Use Case 2 in Media & Communications – Business Cases

Multimedia content Analysis and annotation

KW Partner: CERTH

1 Overview

Challenge
The development and maintenance of large multimedia databases is a difficult task due to the challenges of organising, finding, re-using and distributing multimedia content
Solution
Multimedia is annotated in terms of knowledge extracted from it so that both humans and machines can process large bodies of multimedia more easily
Why a Semantic solution
The Semantic Web offers machine-processable data models for annotating the media and supports semantic search, navigation and reasoning functionalities
Key Business Benefits
Multimedia production and storage is a growing market, particularly among consumers. However, uptake is stalled by a lack of tools to ease media creation and management. Such tools can reduce costs for producers and encourage uptake of media services
Business Partners
Multimedia equipment manufacturers
Network operators
Content providers

Key components

Existing Software
Multimedia ontology management
Security mechanism
Research and Development
Query engine
Media annotation tool
Personalization
Semantic reasoning
Constraint reasoning
Technology locks
Knowledge extraction
Multimedia modelling
Ontology-based reasoning

To address the problems of developing and maintaining large multimedia databases, it is not sufficient to just develop faster hardware or to design more sophisticated algorithms. Rather, a deeper understanding of the information at the semantic level is required. This results in a growing demand for efficient methods for retrieving **semantic information** and **extracting knowledge** from such content, since these are key enabling factors for the management and usability of multimedia content. The goal is to bypass the ineffective and time-consuming process of manual searching and retrieval of multimedia content and use computers to make the content easy to be found and accessible to other parties.

The aceMedia "Integrating knowledge, semantics and content for user-centred intelligent media services[1]" 6th FP Integrated Project focuses on generating value and benefits to end users, content providers, network operators, and multimedia equipment manufacturers, by introducing, developing and implementing a system based on an innovative concept of knowledge assisted, adaptive multimedia content management, addressing user needs. The main technological objectives are to discover and exploit knowledge inherent to the content in order to make content more relevant to the user; to automate annotation at all levels; and to add functionality to ease content creation, transmission, search, access, consumption and re-use. In addition, available user and terminal profiles, the extracted semantic content descriptions and advanced mining methods will be used to provide user and network adaptive transmission and terminal-optimized rendering.

[1] http://www.aceMedia.org

2 Current Practices and Technologies

2.1 Typical business practices

In the following scenario, it is shown how the aceMedia project exploits knowledge inherent to multimedia content and associated textual information to enable new services and user experiences. The example is of a school student who aims to create a retrospective in the form of a multimedia presentation. In order to achieve this goal, she needs the functionalities of multimedia retrieval, annotation and automated organization. Additionally some content involves digital rights and must be paid for. The student's typical use of the system functionality is:

- She wants to create a multimedia presentation including images and videos about her uncle's tennis career. She looks through the textual information associated with local media footage of her uncle to create an overview of his life. From this information, the aceMedia system can select keywords and find further matching media on the Web.
- After paying for and downloading media, the aceMedia system uses the annotations of both the local and purchased media to sort the collected media into themes reflecting different event's in her uncle's tennis career.
- The system's proposed themes can be corrected by the student and the system can use the corrections to propose new themes based on low-level media similarities with the current themes.
- The student performs textual and visual queries to select particular media examples.
- Through built-in concept learning functionality, aceMedia builds up a personal profile of her areas of interest, based on the queries she makes. For instance, facial recognition matching identifies other professional tennis players in the media and the system suggests new themes where her uncle plays against these famous players.
- AceMedia subjects this set to media adaptation, to make it fit on the storage capacity of her aceMedia-enabled mobile terminal, and it is then downloaded onto it.
- On the way to the National News Archives, she transmits media examples of her uncle to the Archive so that automatic searching can start before she arrives. Results scaled down for her terminal are returned and she makes some relevance feedback on them.
- On arrival, she chooses to view the most relevant media results, selects some and purchases them. AceMedia adapts the media for storage on the mobile terminal, and automatically annotates and classifies it. She sends the purchased media to her father who selects some key frames and adds audio and text remarks.
- Finally, she creates a new subset of material based on more specific queries such as "only tennis games mentioned in the press" or "miraculous passes of the ball over the net". From this subset she can now create her presentation and give it at her school.

