Entertainment Computing 9–10 (2015) 11–18

Contents lists available at ScienceDirect

Entertainment Computing

journal homepage: ees.elsevier.com/entcom

SEMU: A semantic knowledge management system of musical information ☆

Dinh Hoa Cuong Nguyen, Ngamnij Arch-int *, Somjit Arch-int

Department of Computer Science, Faculty of Science, Khon Kaen University, Thailand

ARTICLE INFO

Article history:
Received 1 July 2014
Revised 5 June 2015
Accepted 10 June 2015
Available online 24 June 2015

Keywords:
Semantic Web
Ontology
Knowledge management
Musical Information Retrieval

ABSTRACT

The diverse data types of musical information domain including binary and text-based structures create semantic gaps between the entities of different data formats. This leads to difficulties in analyzing, capturing and managing entities of the domain. In this paper, we present a semantic knowledge management system, called SEMU, to efficiently managing musical information. We propose SEMU ontology to capture information extracted from various data types and sources. In order to extract information from raw data, we use Musical Information Retrieval techniques for audio files and Natural Language Processing techniques for text-based formats. We develop a rule-based solution to enrich the system knowledge base. Later, we provide a web application with seamless integration between SEMU knowledge base and user interface to enable users to benefit from the advantages of the SEMU system.

1. Introduction

Building knowledge-based system of musical information domain is a challenge due to the diversity of data formats. While music attracts listeners by its sound characteristics (like timbre, pitch or melody), people use natural language to talk about music and its related entities such as moderato, pop, singer or album to name a few. The complexity increases as musical products have been recorded and distributed in various formats (mp3, wav, au, etc.) and natural language descriptions have been stored in text-based formats. These have created a semantic gap between audios and texts within the musical information domain. In web space, this gap becomes larger with thousands of distributed musical websites. Therefore, reducing this semantic gap and building a musical knowledge base from these diverse data sources introduce challenging problems. To deal with these problems, our approach is based on Semantic Web technology, Musical Information Retrieval and Natural Language Processing techniques.

Musical Information Retrieval (MIR) is an emerging research area which focuses on retrieving information from musical audio files [e.g., 1–4]. In this field, music classification is an important problem with the target to automatically label genres of songs. A number of research works [5–8] applied supervised methods of machine learning to link audio features with related words in natural language. However, these words are not managed by a semantic structure. Hence, combining classification results and textual information about music into a knowledge base will create semantic linkages among entities residing in audios and texts.

Within the domain of music information, there are many entities that can be classified to different classes, for example, singers, composers, albums, or music bands. In web space, this information is mainly distributed in hypertext documents of different websites. Therefore, building a knowledge-based system, which is based on Semantic Web technology, can manage these entities efficiently as well as integrating the information.

In this paper, we devise SEMU – a semantic knowledge management system of music – coping with the problem of integrating musical information from multiple data formats and sources. The SEMU system is built upon cutting-edge Semantic Web technologies (e.g., ontology management, RDF persistent storage, reasoning and rule base). The semantic gap between audios and texts is reduced by combining MIR technique, Semantic Web technologies and Natural Language Processing (NLP) toolkits in analyzing and processing these kinds of data. In addition, we develop a web application, which focuses on seamless integration between knowledge base and the user interface, to enable end-users utilizing the advantages of SEMU system.

The rest of paper is organized as follows: Section 2 describes the literature review. In Section 3, we introduce our methodology including the conceptual framework of SEMU system (subsection

☆ This paper has been recommended for acceptance by Owen Noel Newton Fernando.

* Corresponding author. Tel.: +66 824 619 721.
E-mail addresses: ndhcuong@hce.edu.vn (D.H.C. Nguyen), ngamnij@kku.ac.th (N. Arch-int), somjit@kku.ac.th (S. Arch-int).

http://dx.doi.org/10.1016/j.entcom.2015.06.002
1875-9521/

3.1); the SEMU ontology (subSection 3.2); the audio processing and information extraction modules are explained in subSections 3.3 and 3.4, respectively; and the reasoning approach and rule base of SEMU system are discussed in subSection 3.5. In Section 4, we present the implementation of SEMU system. Finally, Section 5 concludes and highlights future work.

2. Literature review

2.1. Musical Information Retrieval

Audio data are popular web-wide. This kind of data has led to many research problems, for instance, genre classification, instrument identification, or copyrights. MIR studies solutions to above problems. MIR retrieves the information from binary data of audio files by applying signal processing and machine learning techniques.

