Chapter XI
Tools for Automatic Audio Management

Marko Helén
Tampere University of Technology, Finland

Tommi Lahti
Nokia Research Center, Finland

Anssi Klapuri
Tampere University of Technology, Finland

ABSTRACT

The purpose of this chapter is to introduce tools for automatic audio management. The authors present applications which are already available for the users and describe the algorithms and methods behind these applications and their performance. They also discuss the concept of metadata, which is an important prerequisite for modern distributed personal content applications. The variety of automatic audio management tools is wide-ranging. This chapter covers audio segmentation and classification, query by example of audio, music retrieval and recommendation, and speech management, which they consider as being the most important aspects of audio information management. Computational complexity is one major concern in the present era of personal mobile devices and large multimedia collections available on the internet. Therefore they also introduce clustering and indexing techniques which are developed for faster access in large databases.

INTRODUCTION

The huge amount of digital multimedia available nowadays has created a growing need for automatic tools to manage all this data. People have their own personal image, video, and audio databases collected

for example by mobile phones or digital cameras that are always close at hand. Above all these, there is the internet with an enormous amount of information accessible almost free of charge. Managing all this becomes impossible without the proper tools.

The most typical operations with personal databases are organizing the samples or searching a certain kind of sample from the database. Samples can be organized based on several criteria. In the simplest case, they can be arranged in an alphabetical filename order or in a chronological order using the time stamps of the files. A more challenging way to organize the samples is to do it based on their actual content. For example, it is possible to organize speech samples by the speaker or music by the genre or artist. This is called content based classification (Liu, 2007; Lu, 2002) or clustering (Cai, 2008), and will be covered later in this chapter.

Finding similar samples from a database, referred to as query by example, is an operation where the user wants to find samples that resemble a particular query sample. Again, there are different criteria for similarity, and even in the case of content based similarity, one has to define in which sense similarity is measured. One example is *query by humming* (Lu & Zhang, 2001; Song et al., 2002), which retrieves melodically similar samples to an input humming. Another is *query by beat boxing* (Kapur, 2004), which retrieves rhythmically similar samples to the example.

A common denominator in all automatic content analysis techniques is the need for an accurate similarity measure. A number of different similarity measures have been used in the literature (Kashino et al., 2003; Zhou & Hansen, 2000). The choice between them is made based on the task and the requirements like speed, accuracy, or whether temporal changes should be taken into account. Recently, the distance between probability density functions (pdfs) estimated from frame-wise feature vectors is often used as a similarity measure (Helén & Virtanen, 2007; Hershey & Olsen, 2007). Several pdf-based distance measures will be introduced in this chapter.

An important factor in multimedia information management is the concept of metadata. The term refers to structured data that characterizes the information-containing entities. Metadata may include for example spectral features extracted from an audio signal but also the relations that different samples have with each other like the above-mentioned similarity of samples. Metadata is the base on which the multimedia management applications are built.

In this chapter, we first discuss the challenges in handling metadata and how the metadata is extracted from audio samples. Then, we introduce applications which utilize the metadata to perform certain audio management tasks, and finally discuss the future trends in this field.

BACKGROUND

Trend Towards Distributed Personal Content Applications

Content of the data is what makes it valuable for its owner. Therefore, helping the users to find, organize, and share the data by its content is naturally central for the application design. Traditionally, and looking just a couple of years back, the systems utilizing automatic audio content analysis techniques have been more or less stand-alone applications targeted for professional use. Transcribing broadcast television and radio news, for example, has been studied extensively already over a decade (Cook et al., 1997; Gauvain et al., 2002). Companies have also used various speech recognition applications in their automatic feedback systems.

20 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/tools-automatic-audio-management/27798

Related Content

Survey on Prioritizing Test Cases in Various Levels of the Software Development Life Cycle
Gayatri Nayakand Mitrabinda Ray (2021). *International Journal of Information Technology Project Management (pp. 1-28).*
www.irma-international.org/article/survey-on-prioritizing-test-cases-in-various-levels-of-the-software-development-life-cycle/269422

ICT-Supported Gaming for Competitive Intelligence
Jan Achterbergh (2005). *Encyclopedia of Information Science and Technology, First Edition (pp. 1379-1384).*
www.irma-international.org/chapter/ict-supported-gaming-competitive-intelligence/14442

An Approach to Optimize Multicast Transport Protocols
Dávid Tegze, Mihály Orosz, Gábor Hosszúand Ferenc Kovács (2009). *Encyclopedia of Information Science and Technology, Second Edition (pp. 206-211).*
www.irma-international.org/chapter/approach-optimize-multicast-transport-protocols/13574

Topaz Japanese-American Relocation Center Digital Collection: A Case Study
Lindsay Lauridsen (2014). *Cases on Electronic Records and Resource Management Implementation in Diverse Environments (pp. 117-129).*
www.irma-international.org/chapter/topaz-japanese-american-relocation-center/82643

Patents and Standards in the ICT Sector: Are Submarine Patents a Substantive Problem or a Red Herring?
Aura Soininen (2008). *Information Communication Technologies: Concepts, Methodologies, Tools, and Applications (pp. 345-382).*
www.irma-international.org/chapter/patents-standards-ict-sector/22673