This is the published version (version of record) of:


Available from Deakin Research Online:
http://hdl.handle.net/10536/DRO/DU:30004562

Reproduced with the kind permission of the copyright owner.

Copyright : 2001, International Society for Computers and Their Applications
Post-Processing of Multimedia Information—Concepts, Problems, and Techniques

Chengqi Zhang and Zili Zhang
School of Computing and Mathematics
Deakin University, Geelong Victoria 3217, Australia
e-mail: {chengqi, zili}@deakin.edu.au

Abstract

Currently, most research work on multimedia information processing is focused on multimedia information storage and retrieval, especially indexing and content-based access of multimedia information. We consider multimedia information processing should include one more level—post-processing. Here “post-processing” means further processing of retrieved multimedia information, which includes fusion of multimedia information and reasoning with multimedia information to reach new conclusions. In this paper, the three levels of multimedia information processing—storage, retrieval, and post-processing—are discussed. The concepts and problems of multimedia information post-processing are identified. Potential techniques that can be used in post-processing are suggested. By highlighting the problems in multimedia information post-processing, hopefully this paper will stimulate further research on this important but ignored topic.

Keywords: Multimedia, Content-Based Retrieval, Multimedia Information Fusion, Reasoning

1 Introduction

The multimedia revolution started at the beginning of 1990’s. Multimedia is virtually revolutionary in many areas such as communications, computing, entertainment, consumer electronics etc. Although there are still many challenging problems remain to be researched and resolved for the further growth of multimedia, with the advent of relatively cheap, large online storage capacities and advances in digital compression, comprehensive sources of text, image, video, and audio etc. multimedia data can be stored and made available for research and applications. Actually, the technical ability to generate volumes of digital multimedia data is becoming increasingly “mainstream” in today’s electronic world.

With numerous digital libraries, other large multimedia databases, and multimedia-based WWW page available, sophisticated search or retrieval technique are required to find relevant information in these large digital data repositories. These retrieval technique should provide not only fewer bits but also the right bits to users.

We notice that to retrieve the right bits from multimedia data repositories is not the final step in multimedia applications. Further processing of the retrieved multimedia information is needed. In this paper, we call such multimedia information processing post-processing. For example, in our ongoing agent based financial investment adviser project [1], we need to retrieve stock market information. These data include breaking news in audio or text form that affect stock market, closing price of specific securities (table), security price or moving average chart, company revenue (pie chart), company profiles etc. In the financial investment advising multi-agent system, different agents are delegated to gather information in different forms. The decision making agents will make financial investment decisions based on their domain knowledge and the retrieved multimedia information. To this end, we come across two problems: How to fuse information in different media forms and how to reason with the multimedia information to obtain new results. These “how to”s are most important topics in multimedia information post-processing.

In this paper, we propose that multimedia information processing be divided into three levels—multimedia information storage, retrieval, and post-processing. Currently, most work on multimedia information processing is focused on multimedia information storage and retrieval, especially indexing and content-based access of multimedia information. Post-processing of information is nearly ignored. With this observa
The concepts of multimedia information post-processing are identified. The problems involved are highlighted. Promising techniques to solve these problems are suggested. Hopefully, the discussions in this paper will stimulate further research on this topic.

The rest of the paper is organized as follows. The three levels of multimedia information processing are presented in Section 2. Brief summaries of the state of the art of indexing and content-based access of multimedia information are given in Section 3 and 4, respectively. The concepts and problems of multimedia information post-processing are investigated in Section 5. Finally, Section 6 is concluding remarks.

### Three Levels of Multimedia Information Processing

In today's electronic world, there are volumes of digital multimedia data generated every second. The first thing we need to do is to store these data for later use. As the amount of multimedia data is increasing dramatically, some effective and efficient retrieval techniques are needed to find useful and relevant multimedia information from these large digital data repositories such as multimedia databases and multimedia-based WWW pages etc. In some applications, to find the relevant information is the final goal, but in others not. Just like in traditional expert systems, some "known facts" must be given before the systems can reach some conclusions. In many multimedia application systems, the retrieved multimedia information is used as "known facts". Based on these "known facts" and domain knowledge, the systems try to obtain final conclusions the users want. Thus, in typical multimedia information processing, there are three levels: storage, retrieval, and post-processing of multimedia information (Figure 1).

Currently, most of research work in multimedia information processing is concentrated on multimedia information storage and retrieval, especially effective indexing of multimedia data and content-based access to multimedia information, whereas the post-processing is nearly untouched. In the next two sections, we will give brief summaries of indexing of multimedia information and content-based multimedia retrieval. We will then identify the concepts and problems of multimedia information post-processing in Section 5.

#### 3 Indexing of Multimedia Data

To store large volumes of multimedia data, in addition to the requirements of very high capacity, fast access times, and high transfer rates, another very important aspect is effective indexing. This is also closely related to multimedia information retrieval.

Indexing is a fundamental operation for large databases. Index structures are closely related to data representation. This implies that we should have different index structures for data in different media forms such as image, video, and audio. Some typical currently available indexing methods include indexing of string attributes [9], indexing of visual attributes [9], and active indexes [8] etc.

### 4 Content-Based Multimedia Information Retrieval

Information retrieval has been attracting the attention of many researchers. Now in this field, more emphasis was placed on the content-based retrieval and post-processing of multimedia information. Content-based multimedia information retrieval contains two parts: content-based retrieval for multimedia databases and content-based access to WWW multimedia information.