MIR researchers use audio features, which are extracted from audio files by signal processing techniques, as the input of their studies. According to Fu et al. [2], audio features are divided into two layers. Low-level features were at the bottom layer including two classes of timbre and tempo. Timbre captures sound quality whereas the variation and evolution of timbre are described by temporal features. Based on these characteristics, low-level features have been used for music classification. The frequently used features of timbre include Zero Crossing Rate (ZCR), Mel-frequency Cepstrum Coefficient (MFCC), Linear Predictive Coefficient (LPC), Linear Predictive Cepstrum Coefficient (LPCC), Spectral Centroid, Spectral Flux, and Amplitude Spectrum Envelop. Statistical Moments, Amplitude Modulation, and Auto-Regressive Modeling are features of tempo class. The top layer holds middle level features like rhythm, pitch, and harmony which are usually extracted from low-level features.

In MIR research, musical genre classification is one of the most important problems. To solve this problem, supervised methods of machine learning are dominant approach. Generally, the process has four steps. As a first step, audio features are extracted from files. Next, trained datasets are built up. Then, the classifiers are trained with the target of reaching highest accuracy. Lastly, the trained classifiers are used to classify new songs. Different kinds of classifiers have been applied to the problem of music classification or audio discrimination. In the study of Fu et al. [1], audio features are extracted from local windows as trained datasets for Naive Bayes classifiers. Artificial neural network is also used to discriminate kinds of audio files as described in [8]. The study of Wu et al. [7], which deals with speech/music discrimination, uses decision trees to label the audio input. Huang et al. [9] propose a self-adaptive harmony search algorithm to derive the local feature set which are extracted from the original feature set including five acoustic characteristics (i.e., intensity, pitch, timbre, tonality, and rhythm). Then support vector machines (SVM) are used to classify each musical recording. SVM is the most popular classifier that are used in [4–6,10–12]. Although SVM method runs slower than other classifying techniques, the accuracy of this classifier is higher than those of others.

2.2. Semantic Web application and musical information domain

The Semantic Web vision first introduced by Tim Berners-Lee et al. [13] is aimed at extending the current web by providing semantic data with the key component named ontology. This kind of data is machine understandable. Therefore, interesting scenarios of Semantic Web applications have been defined (e.g., information integration, automatic processing data among information systems, knowledge base management). As part of Web data, the value of musical data can also be enhanced by applying Semantic Web technology into the musical information domain, especially in building musical knowledge base.

Although there are huge amount of audio data web-wide, the number of studies about building knowledge base for this kind of data is smaller than those of textual data. Karpouzis et al. build a system based on MPEG-21 framework [14] to handle both binary information and structured representations of digital items. Another study of Döller and Kosch [15] introduces a metadata model for multimedia content relying on XML-based MPEG-7 standard. This model, however, has the limitations in inference capability. Both studies of Karpouzis et al. [14] and Döller and Kosch [15] can manage digital content well but these systems have limitations in reasoning. Oscar Celma and Xavier Serra use Semantic Web technology to solve the problem of music recommendation as presented in [16]. They collect musical information from various web sources to enrich the system knowledge base. In [16], the problem of music recommendation is solved by using FOAF ontology and RSS vocabulary. The solution only focuses on textual data while analyzing audio signal is ignored. Rahman and Siddiqi [17] propose a music annotation ontology, which is named mpeg-7Music, mainly supporting music producers to publish the musical contents.

Ontology-based information systems have been developed in different fields. In which, the problem of ontology population plays the key role in each system. There are different approaches to enrich the knowledge base. For example, in [18], the authors propose a semantic ontology mapping for interoperability of learning resources systems and populated ontological instances from relational database management systems. Another approach is presented in [19], the knowledge base is enriched by crawling and extracting textual information from external websites. Similarly, Valkeapaeae et al. [20] introduce Saha system used to extract concepts from various document types and annotate them. Blythe et al. present their solution to integrating information from heterogeneous sources with one common domain ontology in [21], which focuses on analyzing HTML pages and extracting named entities in the texts. In [16], an ontology-based musical recommendation system, which uses RSS feeds of other websites as its main input data for the system knowledge base, is presented.

In musical information domain, the data types include binary and text-based formats. Hence, a musical knowledge-base system should manage these two types. From above analysis, we see that the approach of building a knowledge-base management system of musical information, which is employed the advantages of Semantic Web technology, is not mature enough. Therefore, we propose our SEMU system to deal with this problem.

3. Methodology

3.1. Conceptual framework

This section presents the conceptual framework of SEMU system, as shown in Fig. 1. The framework, which can be classified into four main processes, is described as follows:

1. *Data collection:* in this process, the raw data of musical information domain are collected. The system inputs, which are used in this research, include musical web pages and audio files. These materials are automatically downloaded web-wide by the web crawler. In addition, audio files can be contributed by users.
2. *Feature information extraction and classification:* This phase is responsible for processing the raw data provided by the process of data collection. For text-based data, the Information Extractor, which uses NLP toolkit (Apache OpenNLP[1]) and

[1] http://opennlp.apache.org/.