

Moody Tunes: The Rockanango Project

Nik Corthaut, Sten Govaerts, Erik Duval

K.U. Leuven
Celestijnenlaan 200A
B-3001 Heverlee

{Nik.Corthaut, Sten.Govaerts, Erik.Duval}@cs.kuleuven.be

Abstract

Wouldn't it be nice if we had a tool that could offer people the right music for a specific time and place? For HORECA (hotel, restaurants and cafés) businesses, providing appropriate music is often not just nice, but essential. Typically this boils down to music that matches a certain situation on desired atmospheres, this will be defined as a musical context (MC). The developed tool, a music player, meeting the specific needs of HORECA, allows creation and management of those contexts. The user creates a musical context by selecting a number of appropriate atmospheres and can fine-tune the context with additional musical properties. The atmospheres are defined by a group of music experts, composed of DJ's, music teachers, musicians, etc., who also manually annotate the properties of all musical content. To assist the music experts, a specially developed tool allows them to categorise and annotate the songs and evaluate their results. We provide insight on how we constructed and implemented our metadata schema and look at some existing schemas. The evaluation shows the economic value of such a system in the specific context of a HORECA business.

Keywords: Multimedia Systems, Music Information Retrieval, Context, Metadata.

1. Introduction

Wouldn't it be nice if we had a tool that could offer people the right music for a specific time and place? For HORECA businesses, providing appropriate music is often not just nice, but essential. Automating this process requires means to describe music and to select music that corresponds to a given description that mirrors the context. Moreover, individual pieces of music must be combined into a playlist and automatically rendered in a nice mix. This way the user can focus on his core task: bartending, rather than acting as a DJ. This is the focus of the Rockanango project (rockanango.cs.kuleuven.be), an 18 month cooperation between the K.U.Leuven (www.kuleuven.be), a Belgian music content provider for HORECA, Aristo

Music BVBA (www.aristomusic.com) and with the financial support of IWT-Vlaanderen (www.iwt.be).

For the music recommendation part, legacy systems are roughly based on two strategies. In collaborative filtering, the musical taste of a specific user is matched against that of all other users. The assumption is that the usage patterns of similar users lead to useful recommendations. This approach requires some degree of feedback. This can be very simple, in a 'do you like this?' style or more indirect, by tracking user behaviour. Audioscrobbler (www.audioscrobbler.net) as used by Last.FM (www.last.fm) and Amarok (amarok.kde.org) rely on this technique.

A more common strategy for music recommendation is content based filtering. In this approach, the actual music and metadata are processed to determine patterns and categorizations. MusicLens (www.musiclens.de), MusicMiner (musicminer.sourceforge.net) and the Music Genome Project (www.pandora.com) rely on this approach. The metadata is provided by feature extraction and manual annotation of the music, either done by professional experts or by a community of users.

Consumer applications like Pandora or Last.FM rely on continuously evolving personal music profiles. Bootstrapping proceeds in a similar way in both systems. First, the user is prompted for an artist or song he likes and then an audio stream with similar songs is created. The user cannot intervene directly in the music selection. However, the user can indicate whether he likes the music or not, which is used to fine-tune the profile and thus taken into account for future recommendations. The emphasis is on finding more music that the user likes. In this way, he can expand and explore his musical boundaries.

Unlike the abovementioned approaches, we focus on commercial use of music in public places, like pubs, restaurants, hotels, etc. In our context, the customer population changes throughout the day and the week. As an example, customers may want to dance in the evening. At other times, a thematic music selection is more appropriate, for instance when there is a big sports manifestation, a Latin night, a disco evening, etc. The bartender needs to adapt the music to certain types of clientele and situations, immediately, without the converging process associated with profile fine-tuning.

In our approach, songs are annotated with roughly twenty metadata fields and MCs are created to define the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2006 University of Victoria

relevant music, by selecting values for the different metadata fields. In this way, full control is in the hands of the context creator. Music experts have created a set of default contexts for different occasions. Bartenders can fine-tune these defaults or create contexts of their own. After the context creation, a playlist can be generated that covers a specific timespan. The user preserves control over the playlist. He can add or remove tracks, change the order, ... After all, it might not be a good idea to play the soccer hymn of the adversary in the club's canteen.

