

# Knowledge based Semantic Annotation Generation of Music

Sunitha Abburu, PhD.  
Professor & Director, Dept. of M.C.A  
Adhiyamaan College of  
Engineering, Tamilnadu.

## ABSTRACT

The storage capacity and the cost of the storage devices gives rise to voluminous music collection management. Various MIR techniques exist, but getting the required songs from large collection of music files, is still a challenging problem. Getting the required songs from voluminous collection with good recall and precision depends on good annotation, indexing and retrieval techniques. Among this annotation plays a vital role in developing an efficient and effective retrieval system. Tag based annotation or the low level feature based annotations retrieval systems performance is very poor, though the processing is very simple. Semantic concept based annotation, indexing and retrieval techniques are trying to fill the gap between the machine understanding and the human preferences. This raises the need for semantic based annotation of film songs. Ontology plays a major role in semantic web and information retrieval. This raises the need for an ontology based annotation generation tool for film songs. The current research designs and implements a tool M-SAGT –Music Semantic Annotation Generation Tool. Which is flexible, user friendly and ontology based semantic annotations can be generated and stored in RDF/XML format. The generated semantic annotations can be used in semantic indexing and retrieval, which will enhance the performance of the retrieval system.

## General Terms

Semantic Information Retrieval.

## Keywords

Annotation, Film music, Ontology, Semantic knowledge.

## 1. INTRODUCTION

Enormous Enormous volume of digitized multimedia content has changed the way the multimedia content are stored, annotated, indexed and retrieved. The enormous growth of the large music database throws many challenges to researchers for effective and efficient management. When searching for a song of user choice, using a search engine such as Google still requires manual work in filtering out the undesired songs. The user is mostly buried with too many irrelevant music files. With the volume of music databases it has become a demanding challenging from users and music lovers to have an efficient annotation and retrieval system.

The performance of the music retrieval system would be good if the semantic concepts of the music are considered. With the advent of the semantic web it is a well known fact that performance would be high with the semantic concepts and the knowledge representation. The recall and the precision of the music retrieval system can be enhanced by considering the semantics knowledge of the songs. However all that knowledge is mostly stored in an unstructured fashion that needs human consumptability. Semantic web era aims to add

information to multimedia content to make it processable by intelligent systems. This is accomplished by the annotation of metadata with ontology based vocabularies which provides the establishment for high recall and precision music retrieval systems. In order to achieve this, knowledge management methodologies and metadata classification schemes for the domain have to be taken in to consideration. It is a challenging issue to align the text based annotation with the semantic concepts in a meaningful context. The current research presents the enriched value of ontology based semantic annotation and the significance of aligning it with the semantic Web. The significance of knowledge management for the musical domain lies in knowledge about music, called metadata. Such metadata is used to annotate the music files and these descriptions about music build the foundation for building knowledge management system and ultimately, for exploiting existing knowledge in order to assist in finding accurate song as per the users choice. To facilitate relevant songs retrieval of large-scale music databases, the detection of musical concepts, or auto-tagging, has been an active research topic. Many songs in large music databases are not labeled with semantic tags that could help users sort out the songs they want to listen to from those they do not.

The rest of the paper is organized as follows. Section 2 gives the literature survey. Section 3 discusses the design of the film music semantic annotation generator toll (M-SAGT). Section 4, gives the classification of semantic concepts of film songs and implementation of M-SAGT. Conclusion and future enhancements are been discussed at the end.

## 2. RELATED WORK

Efficient and Intelligent Music information retrieval is very important and hot topic of 21<sup>st</sup> century. Music information retrieval (MIR) is an emerging research area that receives growing attention from both the research community and music industry. It addresses the problem of querying and retrieving music from large music data according to users choice. Huron [1] discusses the preeminent function of music are social and psychological, the most useful concepts are based on four types if information like genre, emotion, style and similarity.

In traditional keyword assignment the terms are often drawn from a controlled vocabulary. Where as in tagging, no restriction are imposed on the choice of words on the makeup of a tag. A typical tag is a word or short phrase that describes the multimedia resource. There is usually no restriction on the number of tags that can be assigned to an item. Music tags are free text labels associated with artists, genre, emotion, mood, instruments, etc. [2] describe the state of the art in commercial and research social tagging systems for music along with how tags are collected and used in current systems. Researchers need to expand the pool of contextual knowledge about music for efficient annotation. One source for this contextual

knowledge is the set of social tags that humans apply to music. Social tags are the result of collaborative tagging. In a social tagging system, an individual applies short, free text annotations (tags) to items, typically to organize their personal content. These tags can be combined with those created by other individuals to form a collective body of social tags. With a large enough set of taggers generating many tags, a very rich view of the tagged items emerges. Social tags are typically used to facilitate searching for items, exploring for new items, finding similar items, and finding other listeners with similar interests.

