Metadata Transformation and Management with Oracle interMedia

Marco Carrer, Ashok Joshi, Paul Lin, and Alok Srivastava
{*mcarrer,ajoshi,pilin,alsrivas*}*@us.oracle.com*
Oracle Corporation
One Oracle Drive
Nashua, NH 03062

Abstract

The management of multimedia data in object-relational database systems presents a challenging problem for extracting, processing, and managing associated metadata. This information is usually embedded within the media source using proprietary formats, thus not easily accessible in a uniform fashion. This poses a need for a series of structural transformations to ensure that the metadata gathering process produces a unified representation across a multitude of media sources. The ultimate goal of these transformations is to consolidate management efforts for diverse sources of media.

This paper presents the Oracle MediaAnnotator, an extensible architecture developed to be used with Oracle8i interMedia. The MediaAnnotator supports automatic extraction and transformation of metadata into logical annotations. This approach allows for the creation of unified metadata repositories, which can then be used for indexing and searching. The extensible nature of this architecture makes it applicable to any multimedia domain. The MediaAnnotator leverages Oracle8i interMedia to tie together the multimedia data and its logical annotations; this offers greater manageability of media archives and opens the possibility for new applications which integrate multimedia content with other user data.

1 Introduction

Digital multimedia yields a format that is very different from alphanumeric data. While textual information supports the concepts of alphabetical ordering, indexing, and searching, media data does not. Multimedia formats are designed to fulfill playback rather than manageability requirements. In order to add manageability, it is imperative that metadata be associated with the media [2, 3, 4, 5, 6, 9]. The approaches presented in the literature support the idea that, in order to achieve fast data retrieval through queries, media information must be extracted – automatically or manually – from the raw media data and stored in a different, more readily usable format, which will constitute the input for the query engines. This transformation process is known as *media annotation*. The transformation into annotations – which involves a special processing of a set of inputs to provide an enriched set of contents – allows the employment of conventional database systems to manage digital multimedia.

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

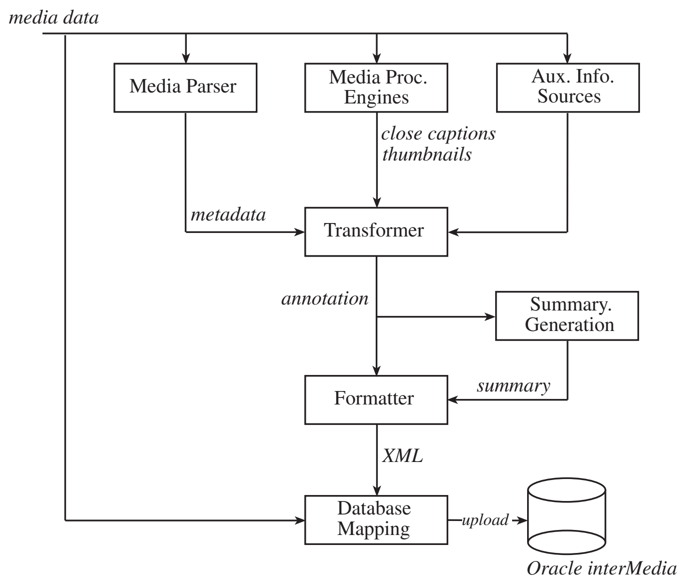

Figure 1: The MediaAnnotator Architecture

The development of Oracle MediaAnnotator has been motivated by the abundance of information already embedded into media sources. This consideration is validated by popular media formats such as QuickTime, from Apple Computer [1], and Advanced Streaming Format (ASF), from Microsoft Corp. [7], both of which are designed to allow for capturing user's as well as system annotations within the format. This information is generally added during media creation and editing. In addition, emerging standards such as Digital Versatile Disc (DVD) and MPEG-7 support even richer metadata, which facilitates the design of automated media management solutions. However, there is no unified way of capturing and using this metadata in applications. The MediaAnnotator addresses this problem by capturing proprietary metadata and transforming it into logical annotations. Its architecture also enables the capturing of associated data which are not available within the media.

Section 2 covers the terminology that will be used throughout this paper. Section 3 offers a detailed discussion of the Oracle MediaAnnotator and, finally, Section 4 concludes the paper.

2 Terminology

For simplicity, we define the following terminology used throughout this paper: raw digital video and audio data are collectively referred to as media data. Descriptions for media data will be referred to as metadata. Metadata and annotations are used interchangeably in this paper identifying a collection of easily indexable attributes (media title, copyright, etc.) or properties.

3 Oracle MediaAnnotator

3.1 Architecture Overview

Oracle MediaAnnotator addresses the need to manage, index, and search digital multimedia by capturing and organizing the associated metadata, which semantically describe media content. This information is captured and re-structured into a searchable form which is suitable for management with Oracle *inter*Media [8].

The MediaAnnotator provides a framework to capture, transform, process, and store metadata. This framework is extensible to allow for user-defined components to be plugged in at every step. Fig. 1 gives a graphical overview of the framework. As shown in the diagram, media data flows through the system along multiple paths. The parser takes media data to extract metadata while the media processing engine generates information based upon media contents. The generated information is then combined with any auxiliary metadata sources, and transformed into annotations. At this stage a summary of the media data can also be generated based on the information captured so far. After being formatted into a searchable representation, annotations and the original media data are uploaded into the database.