This paper is structured as follows: In the second section we define and explain the concept of a MC. In the next two sections we determine what information is useful to model MCs and what we used. We compare our approach with the work of others. The fifth section deals with the implementation of the system and the final 2 we try to validate the system and give a road map for the future.

2. What is a Musical Context?

A personal music collection typically contains a few thousands songs. Manually picking songs from such a collection in order to create a MC is a tedious task that demands expertise and skill. In most tools, the user must know the songs by e.g. artist to find the song. Moreover, this manual approach limits the ability to discover new music, since the selector has to know his music collection pretty well.

With a description of the music we want for a certain context, we make it possible for the computer to select the appropriate music and propose a playlist. Different musical parameters, like genre, rhythm, mood, occasion, year, dance style, etc., describe the music selection for a context. For the dance evening we want for example:

- *Genre*: Dance
- *Mood*: Cheerful, Exuberant, Delirious, Stimulating.
- *Popularity*: medium to very high
- *Party level*: medium to very high.
- *Loudness*: medium to high

By selecting the music that satisfies the description multiple playlists with a specified time span can be generated without having to know the entire music collection. In order to be able to select songs for a MC, we need a detailed description of each song in the collection. Through experience and knowledge, a musical expert knows in which situation he can play a song. The expert has the skills to use appropriate songs for a situation with a certain atmosphere. They actually assign possible atmospheres to the different songs. An example: the description of the atmosphere for a song used for a slow dance between amorous couples could be: romantic, emotional, intimate, etc.

These atmospheres can be recalled if the end user thinks they are appropriate for his specific situation. Basically, selected atmospheres are reused in a different (musical) context and through that, they evoke the atmospheres of the situation they were derived from. These MCs can then be further refined for personal taste with other musical properties, like genre, rhythm, instrument, etc. In conclu-

sion, a musical context is a musical description for situations based on atmospheres and musical properties.

3. Modeling of Metadata

As discussed before, we use metadata to describe a MC. These include the author and the title of the song, the group members of the band, the record label and the genre of the song. However, such basic information is not sufficient when we want to generalise and determine whether the music is good for partying, for a candlelight dinner, for children or for a special occasion.

3.1 What metadata?

3.1.1 Objective and subjective metadata

Many of the metadata elements define characteristics which are objective, such as the song author and title, artists, the release year, the version, ...

Of course, there are a lot more musical characteristics, which tend to be based on personal perspective or differing points of view, hence more subjective [1]. A few examples are popularity, mood, genre, hardness (metal vs. lounge), etc.

3.1.2 An extraction based categorisation

Based on the way the metadata is extracted, we can categorise musical metadata in different groups [2][3]:

- *Editorial metadata*: authoritative experts provide the metadata manually.
- *Acoustic metadata*: the information (usually purely objective) is obtained by analysis of the music file, without any reference to textual or prescribed information.
- *Cultural metadata*: the metadata results from an analysis of emerging patterns, categories or associations from a source of documents produced by the environment or culture.

3.2 Other metadata schema's

Songs can be categorized in many ways, however comparison of different metadata schemas of existing systems learns that they all incorporate the same basic editorial metadata for a song. We will discard this metadata for now. We are interested in the metadata that describes the musical features of a song and subjective metadata like mood and purpose.

In table 1, we collected these metadata elements for 8 systems. We weren't always able to obtain detailed information about the systems mentioned here, and basically relied on [4] and [5]. All of the systems are used for music recommendation, although most people use the All Music Guide mainly as a music encyclopedia. The PATS system [6] is focused on Jazz Music, so it contains a lot of specific metadata for this kind of music, e.g. the ensemble strength, the soloists and the geographical roots. We don't really need this for HORECA.

