Jackson: DJ Software Powered by Musical Metadata

Francis Van Aeken
VAN AEKEN SOFTWARE bvba
francis@vanaeken.com

Abstract

We have built an application that facilitates the creative mixing of recorded songs. The power of our application comes from musical metadata embedded in the recorded songs. We use a proprietary format for the metadata but plan to make it conform to the MPEG-7 standard. We believe that musical metadata will play an increasingly important role in the future. It will power not only creativity tools, but also virtual DJs, personal radio and intelligent music compilations.

1. Introduction

In the hands of a disc jockey, a recorded song is not an end product but a starting point. By building up a meaningful sequence of songs and skillfully interweaving them, a new work of art is created.

Many professional DJs use vinyl because turntables allow for a meticulous control of tempo and timing. Only recently CD decks are gaining popularity because they now mimic the tactile control of vinyl decks.

DJ software has existed for some time, but those packages typically emulate the traditional DJ setup. To properly use those programs one needs a physical controller and the same skills as a traditional DJ.

Computers offer new possibilities. Starting from scratch, our company has taken a approach different from emulation. Our product is the first dedicated DJ software that implements sequencer-style mixing and that exploits the power of musical metadata.

The application automates the less creative but sometimes difficult aspects of mixing. For example, beat-matching, the process of rhythmically aligning two different songs is automatic and perfect. This opens up the art of DJing to a larger public without affecting its creative side.

The underlying technology also allows for the creation of virtual DJs. Products like portable mp3 players can be made more attractive by incorporating these software agents that mix songs according to the preferences and the mood of the listener.

2. The DJ software 'Jackson'

Over the last three years Van Aeken Software has been building the DJ tool 'Jackson' that allows for the easy manipulation and mixing of songs. The application, developed in C++, runs under Microsoft's Windows XP. A demo can be downloaded from [1].

2.1. Playing and manipulating songs

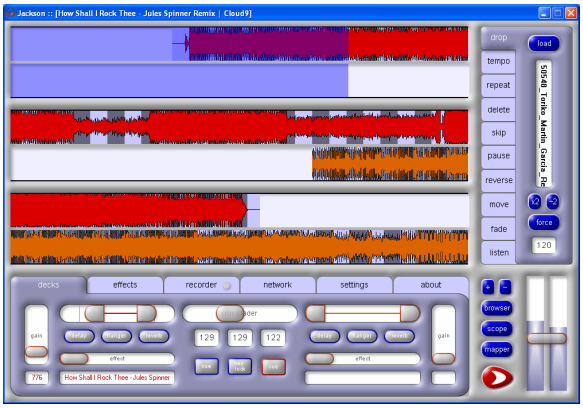

Figure 1. Jackson: main window

Figure 1 shows the main window of the application. Jackson supports two decks. The songs played on deck 1 and deck 2 are visualized as waveforms in different colors. The timeline is divided in three consecutive parts, one on top of another, much as in a music partition. The part of the songs already played is covered by a transparent colored rectangle. This rectangle grows with time until the complete first part is covered. At that point the two lower parts shift up and the 3rd part is replaced with new data. The rectangle disappears and starts growing from the left again.

The users can zoom in and out to see either the details of the waveform or a global overview of the sequenced songs.

Songs are automatically beat-matched and are therefore always rhythmically aligned.

On the right side of the waveforms window there are a number of tabs that cover different tools to manipulate songs and to control the mix. Jackson allows DJs to alter the structure of songs while playing them: parts can be repeated, skipped, paused and reversed.

Below the waveforms window one finds tabs covering the real-time mixer, the effects, the recorder, the network functionality and the user settings.

By using the filters in the mixer window the DJ can fade in and out, bass first or treble first. This way, parts of the spectrum of one song can be combined with parts of the spectrum of another.

The effects window features delay, flanger and reverb effects to spice up songs. All effects are automatically synchronized to the beat.

Through the recorder window the DJ can record the set he or she is playing to hard disk. This music file can then be burned on CD or published on the Web.