The following sections offer detailed descriptions of the components of the system depicted in the Fig. 1.

3.2 Media Parser

The media parser extracts metadata embedded in the digital media according to the file format specifications, which contain instructions on how to parse and extract this information. The MediaAnnotator uses the mimetype of the media source to dynamically load the appropriate parsers for the media. This enables user-defined parsers to be plugged into the MediaAnnotator framework at run-time, thereby extending the range of media formats handled by the system.

3.3 Media Processing Engine

The media processing engine is responsible for generating additional information by analyzing the actual media content. The output of this engine is often time-based; closed-captions, thumbnails, embedded links (URL flipping), and sample clips are some examples. Collectively this information can be viewed as a set of time-based snapshots of the media. Consequently, advanced queries can be performed, producing results with time stamps which can then be used to seek to specific positions within the media.

3.4 Auxiliary Metadata Sources

Auxiliary metadata sources provide the information which is not obtainable by processing the media itself. For example, audio compact discs do not carry any meta-information along with the physical media; it is therefore necessary to gather metadata from auxiliary sources such as user's input, or look-up services on the Internet.

3.5 Transformer

The transformer combines the media information collected thus far to construct unified logical annotations, which contain attribute value pairs as well as time-based samples, describing the media. For example, the logical annotation for an audio compact disc will feature attributes such as the title, artist, duration, and number of tracks as well as audio clips for each track.

The primary task here is to organize the captured attributes and partition them semantically. Each resulting partition constitutes a logical annotation. This structuring of the metadata provides a facilitated method of managing the media. In particular, the effect of such a transformation is an abstraction layer above the diverse multitude of media formats. The client of annotations is shielded from having to understand the format and storage specifications of the original media source. It is now possible to manage the data in a semantically rich manner.

The MediaAnnotator allows users to override the predefined set of annotations or define a completely new set. Similar to parsers, annotations are dynamically loaded based upon the mimetype of the media.

3.6 Summary Generator

Logical annotations can be processed to generate a summary of the media data. The summary generator accomplishes this task according to user's specified guidelines. For example, a song can be summarized by grouping together the performer's name, the song title, and a song clip. The summaries are especially useful for quick browsing of media catalogs.

3.7 Formatter

The formatter is responsible for transforming the logical annotations as well as the summaries into a form which is searchable and manageable by databases. A well-defined XML structure is used by the MediaAnnotator to store this information, hence a unified representation for the metadata is achieved.

3.8 Database Mapper

Database Mapping is the final step of the transformation chain and completes the database population process. During this step, the MediaAnnotator uploads the media and the associated XML document, produced by the formatter, into the Oracle database. Database mapper leverages the media support offered by Oracle *inter*Media, the design of which allows for simultaneous storage of the actual media data and its corresponding metadata. The MediaAnnotator maps the physical properties captured in a logical annotation, to fields of an *inter*Media object. In addition, the XML representation, which includes content attributes, is also stored within the object. As a result, a self-contained repository, for the media data and its description, is created in the Oracle database. This repository can now be indexed with *inter*Media indexing techniques [8], enabling advanced searches on the multimedia data. The subject of indexing techniques is beyond the scope of this paper and is not discussed.

4 Conclusion

In this paper, we have presented an architecture which addresses the problem of managing multimedia data through a set of transformations. The extensibility of the MediaAnnotator makes it applicable to a wide variety of media domains. The MediaAnnotator is extensible in both the understanding of new media formats, and the grouping of attributes into meaningful logical annotations. The media and its logical annotations are stored into a single uniform repository within an Oracle database. This final transformation to a database stored structure, such as an Oracle8*i* *inter*Media object, takes advantage of the built-in textual and object management functionality of an OR-DBMS and completes the process of a semi-automated solution for media asset management.

References

[1] Apple Computer Inc. *QuickTime*. Inside Macintosh. Addison-Wesley Publishing Company, 1993.

[2] K. Böhm and T. C. Rakow. Metadata for Multimedia Documents. *SIGMOD Record*, 23(4):21–26, December 1994.

[3] M. Carrer, L. Ligresti, G. Ahanger, and T.D.C. Little. An Annotation Engine for Supporting Video Database Population *Multimedia Tools and Applications*, 5(3):233-258, November 1997.

[4] G. Davenport, T. A. Smith, and N. Pincever. Cinematic Primitives for Multimedia. *IEEE Computer Graphics and Applications*, pages 67–74, July 1991.

[5] W. I. Grosky, F. Fotouhi, and I. K. Sethi. Using Metadata for the Intelligent Browsing of Structured Media Objects. *SIGMOD Record*, 23(4):49–56, December 1994.

[6] W. Klaus and A. Sheth. Metadata for Digital Media: Intruduction to the Special Issue. *SIGMOD Record*, 23(4):19–20, December 1994.

[7] Microsoft Corporation *Advanced Streaming Format (ASF) Specification*. Public Specification Version 1.0, 1998.

[8] Oracle Coporation *Oracle8i interMedia Audio, Image, and Video User's Guide and Reference* Release 8.1.5 (A67299-01), 1999.

[9] Y. Tonomura, A. Akutsu, Y. Taniguchi, and G. Suzuki. Structured Video Computing. *IEEE Multimedia*, pages 34–43, Fall 1994.