Presentation at Petamedia Industry Workshop

Content based video search applications

Some examples from Quaero and Elsewhere

September 30th 2009

Pieter van der Linden,
Program Manager
Thomson
The global picture

- An observation about the market and underlying trends:
  - The volume of universally available digital information has exploded
  - New consumer media (PC, TV, handheld devices, etc.) have proliferated and multiplied
  - Internet becomes the privileged information space.
  - Search tools are the standard for accessing and using content.
    - Video search is a reality today
Why Bother about search

- Because it is useful to the public and to professionals
  - Preferred access means to information on the Internet, Intranet and professional environments.

- And because it generates tremendous economic value

---

Internet advertising

% of 2008 Second-Quarter Revenues
- Digital Video*: 2%
- Banner Ads*: 21%
- Lead Generation: 7%
- E-mail: 2%
- Rich Media*: 7%
- Sponsorship*: 2%
- Classifieds: 14%
- Search: 44%

Total – $5.7 Billion

Source: PWC

Enterprise solutions

<table>
<thead>
<tr>
<th>Forecast (M$)</th>
<th>2006</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>717</td>
<td>1,219</td>
</tr>
</tbody>
</table>

Source: Gartner

In Europe..
Video and Image search
A reality Today

- Internet: the Youtube example
  - #2 for search queries in August 08 - >100% growth year to year.

Search queries (MM) Aug-08

<table>
<thead>
<tr>
<th>Total Expanded Search</th>
<th>17,271</th>
</tr>
</thead>
<tbody>
<tr>
<td>Google Sites</td>
<td>10,158</td>
</tr>
<tr>
<td>Google</td>
<td>7,594</td>
</tr>
<tr>
<td>YouTube/Other</td>
<td>2,564</td>
</tr>
<tr>
<td>Yahoo Sites</td>
<td>2,427</td>
</tr>
<tr>
<td>Yahoo</td>
<td>2,393</td>
</tr>
</tbody>
</table>

Source: Comscore

- Video and Image search
  - A reality Today
Content understanding challenge

Describe into words that are useful for search

*) Pwned: 88700 results on Youtube

Source: Alex Hauptmann
Highlight on video search.  

How much progress is realistic?

- Current deployments mainly rely on keyword search using textual context for web and editorial metadata for media applications
  - Some pioneering applications using audio transcription search from Blinkx, Exalead, Google….

- Over the next years significant progress is expected on base technologies
  - A result of the combination of algorithmic improvement, enhanced methodologies and international cooperation, researcher genius and increasing computer power.
  - Nevertheless technologies are expected to remain far beyond human brain.
  - Some examples:

<table>
<thead>
<tr>
<th>Sample technology</th>
<th>Status</th>
<th>5 year likely improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object recognition</td>
<td>90% success on rigid objects/random results on non rigid objects.</td>
<td>Non rigid objects to catch up with rigid object.</td>
</tr>
<tr>
<td>Scene segmentation</td>
<td>&lt;60% success best case.</td>
<td>20-25% improvement.</td>
</tr>
<tr>
<td>Speech transcription</td>
<td>50-60% success on conversational speech</td>
<td>Improvement by 20-25%.</td>
</tr>
<tr>
<td>Translation</td>
<td>30-35% score BLEU</td>
<td>10% increase.</td>
</tr>
</tbody>
</table>

Most of these technologies have been investigated for a long time*. Targeted solutions combining multimedia analysis, social tags and editorial information, on precise needs or opportunities are expected to succeed.

*) In 67 Marvin Minsky assigned a student to solving the computer vision problem over the summer.
The Quaero Program

- **A collaborative research and development program**
  - Focused on automatic extraction, analysis, classification and use of multimedia, multilingual content
  - To facilitate access to content

- **6 application projects** lead by industrial “champions” aiming at identified business targets
  - So far … other applications may be added later.