Through the network window, the user can configure the networking functionality of the application. Different computers running Jackson can be synchronized over TCP/IP. Several DJs can jam together while all played songs are automatically aligned to the beat.

The application also interfaces to standard midi controllers and custom controllers based on Measurement Computing's PMD-1208LS and Silicon Labs' C8051F320 controllers. Most DJs favor this type of controllers over the mouse. Not only are they better adapted to the type of control needed but they also allow the DJ to change multiple parameters at the same time.

Jackson supports the use of two different sound cards simultaneously. The output of one soundcard is directed towards the audience. The output of the other card is connected to headphones. Using such a configuration, one can cue like a traditional DJ or listen to parts of the mix in the future.

2.2. Selecting songs

A DJ is above all a selector. When mixing, nothing is more important than the selection of the songs. The browser, shown in figure 2, assists the user in this essential task.

The browser supports different virtual crates, corresponding to directories in the file system. Songs from a crate are displayed and ordered according to different criteria like title or tempo.

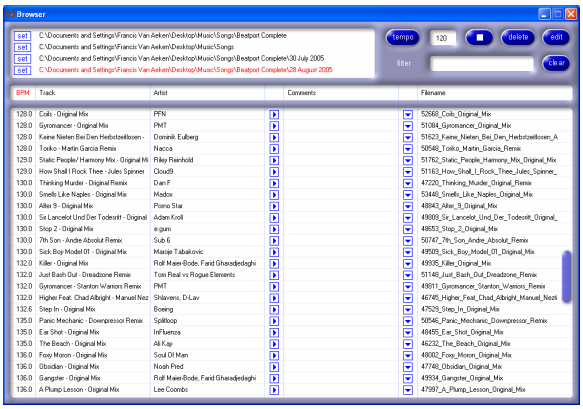

Figure 2. Jackson: browser

The browser features a previewer which is shown in figure 3. The previewer shows the rhythmic structure of any song and that lets the DJ listen to any part of it.

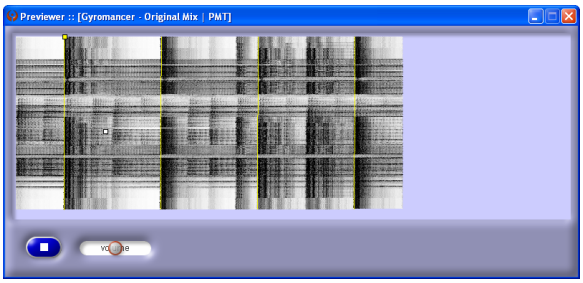

Figure 3. Jackson: previewer

The previewer shows a song as a grid of pixels of which the intensity corresponds to the energy of the signal in time. Black pixels indicate high energy while white pixels indicate the lack of energy, i.e. silence. Each row of the grid represents one measure of music. Time goes from left to right and from top to bottom.

Thanks to the visual representation in the previewer, the DJ can immediately see the rhythmic structure of the song and its evolution in time. An experienced user can readily identify the song in Figure 3 as a fixed-tempo break-beat song having a major breakdown after one 3rd of the song. Indeed, the beats vertically align and form straight vertical lines, meaning that each measure has exactly the same length. Also, the kick drums and snare drums (black elements) do not divide each measure in four equal parts. The pattern is more irregular, suggesting a break-beat rather than a four-to-the-floor song. Consecutive measures (lines) that have no black elements constitute breakdowns.

2.3. Analyzing songs

The key feature of Jackson is that songs are automatically beat-matched and therefore always play in sync. Musical metadata makes this possible.

Before a song can be played it must be rhythmically analyzed. Our application comes with a sophisticated beat-mapping tool that makes this an easy and instructive process. Figure 4 shows the beat-mapper's window.

The beat-mapper uses the same visual representation as the previewer. The system initially estimates the tempo. The user can then adjust the tempo or add markers on the onset of beats to take into account changes in tempo.

For electronic dance music the automatic tempo estimation is almost always 100% on the mark. In these cases, the user only has to put a marker on the first beat of a measure for the system to have complete information about the rhythmic structure of the song.