Recently, automatic audio tag annotation has been a raising and active research topic [3][4]. Social tags [5] have started to play a key role in the development of “Web 2.0” technologies and have become a major source of information for recommendation. Consequently, music tag classification seems to be a more complete and practical way for musical information classification. Given a music clip, the tagging algorithm can automatically predict tags for the music clip based on the models trained from music clips with associated tags collected beforehand. [6] automatically annotates audio clips with tags and retrieve relevant clips from a music database by tags. The given audio clip is divided into several homogeneous segments by using an audio novelty curve, and then extracts audio features from each segment with respect to various musical information, such as dynamics, rhythm, timbre, pitch, and tonality. If the words that apply to a song can be predicted from audio, then those predictions can be used both to automatically annotate a song with tags, allowing users to get a sense of what qualities characterize a song at a glance. Automatic tag prediction can also drive retrieval by allowing users to search for the songs most strongly characterized by a particular word. Codeword Bernoulli Average model [7], which predicts the probability that a tag will apply to a song based on counts of vector-quantized feature data extracted from that song. In contrast to using probability models, Eck et al. [8] use Ada Boost to automatically generate audio tags for music recommendation. When attempting to annotate music, it is important to consider both acoustic content and social context. [9] explores techniques for collecting and combining multiple sources of such information for the purpose of building a query-by-text music retrieval system. The method considers two representations of the acoustic content (related to timbre and harmony) and two social sources (social tags and web documents). These sources of music information are useful for semantic music retrieval. The current challenge the researchers are facing is, How to represent the music genre taxonomy? Music genre is widely used to classify music and to aid in retrieval and discovery. However there is no general agreement as to the correct genre taxonomy. Traditionally, few music information retrieval (MIR) researchers focused on musical information classification with respect to genre, mood, instrumentation, quality, etc. Automatic music style classification is an important, but challenging problem in music information retrieval. It has a number of applications, such as indexing of and searching in musical databases. Traditional music style classification approaches usually assume that each piece of music has a unique style and they make use of the music contents to construct a classifier for classifying each piece into its unique style. [10] proposes a multi-label music style classification approach, called Hypergraph integrated Support Vector Machine (HiSVM), which can integrate both music contents and music tags for automatic music style classification.

Douglas [11] presents a system that can both annotate novel audio tracks with semantically meaningful words and retrieve relevant tracks from a database of unlabeled audio content given a text-based query. Douglas collected a data set of 1700 human-generated annotations that describe 500 Western popular music tracks. For each word in a vocabulary, this data is used to train a Gaussian mixture model (GMM) over an audio feature space. “query-by-text” is been used to retrieve appropriate songs for a large number of musically relevant words. Efficient music information retrieval of Carnatic music is been discussed in [12]. A methodology for Carnatic music singer identification is proposed and implemented. A new set of features like just tempered, varying pitch, varying interval of the octave, etc are identified to identify the singer. The coefficients are called as Carnatic interval cepstral coefficients (CICC) since they are based on Carnatic music’s octave interval. These coefficients are used to construct a GMM model which forms the basis for singer identification.

Mood annotation of music is challenging issue as it concerns not only audio content but also extra-musical information. It is a representative research topic about how to traverse the well known semantic gap. [13] propose a new music-mood-specific ontology. Novel ontology-based semantic reasoning methods are applied to effectively bridge content-based information with web-based resources. Performance analysis is a common issue in the musical field, especially for instrument teachers and musicologists. That is why [14] present an ontology (technical aspect) and a descriptive model (pragmatic aspect) to allow music teachers and students to create semantic annotation on musical performances. The proposition is based on real music lesson analysis and includes six main concepts to precise Tone, Harmony, Gesture, Dynamic, Appropriation and Structure aspects of a given performance.

[15] discuss the use of knowledge for analyzing and retrieving music contents semantically. Context-based Music ecommendation (COMUS) ontology grounds desired user emotion state from context and user preference information in the ontology. COMUS is a music dedicated ontology in OWL constructed by incorporating domain specific classes for music recommendation into the Music Ontology, which include situation, mood and musical features. More specifically, describes the ontologies of mood and situation in music using low-level features like pitch or duration and musical factors like tempo or rhythm.