Table 1: comparison of existing metadata schemas

Magnatune src: magnatune.com editorial	MusicLens src: musiclens.de editorial/cultural	All Music Guide src: manual search editorial/cultural	MoodLogic src: from application editorial/cultural
mp3 genre ① magnatune genre	genre ① volume(silent-loud) tempo(slow-fast) ⑤ voice(voice-instrument) ⑥ size(solo-orchestra) purpose(listening-dance) ③ sex(female-male) age(0-70) mood(smile-angry) ② colour of song	genre ① styles moods ② similar artists ⑦ influenced by followers see also instruments ⑧ country	genre ① type recording voice ⑥ sound quality similar artists ⑦ energy energy level beat tempo ⑤ danceable mood(negative-positive) ② memorability of melody lyrics language lyrics topic lyrics explicit (x-rated) instruments ⑧ general type instruments mood description ②
iTunes src: weblog [5] editorial	AudioInsight src: article [4] editorial/cultural	Sony's StreamMan src: streamman.net editorial/cultural	
genre ① rating ④	genres ① subgenre microgenres	genre ① mood ② atmosphere ③ decade rating ④	

People seem to use genre ① as the main categorizer. All systems have it; some use only that apart from the basic metadata, e.g. Magnatune. We need it too for MCs.

Another interesting data element, which reoccurs in the other systems, is mood ②. AllMusic, Moodlogic and StreamMan use a value space for the mood description. MusicLens uses an interval between positive feelings and negative, as does MoodLogic on top of a mood description. As abovementioned, this is what we will use too to describe musical contexts. StreamMan and MusicLens take this a bit further by enabling to describe the purpose ③ (ranging from listening to dance). We embedded this in our system in the dancability data element. This is an essential factor in the HORECA.

MusicLens also takes the listener into account by selecting music based on the listener's sex and age. When there are a lot of people present, describing the listeners becomes very hard. We put the stress on describing the music, not the people.

There are more similar elements, like tempo ⑤, instruments ⑧, similar artists ⑦, ... MoodLogic features other elements that can be very useful to describe the desired music, e.g. language, the energy of a song, the topic of the lyrics, ... All Music Guide recommends music based on similar artists and an influence network between artists, we however believe that we should work on song level for recommendations.

3.3 Incremental approach

3.3.1 Why?

A lot of the objective metadata can be retrieved from existing databases, like FreeDB (www.freedb.org). Other objective metadata fields, like for example the duration, the beats per minute (BPM), the intro end and the beginning of the outro, can be calculated by the computer

(www.un4seen.com/bass.html) [7]. Subjective metadata is typically hard to determine by computer. Collaborative filtering techniques are an alternative, but require feedback and a large user population. A pub owner is typically too pre-occupied to give feedback to the system.

Because we don't have the critical mass to converge to valid metadata, we decided to annotate the music manually a priori. Later on, when we have a solid metadatabase, we can automate annotation with feature extraction techniques for data member estimation.

In the beginning, we had a very long list of possible metadata elements, but it was not clear what metadata elements were needed or useful to describe a MC. Some musical properties that are subject to personal taste can be hard to reach consensus on. The multitude of musical metadata elements and songs often implies that it is hard to let the categorization converge to a uniform tagging over the whole music collection. Experience learned that this is mainly due to the fact that many people work together on the same task and during a long time period, in which they find new insights and change their minds.

3.4 Construction of the metadata schema

We opted a flat structure for the metadata elements with independent value spaces, allowing the structure to grow, adapt and change. This allows flexibility in the structure and favours the modeling capabilities for a MC.

By working closely together with the music annotators, we built a software metadata input tool that allowed them to experiment with many musical parameters and play with the parameters' modeling powers to create a MC. The card sorting idea from the usability-testing world lies at the basis of the metadata input tool (<http://www.iawiki.net/CardSorting>).

On the left side of Figure 1 they have a working set of songs available to annotate and listen to. They can then select a data element (2nd column) to work on, e.g. genre.

Then they will be presented with the value space of this data element and add songs to these values (cards), e.g. pop, rock, metal, ... The experts can also add data elements and edit their value spaces. A visualisation tool [8] is used for evaluation purposes, which clusters the songs based on musical parameter constraints and let them compare their work with that of the other team members.

coldplay-clocks ac/dc-tnt pez-wild gypsy fela kuti-zombie REM-bad day sting-roxanne blur-song 2 korn-word up slayer-epidemic cher-bang bang flaviium-nightlife	genre ▶	pop	korn-word up slayer-epidemic
	subgenre	rock	
	mood	metal ▶	
	voice	jazz	
	party level	club	
		trance	

Figure 1: paper mock-up of metadata input tool.