In content-based retrieval for multimedia databases, there are two principal ways for the representation of queries, namely, "query-by-subject/object" and "query-by-example" (QBE) [4]. Currently, many more studies have been done in relation to content-based retrieval that refers to a single non-textual data. However, we think content-based retrieval studies for multimedia databases should pay more attention to the multimedia content that is associated with heterogeneous types of data.

For content-based access to web-based multimedia information, the most interesting work is the Content-Based Access to Multimedia Document effort (CBAM)
of Microelectronics and Computer technology Corporation (MCC) [7]. The CBAM effort is an outgrowth of the MCC InforSleuth project, wherein technologies have been developed for using agents to achieve semantic access to networks of heterogeneous and diverse information sources. The key to accomplishing CBAM applications in this manner is the semantic content extraction algorithms encapsulated in multimedia resource agents and the advancement of ontologies for content correlation to better handle temporal, spatial, and uncertain aspects of multimedia content analysis.

5 Post-Processing of Multimedia Information

It is worth noting that there are many researchers working on (multimedia) information retrieval (especially content-based retrieval), whereas the processing of retrieved multimedia information is nearly ignored. Here, we try to highlight the problems so as to trigger further research on this interesting field. We consider that the most important topics in multimedia information post-processing are fusion of multimedia information and reasoning with multimedia information.

Fusion of Multimedia Information. Fusion of information is about how to use simultaneously pieces of information provided by several sources in order to come to a conclusion or a decision. Currently, most approaches in information fusion can only deal with information in alphanumeric form. In many multimedia application systems, we may have information in different media forms such as text, image, audio, and video to describe the unique object.

Intuitively, multimedia information fusion should be based on some effective transformation among information in different media forms and media abstractions. Some work has been done on transformation [7] and media abstractions [10]. A framework for multimedia information fusion is shown in Figure 2. In Figure 2, the transformations from data to abstracted representations are indicated by circles.

As illustrated in Figure 2, multimedia information fusion is feasible when information from different sources can be converted into similar representations (indicated by several circles in the same horizontal row). The fusion then takes place among assertions. We call such fusion horizontal fusion because it combines information abstracted from different media encoded in the same uniform representation. Another type of fusion is applicable to data from similar media with different abstracted representations so that they can be combined and checked for consistency etc. We call such fusion vertical fusion because it combines information having different representations at different levels of abstraction.

Horizontal fusion can be accomplished with the help of an artificial neural network due to its ability to combine information abstracted from different media and adequately encoded in the same uniform representation. The active index [8] can be used in vertical fusion due to its ability to obtain information from different sources and actively connecting them by dynamic linking (using index cells). Vertical fusion is associative and combines information in different representations. An artificial neural network with fixed connections is not as appropriate as an active index with flexible connections.

Reasoning with Multimedia Information. As we have pointed out that in many multimedia applications, to retrieve the relevant information is not the final step of these applications. We need to use the retrieved information (in different media forms) as known facts to do some reasoning based on domain knowledge. As these known facts are represented in different media forms (not only in alphanumeric form), traditional reasoning algorithms cannot be applied directly. We call such kind of reasoning reasoning with multimedia information. It is different from visual reasoning and spatial reasoning [5] (chapter 11) etc. used in multimedia information retrieval (especially image retrieval).

To solve the problems of reasoning with multimedia information requires the contribution of many research disciplines such as image/video analysis and processing, pattern recognition and computer vision, multimedia data modeling, multidimensional indexing etc. The following technologies are promising to partially solve this problem: (1) Effective transformations among different media; (2) Symbolic projection theory [5]; (3) The Tele-Action Object (TAO) techniques [8].

Example. Recall the financial investment adviser project we mentioned in Section 1. To give advice about stock buying/selling we need to do some reasoning based on the moving average chart of a specific security as well as many other analyses. There are two main trading rules for moving averages: (1) A buy signal is given when the price moves up and crosses over the moving average from below; (2) A sell signal is given when the price moves down and crosses over the moving average from above.

Here, it is more natural and convenient to represent the condition parts of these rules as well as the "known facts" (retrieved moving average charts of some specific securities) by graphics (charts). This is one form of reasoning with multimedia information. For this specific problem, it is actually a problem of reasoning with still images. We employ symbolic projection theory to do
with reasoning. We represent the moving average chart 2D strings in symbolic projection, and then use 2D string matching algorithm to accomplish the reasoning. The details will be discussed in another paper.

### Concluding Remarks

As we noticed that to retrieve the right bits from multimedia sources is not the final step in many multimedia applications, we proposed that multimedia processing be divided into three levels—multimedia information storage, retrieval, and post-processing. Currently, most of research in multimedia processing is focused on the first two levels, whereas the post-processing is nearly untouched.

We then identified the concepts and problems in multimedia information post-processing. Two most important topics—multimedia information fusion and reasoning with multimedia information—were highlighted. Promising techniques to solve the problems in multimedia information post-processing such as effective transformations, active index, symbolic projection theory, and TAO were outlined.

To tackle the problems in multimedia information post-processing requires the contributions of many research disciplines. We have conducted some research work on reasoning with still image information by using symbolic projection theory. We hope the discussions in this paper will stimulate further and more research on this interesting topic.

### References