2.2 System requirements Analysis

Considering the identified use cases in this scenario, we can derive the system requirements:

Annotating Content

The system shall be able to both automatically produce annotations of media from the basis of low level feature analysis and manually allow annotation of the media by the user with the

selection of descriptive terms from controlled vocabularies in order to avoid ambiguity and support system processing. Hence we require

- means to derive concepts from low-level media features, both textual and non-textual
- means to determine the domain of selected media
- means to acquire and use a controlled vocabulary from that domain
- means to suggest to the user appropriate terms from the vocabulary for the annotation

Retrieving Annotations

The system shall be able to retrieve the annotations produced and stored with the media on the basis of queries modelled on the vocabulary of those annotations and present them to the user organized according to the organization of those terms within the controlled vocabulary. Hence we require

- the storage of references between annotations and the media they are annotating
- the extraction of annotations from the content store separate from the media they annotate
- an user interface which supports the user in expressing their query in the vocabulary of the annotations, or a means for the system to interpret a query in respect to that vocabulary
- a presentation component which can represent the relationship of terms from the domain to the user

Retrieving Content

The system shall be able to retrieve media on the basis of a query referring to the annotation of that media, or on the basis of low-level feature analysis, or a mix of both approaches. Media retrieval may be performed on external distributed sources (i.e. on the Web) where annotations may be unavailable or expressed using a different vocabulary. It will also need to take into account the different devices on which retrieved media will be accessed and the need to preserve bandwidth by giving users the option to avoid the retrieval of large media files. Hence we require:

- the internal capability to map between terms in different vocabularies, or from terms to low-level features of media, or from low-level features of media to terms
- automatic low-level feature analysis on media which is found to have no annotation
- the performance of media adaptation to ensure retrieved media can be delivered and displayed to the user on his or her device
- provision of a preview of retrieved media (e.g. thumbnails, brief description, short clips) in order to allow the user to make a pre-selection of desired content before downloading/accessing full media files

Paying for Content

The system shall be able to determine if certain media is copyright protected and hence has some stated restrictions on user access e.g. normally requiring digital payment before it can be downloaded or viewed. It shall also be able to negotiate to have access to the media by meeting the stated access requirements e.g. by making the payment to the copyright owner. In order to avoid any problems, the user shall be kept in the loop and shall have to give permission before the system carries out any operations which affect the user e.g. before payment from the users credit card or accessing restricted content. Hence we require:

- support for digital rights management (DRM), correctly interpreting different access mechanisms and rules possibly expressed in different formats
- the permission of the user to be given before specific access mechanisms are executed e.g. digital payment
- functionality for the support for (secure) communication between different, possibly heterogeneous, digital payment and content access systems

Organizing Content

The system shall be able to use a personal profile of the user, determined (semi-) automatically from the users actions e.g. queries, to carry out a personalized organisation of a set of selected media so that it can be better managed by the user. This management could take place as part of the retrieval functionality (presentation of results) or independently by a request from the user. The content management has the aim of promoting the media most relevant (or of most interest) to the user and to supporting the navigation and location of media through collecting related media into thematic groups. It may also be able to make proposals to the user in terms of new thematic groups of media. Hence we require:

- functionality to determine a user's personal profile through user actions
- means to use that profile to organize media – this means being able to determine relationships between terms in the user profile and terms in the media annotation
- the interpretation of media annotations based on the domain of that media in order to organize the media in a conceptually-related (and hence intuitive to the user) form
- the facility to make proposals to a user of new approaches to organizing media e.g. the creation of thematic sub-folders when there is too much media in a single folder

2.3 Review of the current systems

Current systems that could be reviewed include PC-based media organization tools or the software packaged with PVRs for managing recorded audio and video. However existing systems are not in a position to offer the level of semantic-based functionality described in this scenario.