[16] describe a novel approach to apply text-based information retrieval techniques to music collections. Tracks are represented with a joint vocabulary consisting of both conventional words, drawn from social tags, and audio muswords, representing characteristics of automatically-identified regions of interest within the signal. [17] investigate an approach to a music search engine that indexes music pieces based on related Web documents. This allows for searching for relevant music pieces by issuing descriptive textual queries. The effects of incorporating audio-based similarity into the text-based ranking process – either by directly modifying the retrieval process or by performing post-hoc audio based re ranking of the search results is been examined . The aim is to improve ranking quality by including relevant tracks that are left out by text-based retrieval approaches.

### 3. SEMANTIC ANNOTATION GENERATION FOR FILM SONGS

For music lovers, TV show editors and other group of music users finding suitable songs for a particular purpose is a big challenge. The rapid growth of technology resulted in voluminous storage and retrieval of music files. While voluminous music collection exists, techniques for semantic annotation and semantic retrieval are limited.

Many annotation and retrieval schemas use tag based annotation and searching but not much more useful to the user in finding the desired songs. The recent research trends promote the use of explicit semantic knowledge as a way. The semantic web and the associated standards like RDF, helps represent background knowledge of any multimedia. This paper, explores the use of semantic knowledge contained in ontologies to annotate and search with in a large collection of music files. This paper discusses the development of an annotation tool (M-SAGT) that helps to formulate ontology based semantic annotation of film songs.

All annotations generated by M-SAGTool are stored in RDF/XML form. Since RDF/XML format is the most widely used format in the semantic web and widely accepted by the W3C. Protégé is an ontology editor that is used to develop ontologies. Protégé (<http://protege.stanford.edu>) was developed by Mark Musen's group at Stanford University. Protégé, the ontology editor can be used to design classes in flexible style and to organize classes as hierarchy. The protégé editor supports the construction of ontologies and saves the ontology in RDF/XML format. Ontologies used in the annotation tool are specified in RDF schema using protégé.

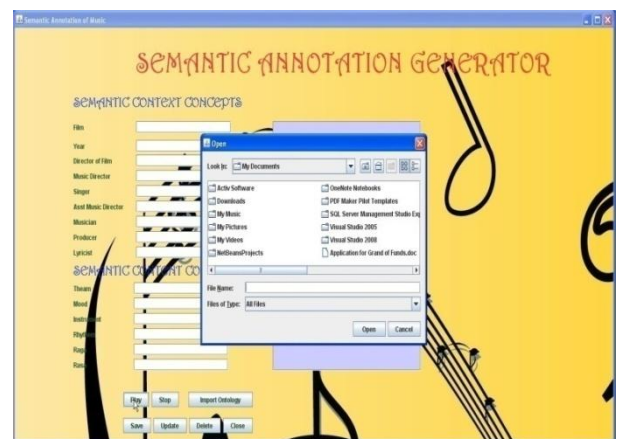
The M-SAGT helps the user to generate ontology based semantic annotation of music files. The tool supports browsing the songs from the personalized collection, play the song and generate the semantic annotation using high level features and storing annotation in an XML file format. The semantic annotation can be generated using the semantic features listed in the tool interface. The tool listed out various semantic concepts pertaining to cinema music. [18] have presented a novel framework of utilizing concept correlations to improve musical concept detection. To improve musical concept detection, concept correlations are used. Yi-Hsuan proposes to formulate concept detection as an ordinal regression problem to explicitly take advantage of the ordinal relationship between concepts and avoid the data imbalance problem of conventional multi-label classification methods. In M-SAGT user has free hand to use these features, where all the extracted semantics for a particular concept can be drawn as a drop down list. For new semantic concepts or new description user has the choice to import the ontology pertaining to that semantic concept.