3.5 Who works on the metadata?

The choice of the people who work on the metadata will determine the categorisation and hence also the usefulness of the fields to describe a MC with. The success or failure of a playlist generated from a MC depends highly on the quality of the metadata.

It is very hard to gather such a team. When we use for example a group of musicologists, we will probably get historically correct data, e.g. a classical reactionary work can historically have an aggressive and dark mood, but we can perceive it today as vibrant and enthusiastic instead.

On the other hand, if we just let anyone annotate music, we get a lot of data from people with different backgrounds and musical tastes. The major disadvantages are possible divergence in the selection of appropriate values from the value space and the need for a substantially large user base to support this. This would require a lot of time to bootstrap and due to the relatively short duration of this project (18 months) we chose another option. The idea to create a community collaborative system has not been abandoned and might be embedded in a follow-up project.

We decided to work with a diverse group of music experts, including DJ's (from different areas: clubs, discos, parties, weddings, fashion shows, etc.), music teachers, dance teachers, musicologists, musicians, producers, etc. These music experts have expertise in their own musical sub-domain. Through their diversity, experience in the field and their work as a team they can unveil the correlation between musical properties and atmospheres without being all too rigorous or fundamentalistic. For most people reggae is associated with the typical backbeat guitar riff though not all reggae songs have this peculiar pattern. The end goal is to fulfill the expectations of the users of our system, which are not music experts. The old saying 'Equality is fitness for purpose' is still valid

(<http://www.qualityresearchinternational.com/glossary/fitnessforpurpose.htm>).

3.6 Evolution of the metadata schema

In the beginning we presented the music experts a selection from the metadata fields as seen in Table 1 and asked them for their opinion and additions. This is how we came up with the left column of Table 2 (this is also a reduced schema like in Table 1).

As shown in the first column we took a lot of data elements from the existing systems, e.g. genre, mood, dance level, ... The experts added rhythm style and party level because they found these important to select dance music and are not present in the other systems. The geographical location was important for the music distribution, because some regions have very local music needs, carnival music for example is very region bound.

Table 2: evolution of the metadata schema.

Original Metadata Schema	Present Metadata Schema
<u>version</u>	<u>version</u>
<u>language</u>	<u>language</u>
<u>tempo (bpm)</u>	<u>bpm</u>
<u>voice</u>	<u>voice</u>
<u>genre</u>	<u>genre</u>
<u>most significant instrument</u>	<u>instrument</u>
<u>mood</u>	<u>mood</u>
<u>rhythm style</u>	<u>rhythm style</u>
<u>dance level</u>	<u>danceability</u>
<u>party level</u>	<u>party level</u>
<u>global popularity</u>	<u>global popularity</u>
<u>timeless</u>	<u>actuality</u>
<u>geographical location</u>	<u>geographical region</u>
dynamic	occasion
original/cover	type
actuality based on music charts	loudness
	dance style
	continent
	subgenre

We compared the original schema with the present one in Table 2 by underlining the identical data elements and putting them next to each other. It is clear that the biggest part of the original metadata schema is still there. We added a few extra data elements, like the occasion to a play a song (Christmas, Halloween, ...), the loudness of a song (to be able to select some quiet music for dinner) and the dance style for more control on party music. On the other hand we merged timelessness and music chart actuality together. The new data element describes whether it's a hit (in the music charts) or whether it's new, recent or from the archive. This is especially handy for clients to check the latest music on their system.

Generally speaking, a few data elements were added to the set of the existing systems, which are typical for usage in the HORECA.

3.7 Problematic cases

3.7.1 Genres

There are many music genres available (http://en.wikipedia.org/wiki/Category:Musical_genres) making it hard to decide which is the most appropriate one. Most of the time this selection reflects the personal opinion of the music expert. A song by U2 for example could be categorised in Rock or in Pop. Neither case is ideal. Therefore the experts decided to create a genre and a subgenre categorisation. The former contains main genres, like Pop, Rock, Jazz, Rhythm & Blues, Classical, etc. The subgenre can contain songs from different genres and thus take care of the border cases between main genres. For the U2 case, all their songs are categorised in genre Rock and the harder songs in the subgenres Rock Café and the ones closer to Pop in Pop Café.