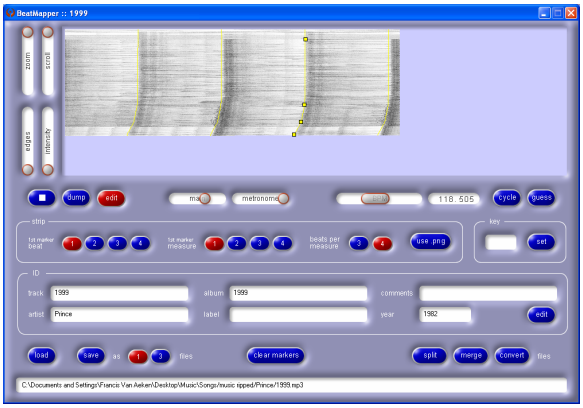

Figure 4. Jackson: beat-mapper

Different from electronic music most songs played live feature changes in tempo. Figure 5 shows the beat-map of the song Gigantic by the Pixies. Small squares indicate the position of markers. As one can readily see, steady tempo is not the trademark of this fine group. Indeed, in this case the beats no longer form black vertical lines. Beat-mapping such a song requires placing many more markers than the single one needed for electronic songs.

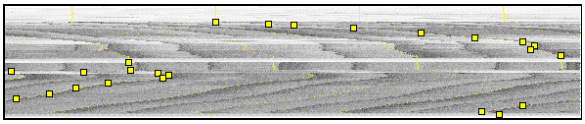

Figure 5. Pixies: Gigantic

Once a song is analyzed, the musical metadata is embedded in it and the DJ can mix and manipulate the song in all freedom.

This approach works with all styles of music, whether the tempo is fixed or not. Such styles often pose problems for traditional DJs. Thanks to a time-stretcher built into the application, songs in different styles and tempos can be combined into one rhythmically and harmonically consistent whole.

Apart from rhythmical metadata, one can also specify harmonic metadata describing the musical key of the song. Editorial metadata such as the name of the songs or its label can also be edited and embedded with the beat-mapper.

3. Musical metadata

Our metadata describes the musical structure of the song. We store the position of the individual beats and specify how they are grouped in measures and groups of measures. We also store a visual representation of the song, editorial and harmonic information, settings related to the beat-mapping process and (if applicable) the positions of mp3 frames.

When we started building the application, we were focusing on practical results rather than on standards. MPEG-7 still was fairly academic and few people were using it.

Given the market a couple of years ago, we focused on the WAV, MP3 and WMA formats. MP3 (or rather ID3v1 & ID3v2) and WMA have some support for metadata, but do not define fields for all metadata we use. Moreover, the two metadata formats are not interchangeable. Given all those factors, we decided to design our own format. We encode our metadata in a temporary file using a format structured much like the excellent PNG file format. This file is then embedded in the audio file. Optionally one can embed a second file that offers an alternative visual representation of the song.

4. MPEG-7

The more the metadata is accessible to different applications, the more its power can be leveraged. As standards mature, we must adhere to them as much a possible.

MPEG-7 is currently the most general standard to describe multimedia material [2][3]. Different from other MPEG standards, it focuses on metadata and not on the encoding of the data itself. MPEG-7 uses

Descriptors to describe low-level features and Description Schemes to describe higher-level features.

MPEG-7 specifies the Audio Waveform Descriptor that maps directly on the part of our metadata that visually describes the song.

On a higher level MPEG-7 specifies a Segment Description Scheme that represents the spatial, temporal or spatiotemporal structure of the audio-visual content. We can use a hierarchy of segments to describe the musical structure of a song. A song can be described as a sequence of sections composed of measures. Each of these is built up of beats. Unfortunately there seems to be no way to explicitly specify the musical meaning of this hierarchy (although the segments can be associated with low-level Descriptors and the relations between segments can be detailed). The mismatch seems to stem from the fact that the main goal of MPEG-7 is to facilitate searching and querying of material rather than their manipulation. This is an issue that we have to investigate further.

