Chapter 12
Abstractions in Intelligent Multimedia Databases:
Application of Layered Architecture and Visual Keywords for Intelligent Search

Ranjan Parekh
Jadavpur University, India

Nalin Sharda
Victoria University, Australia

ABSTRACT

Semantic characterization is necessary for developing intelligent multimedia databases, because humans tend to search for media content based on their inherent semantics. However, automated inference of semantic concepts derived from media components stored in a database is still a challenge. The aim of this chapter is to demonstrate how layered architectures and "visual keywords" can be used to develop intelligent search systems for multimedia databases. The layered architecture is used to extract meta-data from multimedia components at various layers of abstractions. While the lower layers handle physical file attributes and low-level features, the upper layers handle high-level features and attempts to remove ambiguities inherent in them. To access the various abstracted features, a query schema is presented, which provides a single point of access while establishing hierarchical pathways between feature-classes. Minimization of the semantic gap is addressed using the concept of "visual keyword" (VK). "Visual keywords" are segmented portions of images with associated low- and high-level features, implemented within a semantic layer on top of the standard low-level features layer, for characterizing semantic content in media components. Semantic information is however predominantly expressed in textual form, and hence is susceptible to the limitations of textual descriptors – viz. ambiguities related to synonyms, homonyms, hypernyms, and hyponyms. To handle such ambiguities, this chapter proposes a domain specific ontology-based layer on top of the semantic layer, to increase the effectiveness of the search process.

DOI: 10.4018/978-1-61350-126-9.ch012

INTRODUCTION

Over the last couple of decades a number of multimedia applications such as: digital photo albums, computer based training, games and entertainment, online galleries, medical applications, and information kiosks have led to the growth of large repositories of digital media. Muller (2004) states that "The Radiology Department of the University Hospital of Geneva alone produced more than 12,000 images a day in 2002" (p. 1).

In this scenario a fast and efficient search and retrieval mechanism from these repositories assumes a fundamental importance, as a repository without a retrieval mechanism is comparable to a library without a catalog – even though all the information is present, it is practically inaccessible to somebody with a specific search criteria.

Therefore, the issue of having a multimedia repository that can be searched efficiently has become of paramount importance, in other words, "The question we now need to answer is how to build a multimedia database around a multimedia repository" (Baral, 1998, p. 38).

Otherwise, "Much similarly to the case of books in a library that have not been indexed, information stored in a multimedia archive that cannot be searched, identified, and accessed easily is practically unavailable" (Wallace, 2006, p. 34).

Earliest attempts in building multimedia repositories amenable to search and retrieval has been though textual annotations. However, many drawbacks to this approach soon became apparent; some of these include:

- It requires manual annotation and processing, hence it becomes time consuming.
- Search results are dependent on exact matching of text strings, hence the presented results are not complete.
- Search efficiency is limited by subjective nature of textual descriptions; hence it can lead to false and missed matches.
- Support for Query-by-Example (QBE) is not possible; hence the user needs to be good at formulating the query.

To overcome these limitations the research community focused on Content Based Storage and Retrieval (CBSR) techniques, in which features are extracted directly from the media files by automated algorithms, and used as metadata for their subsequent search and retrieval. A number of research prototype systems have been developed, for example: QBIC (Niblack, 1993), CORE (Wu, 1995), PhotoBook (Pentland, 1996), VisualSEEK (Smith, 1996), and the Digital Library Project (DLP) of the University of California (Belongie, 1997).

However, there are still a number of unresolved research issues, such as the following three:

Compatibility: Various systems developed independently are potentially incompatible with each other. Therefore, there is a need for integration with a uniform approach and a single point of access. However, this is not easy as pointed out by Liu (2001, p. 235): "Since each kind of multimedia data … has its own characteristics, it is very difficult to develop a uniform method for content-based retrieval." How to achieve a single point of access is also not fully resolved, as articulated by

Wallace (2006, p. 34): "The development of single points of access, providing common and uniform access to their data, despite the efforts and accomplishments of standardization organizations, has remained an open issue …"

Layered Architecture: Since multimedia metadata can belong to various levels of abstractions, there is a need for a layered or hierarchical approach to data analysis for a better understanding of media content. Comaniciu (2002, p. 603) points out that: "To improve performance, the execution of low-level tasks should be task driven i.e. supported by independent high level information."

One of the most widely used multimedia metadata standard is MPEG-7, however,