3.7.2 Popularity of a song

The relevance of the popularity depends highly on the user group. Each genre has its 'anthems', how many people know the main hit from the gabber scene, let alone appreciate it? To cope partially with this problem we came up with the idea of global popularity made for the mainstream user.

4. Modeling of a Musical Context

4.1 The idea

Say all the songs in the collection have values for all the metadata elements in the metadata schema. Now we want to create a MC description.

We can use the data members and their values to describe the MC. If you select values of musical parameters, you will get music that satisfies these parameter values. E.g. if we select genre: jazz and soul and mood: relax and intimate, we can get a playlist that has jazz or soul music that is relax or intimate.

We can take this one step further, say we want romantic music for a candlelight dinner and we select genre: pop and chansons, language: English and French and mood: romantic, of course. Now we get a playlist with English and French romantic songs. But we like the English songs more then the French and we want to be able to alter the ratio of French and English music. So you could create two MC descriptions, one for the English music and one for the French music. We call these subcontexts. Each subcontext has a certain weight ensuring for instance 20% French and 80% English music. This makes it also easier to create a complex MC (divide-and-conquer).

4.2 More abstract

Consider a set of musical parameters. Each parameter has a domain. Note that the domains are not necessarily compatible.

Each song has values for the musical parameters within the proper domain. A subcontext is defined as a subset of

parameters with selected values within their domain and a weight. This weight allows modeling things such as 20% French and 80% English songs. A musical context is defined as a collection of subcontexts. A song belongs to a subcontext if and only if the song contains at least one value that also occurs in the selected values of the subcontext, for each parameter in the subcontext. A song is part of a context if it belongs to the union of songs belonging to the subcontexts. When generating playlists from a context, the weight has to be taken into account.

Say for example we create a subcontext by selecting: genre: Pop & Rock, year: 2004 till 2006, mood: happy, summer & exuberant. To get all the songs contained by this subcontext description, we select the songs that have as genre Pop OR Rock AND the songs with mood happy OR summer OR exuberant AND the songs from the year 2004 OR 2005 OR 2006.

5. Implementation of the HORECA System

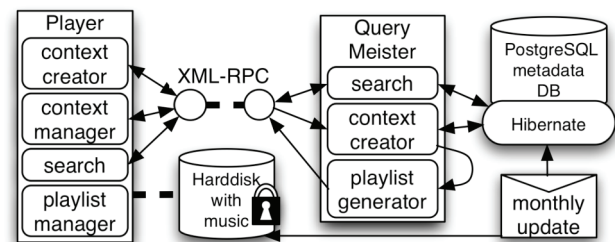


Figure 2: architecture of the HORECA player.

For HORECA businesses we developed a stand-alone system (Figure 2) consisting of a music player, a database system, an encrypted music archive and a set of services for querying and context retrieval. To distribute the right music to the right place, the customers are clustered in groups with similar music styles by doing an inquiry on their preferred musical taste. Monthly music, metadata, playlist and context updates are sent on a CD by mail. Downloading updates is a future project.

In the player the user can search in a Google-like fashion for artists, albums, titles, etc., get suggestions for similar music and can add the found results to playlists and manage them. To create a context the user can select the wanted values from the different data fields and generate a playlist or he can use a context provided by the experts.

On top of a PostgreSQL database (www.postgres.org) lies a Hibernate object-relational mapping (www.hibernate.org). Hibernate enables us to transparently communicate with the database through objects and offers different object-oriented query languages. One of these, the Criteria language, is used to dynamically build the query for the context creation and search service.

An XML-RPC layer takes care of the communication between the query part in Java and the player part in .Net. By placing the database and services on a server and letting users connect through XML-RPC over a network we can easily create a distributed system.

6. Evaluation

6.1 Why contexts are a great idea in HORECA businesses

Experience learned that the most important factors that the HORECA wants to offer their clientele are continuity of adequate music and the possibility to comply with customer requests. From the customer's perspective this is enough and this was more or less feasible with the old system. The pub or restaurant had a music player, the PCDJ Red (www.pcdj.com). Aristo Music provided the music and manually created playlists. User profiles were created dependent on the preferred style: rock, general pop, lounge, lounge/dance, restaurant, dance, schlager, youth house, background music, a general profile and custom profiles. Each of those profiles had variants depending the possible location of the business: Flemish or Wallon part of Belgium, the Netherlands or Luxemburg. The playlists and music are matched with the profiles in Aristo Music's ERP system.