At the systems level, MPEG-7 metadata is written as XML, but can be embedded as compact BiM.

5. Other applications of musical metadata

There are relatively few DJs in the world. A DJ has to invest in equipment, has to acquire skills and has to continuously update his or her music library.

Technology like ours can substantially lower the threshold for people to start mixing: the investment in equipment is minimal if the person has a computer already. Also, the DJ no longer has to acquire purely technical skills like beat-matching. Still, not everybody might feel the burning need to become a DJ.

However, most people do love to listen to music and musical metadata can enhance today's listening experience. For example, it can be at the basis of virtual DJs that build musically meaningful sets according to the taste and mood of the listener.

5.1. The changing listening experience

To listen to music, people have been going to concerts and parties, switching on the radio or television or putting on a cassette, vinyl album or CD.

In all these cases user control over the experience was limited. In the case of cassettes, albums and CDs the format of the physical carrier determined the duration of the experience and the sequencing of the individual songs.

Electronic delivery of music over the Internet is doing away with the limitations of a physical carrier. People typically download individual songs rather than albums and sequence them the way they want to.

This new freedom implies new responsibilities: people have to be their own DJs now. Luckily, in the near future they can choose to delegate this responsibility to virtual DJs.

5.2. Virtual DJs

The technology we have developed can be applied to the development of virtual DJs that produce mixes automatically. We expect to find virtual DJs in desktop PCs, hi-fi equipment and portable audio players. Radio stations also will appreciate this technology.

We are currently building a first version of a virtual DJ on top of Jackson. Following a tempo trajectory and a virtual crate specified by the user, the DJ will create a custom mix.

The system will be driven by rhythmic and harmonic metadata so that the mix will be technically flawless. For an optimal listening experience, however, we will have to add metadata about the cultural and emotional aspects of the songs.

5.3. Personal radio

A virtual DJ does not need to be embedded in a hardware device. Nor does it need to run on a personal computer. It can just as well reside on a server that people can connect to for a personal radio experience.

In this business model people pay per time unit rather than per song. It is clear that the quality of the virtual DJ must be high for such a system to have appeal. Again, apart from metadata on the rhythmic and harmonic level, metadata describing the mood and the (sub)cultural identity of songs will drive these systems.

5.4. Intelligent compilations

We also see the potential of intelligent compilations of music (on CD-ROM for example) in which a collection of songs and one or more virtual DJs are combined.

Depending on the preferences of the listener, including the choice of DJ, a different mix will be produced. This new listening experience will be interactive and dynamic rather than passive and static.

6. Distribution of musical metadata

Currently no online distributor of music sells songs having musical metadata embedded in them. It is up to the end user to add the metadata to the files. This typically means musically analyzing the song or retrieving the metadata from another place (like a central server). Our application comes with an excellent tool to analyze music, but songs having an irregular tempo can take some time to analyze.

It is obviously in the interest of the user that this metadata is included at the source. Given the wealth of applications that can benefit from musical metadata, we expect many distributors in the future to embed metadata in the songs they sell. Standardization will be key to the success of musical metadata. It is therefore important that companies like us work closely together with competitors and standards organizations like the MPEG-7 Consortium.

7. Conclusion

Computer technology offers us great opportunities to enhance our consumption of music. The application 'Jackson' demonstrates that recorded songs do not have to be final products, but can be raw material to play with. Musical metadata is the key to this functionality. The MPEG-7 standard offers us a common language to write the metadata in. Many other applications can benefit from this metadata and we hope that music publishers and distributors will grasp this opportunity sooner rather than later.

8. References

[1] http://jacksondj.com

[2] François Pachet, “Knowledge Management and Musical Metadata”, *Encyclopedia of Knowledge Management*, Schwartz, D. Ed. Idea Group, 2005.

[3] José M. Martínez, *MPEG-7 Overview (version 9)*, ISO/IEC JTC1/SC29/WG11N5525, Pattaya, March 2003.