### 4. EXTRACTION OF KEY SEMANTIC CONCEPTS

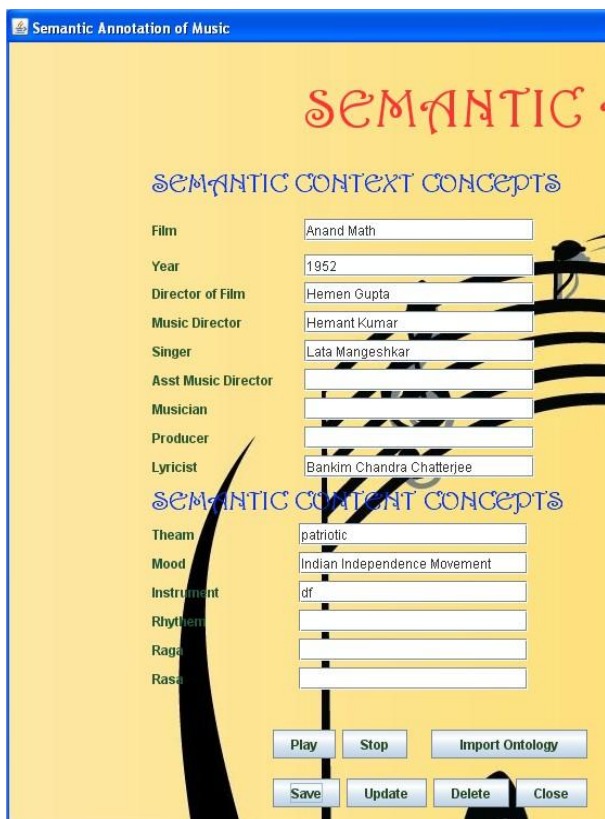


**Fig. 1 Semantic Annotation Generator Tool for film Music**

Few semantic concepts are common for all film songs like singer, music director, film, year, lyricist, theme, mood etc. It is been observed that these above common semantic features gives rise to a common frame work for the semantic concept classification. Semantic concepts for film songs can be categorized as semantic content dependent and context dependent features. The context dependent features are those that do not depend on content but derived from the context of the song. These are like singer, film, music director and so on. Content dependent semantic concepts are those concepts which are derived from the song. They are like theme, mood, rhythm, raga, rasa, instruments used etc see figure1. The human annotator has to identify the semantic concepts of the song. As Indian films has got its own style and diverge topics for songs leads me to develop an annotation tool that uses the existing music ontology. The ontology based music annotation improves the performance of the annotation which in turn improves the efficiency of the semantic retrieval system. This also fills the gap between the low level and high



**Fig. 2 Users choice to select a song**



**Fig.3** Figure showing the annotation of a film song

level semantic concepts. The tool supports browsing the songs from the personalized collection see figure 2. A song annotation should contain at least one semantic concept from any category. But effective annotation and retrieval techniques demands for recording of all semantic features. To facilitate this, the tool provides the list of concept classification, concept features and concept instances see figure 3. We often need to incorporate definitions already available, example instruments used in songs. Every instrument need not have definitions for every song. Usually the film songs use instruments and all instruments in all songs need not be defined. This will create redundant information about the same type of instrument. The performance of the system will degrade with this kind for definitions. To annotate



**Fig.4** Import Ontology for a semantic concept for annotation

the instrument concept, existing instrument ontology could be effective. Since ontology construction is continues ongoing process and the inherent feature of the ontology is its flexibility and extendibility. Semantic concepts can be added at any moment. This broadens the scope of effective annotation, indexing and retrieval system. In M-SAGT the user is also privileged to import any exiting music ontology see figure 4. And the semantic annotation is generated with the help of the ontology, which represents the semantic knowledge of the film songs. The generated annotations are stored in RDF/XML format. Retrieval with a search facility using RDF can read the RDF notation and search for required songs.

The annotation can be generated in two ways, annotator can use free text, the tool checks for the correctness and similarity of the terms entered to annotate to the concept terms with respective concept ontology. If the similarity is above the threshold the description given by the annotator is accepted or the user can use the dropdown list to select the semantic concepts. The other method is to click on the import ontology to brows the existing ontologies for that category of semantic concept. The tool also has the facilities like browsing a new song, storing the annotation, updating of annotation or deletion of annotation. The semantic annotations generated using this tool, could be used for semantic indexing and semantic retrieval with which the effectiveness and the efficiency of the retrieval system will be improved.

## 5. CONCLUSION

Voluminous music collection through various storage and retrieval challenges to the researchers. One among them is the semantic annotation generation using ontology. Semantic annotation generation for film songs is an research area where very less work has been done. The current research focuses on semantic annotation generation using ontology. Ontology based semantic annotation plays a major role in effective and efficient music information retrieval. The semantic concepts are identified and categorized in to two major classes, semantic context dependent and semantic content dependent. This paper discusses the design and implementation of music semantic annotation generation tool M-SAGT. The generated annotations are RDF/XML file format, this format is accepted worldwide and accepted by W3C. The generated annotation can be used in the retrieval systems which are using the RDF/XML form searching. Further research could be conducted on semantic index generation and classification of music songs, with which the performance of the retrieval system will be enhanced.