- **A shared research structure**
  - A broad research scope
  - Systematic evaluation of scientific and technical progress
  - Extensive resources for annotating large collections of multimedia data
### Six projects with application targets

From content providers to consumers
Sharing resources and know-how

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1-</strong> Digitisation and content enrichment</td>
<td><strong>2-</strong> Digital media asset management</td>
<td><strong>3-</strong> Media monitoring &amp; analysis of social impact</td>
<td><strong>4-</strong> Personalised video</td>
<td><strong>5-</strong> Search engines</td>
<td><strong>6-</strong> PC, Mobile portals</td>
</tr>
</tbody>
</table>

**Who steers**

<table>
<thead>
<tr>
<th>Jouve</th>
<th>Thomson / INA</th>
<th>Yacast</th>
<th>Thomson</th>
<th>Exalead</th>
<th>France Télécom</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>[Image 557x63]</td>
<td>[Image 0x0]</td>
<td>[Image 316x48]</td>
</tr>
</tbody>
</table>

**Expected results**

<table>
<thead>
<tr>
<th>Jouve</th>
<th>Thomson / INA</th>
<th>Yacast</th>
<th>Thomson</th>
<th>Exalead</th>
<th>France Télécom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software and services for editors, patent offices and libraries</td>
<td>Software for broadcasters, media companies, audiovisual archives</td>
<td>Cross media platform and B2B services to analyse media social impact</td>
<td>Software for telecom operators, retailers and enterprise video</td>
<td>Multimedia search engine</td>
<td>New generation of access services to audiovisual content</td>
</tr>
</tbody>
</table>

---

**Shared research structure**

*Coordinated by CNRS and RWTH*

- Technologies for analysing audio, music, image, video content.
- Technologies for natural language analysis and translation
- Content protection technologies

---

*September 30th, 2009*
Following the European approval, the Quaero core developments started in May 2008.

- Research and development program stretching over five years
- A budget of about €200 million
- Assisted by the French State through the public agency
- Involving 24 partners

First year achievements

- Project launch: about 300 people at work
- More than 100 scientific publications
- Exalead and LTU face detection in images service
- Orange 2424actu.fr Beta service
- And several other technology demonstrators available
  - Voxalead, Face recognition, Translation, Image search, Celebrity search
Partners

- Private enterprises
  - Bertin, Exalead, France Télécom, Jouve, LTU Technologies, Synapse Développement, Thomson, Vecsys

- Public research laboratories
  - CNRS-LIMSI, CNRS-IMMI, CNRS-INIST, INRIA, IRCAM, IRIT, Institut Telecom, LIPN, MIG-INRA, Université Joseph Fourier, University of Karlsruhe, RWTH university, Aachen,

- Public institutions
  - BnF, DGA, Ina, LNE

Coordinated by Thomson
Quaero R&D: Close interaction between research and industry

- Industrials and research organizations cooperate actively and effectively to develop demonstrators aiming at identified business targets and increase the state of the art in concerned technology domains.

- **Objective measurement** of results through systematic benchmarking
  - Establishment of a objective assessment on the gap between industrial demand and capacities of technology supply.

- **Strong investment in production of large corpora**
  - Data representative of the target applications sectors
  - Manual and computer assisted annotation
Technology objectives
Facilitate access to content

The issue
Provide the user with *useful information* in spite of the fact that his request is possibly poorly formulated and typically *unanticipated*.

Metadata is key
Exploit available metadata. And enrich it by combining editorial information, automatic annotation and social tagging.
Technology challenges*

- Podcast and conversation speech transcription
- Speaker identification
- Audio track segmentation
- Detection and identification of faces, persons, objects
- Similarity classification
- Handwritten O.C.R
- Names entities detection and classification
- Question Answering (who, why, how)
- Translation of text and speech
- Meter/rythm/key extraction
- Genre/style classification
- Music summarization
- Scene detection
- Action and event tracking
- Person and object tracking
- Multimodal fusion
- Fingerprinting and protection.

*) A subset to be refined and adjusted in accordance to application needs.
Video Search in Quaero

First examples of results

Face tracking and recognition technology by University of Karlsruhe

Hazim Kemal Ekenel

Event detection by INRIA

Ivan Laptev

Voxaleadnews, Audio track transcription and search demo by Exalead.

2424actu.fr, News portal beta by Orange Labs.
Video Classification & Search
A peek into Future TV services.

Using Audience Characterization
- How many
- Who, Gender/Age
- Mood.....

Video recommendation and targeted advertising

And model based classification
Advanced Audio access
A glance on future developments
Many Thanks

More information on

http://www.quaero.eu

(http://www.quaero.org)