Switching contexts is as simple as changing a playlist, one push of a button. Typically visiting a HORECA business takes between a few minutes up to a few hours. Bartenders stay through the day in the same working place, leaving them bored with the same playlists over and over. The use of contexts and a newly developed player offers an improvement on those matters, hence providing a direct perceived incremental benefit [9]. Music selection is better and more varied over time. The use of contexts is as easy as using playlists. In fact, a context can be seen as a variable playlist and is presented this way in the GUI. The HORECA workers can edit and create contexts on the spot, taking into account the given circumstances.

The new approach seems to hold. Since the introduction at a HORECA fair in November 2005 (www.horecaexpo.be), already 200 installations have been set up. Owners of the old system who are confronted with the new system indicate the willingness to migrate. Users of the new system do not want the old system back. The main reasons given for this behaviour are the diversity in music, the simplicity and ease of use. The main reason why the new product is not purchased is the lack of DJ functionality. This is currently under development.

7. The Future is Bright

With a version of the tool running in a few hundred HORECA businesses and a metadatabase of about 25000 songs, the foundations have been built for truly interesting points of improvement.

Playlist generation and rendering can be taken to a higher level. Gradual shifting from one MC to another

could be achieved with transition tracks. Also, the default cross-fade between tracks should be replaced by automatic mixing techniques.

Since June 2006, we have started the next phase of this project, currently we are focusing on the following topics. We are evaluating actual usage and the quality of the metadata. This includes determining how to deal with demographics and psychological matters regarding musical taste [10]. What parameters are really important for describing mood?

The painstaking task of entering metadata can be relaxed with music information retrieval techniques (<http://www.iaa.upf.es/mtg/clam/>). These can be embedded to automate the process of annotating music. This automation can combine suggestions for values, to assist the music experts, and calculating objective meta-data. Another way to get and evaluate metadata would be through setting up a community to gather metadata and to compare the results of collaborative filtering with the existing database. Both topics will be researched in the new phase.

References

- [1] E. Duval, W. Hodgins, S. Sutton, S.L. Weibel, "Metadata Principles and Practicalities" in *D-Lib Magazine*, April 2002, vol. 8, no. 4.
- [2] F. Pachet, A. La Burthe, J-J. Aucouturier, A. Beurivé, "Editorial Metadata in Electronic Music Distribution Systems: Between Universalism and Isolationism", published in *Journal of New Music Research* 2005, vol. 34, no. 2.
- [3] F. Pachet, "Knowledge Management and Musical Metadata", in *Encyclopaedia of Knowledge Management*, Schwartz, D. Ed. Idea Group, 2005.
- [4] S. Hansen, "Verzoeknummer – Hoe werken muziekadviesystemen?" in *C'T, Magazine voor computer techniek*, April 2006, Dutch, pp. 104-107.
- [5] The media metadata workspace, <http://www.eu.socialtext.net/mediametadata/index.cgi>
- [6] S. Pauws, B. Eggen, *PATS: Realization and User Evaluation of an Automatic Playlist Generator*, ISMIR 2002 3rd Int. Conf. on Music Inf. Retr.
- [7] X. Amariain, J. Massaguer, D. Garcia, I. Mosquera, *The CLAM Annotator: A Cross-platform Audio Descriptors Editing Tool*, ISMIR 2005 6th Int. Conf. on Music Inf. Retr.
- [8] J. Klerkx, M. Meire, S. Ternier, K. Verbert, E. Duval, *Information Visualisation: Towards an Extensible Framework for Accessing Learning Object Repositories*, World Conference on Educational Multimedia, Hypermedia & Telecommunications, ED-MEDIA 2005.
- [9] E.D. Scheier, *About this Business of Metadata*, ISMIR 2002 3rd Int. Conf. on Music Inf. Retr.
- [10] A. Uitdenbogerd, R. van Schyndel, *A review of Factors Affecting Music Recommender Success*, ISMIR 2002 3rd Int. Conf. on Music Inf. Retr.