## 6. ACKNOWLEDGMENT

This work has been partly done in the labs of Adhityamaan College of Engineering where the author is working as a Professor& Director in the department of Master of Computer Applications. The author would like to express her sincere thanks to Adhityamaan College of Engineering for their support rendered during implementation of this module.

## 7. REFERENCES

- [1] D.Huron, “ Perceptual and cognitive applications in music information retrieval “ in Proc. Int. Symp. Music Information Retrieval, 2000.
- [2] P. Lamere, “Social Tagging and Music Information Retrieval,” Journal of New Music Research, vol. 37, no. 2, pp. 101–114, 2008.

- [3] D. Eck, P. Lamere, T. Bertin-Mahieux, S. Green, "Automatic Generation of Social Tags for Music Recommendation", NIPS, 2007.
- [4] D. Turnbull, L. Barrington, D. Torres, and G. Lanckriet, "Semantic Annotation and Retrieval of Music and Sound Effects," IEEE Trans. on Audio, Speech and Language Processing, vol. 16, no. 2, pp. 467–476, 2008.
- [5] Zhi-Sheng Chen Jyh-Shing Roger Jang, "On the Use of Anti-Word Models for Audio Music Annotation and Retrieval", IEEE Transactions On Audio, Speech, And Language Processing, VOL. 17, NO. 8, November 2009, pp 1547 -1556.
- [6] Hung-Yi Lo; Ju-Chiang Wang; Hsin-Min Wang," Homogeneous segmentation and classifier ensemble for audio tag annotation and retrieval ",IEEE International Conference on Multimedia and Expo (ICME), 2010 , Page(s): 304 – 309.
- [7] M. Hoffman, D. Blei, and P. Cook, "Easy as CBA: A simple probabilistic model for tagging music," ISMIR, 2009.
- [8] D. Eck, P. Lamere, T. Bertin-Mahieux, S. Green, "Automatic Generation of Social Tags for Music Recommendation", NIPS, 2007.
- [9] Douglas et.al., "Semantic Annotation and Retrieval of Music and Sound Effects", IEEE Transactions On Audio, Speech, And Language Processing, VOL. 16, NO. 2, February 2008, pp467-476.
- [10] Rajeswari Sridhar, T. V. Geetha , "Music Information Retrieval Of Carnatic Songs Based On Carnatic Music Singer Identification" , 2008 International Conference on Computer and Electrical Engineering, pp 407-411.
- [11] Jun Wang,et.al., "Enriching Music Mood Annotation By Semantic Association Reasoning", 2010 IEEE 2010, pp1445-1450.
- [12] Véronique, et.al., "An Ontology for Musical Performances Analysis" -Application to a collaborative platform dedicated to instrumental practice, 2010 Fifth International Conference on Internet and Web Applications and Services,2010, pp 538-543.
- [13] Seheon Song, et.al., "Music Ontology for Mood and Situation Reasoning to Support Music Retrieval and Recommendation", IEEE 2009 Third International Conference on Digital Society,2009, pp 304-309.
- [14] Mark Levy, Mark Sandler, "Music information retrieval Using social tags and audio", IEEE Transactions On Multimedia, pp 1-14.
- [15] P. Knees, et. Al., "Augmenting Text-Based Music Retrieval With Audio Similarity", 10th International Society for Music Information Retrieval Conference (ISMIR 2009) pp 579-584.
- [16] Yi-Hsuan Yang, Yu-Ching Lin, Ann Lee, Homer Chen," Improving Musical Concept Detection By Ordinal Regression And Context Fusion", 10th International Society for Music Information Retrieval Conference (ISMIR 2009) , pp 147-152.
- [17] P. Knees, et. Al., "Augmenting Text-Based Music Retrieval With Audio Similarity", 10th International Society for Music Information Retrieval Conference (ISMIR 2009) pp 579-584.
- [18] Yi-Hsuan Yang, Yu-Ching Lin, Ann Lee, Homer Chen," Improving Musical Concept Detection By Ordinal Regression And Context Fusion", 10th International Society for Music Information Retrieval Conference (ISMIR 2009) , pp 147-152.