
Cambridge CB2 1QA
ENGLAND

krw@orl.co.uk
http://www.orl.co.uk/
ORL Funding & IPR

- Olivetti from 1986
- Digital 1990-1992
- EU ongoing
- Oracle from 1996

- Founded in 1986
- Andy Hopper, Director
- UK limited company
- 36+ full time research staff

- Cambridge University
- England
- Cracow University
- Poland

- IPR
- £

June 1997
ORL Research Directions

◆ ATM Networking
  – Fixed ATM systems
  – Mobile ATM systems

◆ Multimedia
  – ATM peripheral modules
  – Applications environments

◆ Mobile systems
  – Wearable computers
  – Supporting user mobility

◆ Network computing
  – Stateless terminals
  – Centralised service clusters

◆ Sensor-driven computing
  – Active environments
  – Spatial indexing

◆ Information retrieval
  – Speech and video indexing
  – Multimodal querying/browsing

“Global Personalisation”
ATM Networked Multimedia

Direct Network Peripherals such as cameras, LCD video tiles, discs and audio I/O plug into the lab’s ATM network, enabling experiments with Video Mail amongst other things.
Video Mail Retrieval

◆ **ORL Video Mail Archive**
  - 1000+ messages

◆ **Problem**
  - retrieval means watching all of archive

◆ **Solution**
  - VMR

◆ **Participants:**
  - ORL (*Networked Multimedia Systems*)
  - CUED (*Speech recognition - Steve Young*)
  - CUCL (*Information Retrieval - Karen Sparck Jones*)
VMR Project Goals

- To develop robust unrestricted word spotting algorithms for use in audio and video document retrieval.
- To adapt existing text-based information retrieval techniques to work effectively on voice data.
- To develop and demonstrate a practical system providing video document retrieval using voice.
VMR Storage and Retrieval System

VMR User Interface

Video Mail Browser

Search request

Retrieval Engine

List of Term Indexes

Lattice Scanner

on-line

off-line

New Mail Message Archive

Phone Recogniser

Phone Lattices
User Interface - Requests

Statistical Information Retrieval methods are used to rank potentially relevant video mail for a given request
User Interface - Browsing

In the browser the horizontal timeline shows potential word hits, highlighting each word. Random access playback can be started by clicking anywhere in the timeline.
Follow-on project: MDR

- MDR = Multimedia Document Retrieval
- “son of VMR”
- CUED, CUCL, plus limited ORL & Entropic
- Heavily speech oriented: primary aim is broadcast news retrieval using audio classification and both large-vocab and phone-lattice recognition
Follow-on project: MMIR

- **MMIR** = MultiMedia Information Retrieval
- **ORL only** (6 full-time research staff)
- **Aim is true multimodal indexing, querying, and browsing of heterogeneous collections, using image and video analysis in addition to speech recognition**
- **Cross-fertilisation with MDR likely**
MMIR: Data

- Harvard lectures
- Television News capture
- Online daily newspaper capture (text and images)
- Still images on the web
- Audio and video on the web
- BBC Monitoring Service, Film archives, ...

cf DEC MediaVista
MMIR: Analysis

- **Audio:**
  - classification
  - speech: large vocabulary, phone-lattice

- **Image/Video:**
  - image parsing (edges, regions, colour, texture, shape grammar, ...)
  - shot parsing, key frame detection, motion
Image Parsing

- Multiscale Edge Detection
- Clever Region Growing
- Colour, Texture, Shape
- Quadrant assignment
MMIR: IR & UI

- How to adapt IR models for multimodal indexing?
  - vector, probabilistic, inference net
  - relevance feedback and query expansion

- How to construct UI for multimodal queries/browsing?
  - query construction
  - visualization of (intermediate) results space

These two aspects are intimately related and equally important!
MMIR: System

TSL Npoint 1000 MPEG Camera

Data
- TV
- VCR

Web
- Scanner
- CD-ROM

Text

Oracle 7/8
- Database
- Video Server

CORBA
- omniORB2
- Orbix

IR
- Isearch (CNIDR)
- Inquiry (Sovereign-Hill)
- Okapi (City University)
+ Experimental Mods

Java
- Intel JMF
- OrbixWeb

Analysis
- Text
- Audio
- Image/Video

7-processor Ultra Enterprise Server with Terabyte store (currently 0.1 !)

UI
- Queries
- Results
- Browsing
MMIR: Plans

- **3Q97 sub-goal:** still-image retrieval prototype

- **Ongoing parallel activities:**
  - data collection & digitizing
  - speech training & language model development
  - video shot parsing & key frame detection
  - CORBA component wrappers & transport experiments
  - Oracle database integration
  - IR techniques R&D
  - UI design R&D

- **3Q98 sub-goal:** speech+image video retrieval prototype