

# Interactive Music with Active Audio CDs

Sylvain Marchand, Boris Mansencal, and Laurent Girin

LaBRI – CNRS, University of Bordeaux, France  
{sylvain.marchand,boris.mansencal}@labri.fr  
GIPSA-lab – CNRS, Grenoble Institute of Technology, France  
laurent.girin@gipsa-lab.grenoble-inp.fr

**Abstract.** With a standard compact disc (CD) audio player, the only possibility for the user is to listen to the recorded track, passively: the interaction is limited to changing the global volume or the track. Imagine now that the listener can turn into a musician, playing with the sound sources present in the stereo mix, changing their respective volumes and locations in space. For example, a given instrument or voice can be either muted, amplified, or more generally moved in the acoustic space. This will be a kind of generalized karaoke, useful for disc jockeys and also for music pedagogy (when practicing an instrument). This demonstration shows that this dream has come true, with active CDs fully backward compatible while enabling interactive music. The magic is that “the music is in the sound”: the structure of the mix is embedded in the sound signal itself, using audio watermarking techniques (patent pending).

**Key words:** interactive music, compact disc, audio watermarking, source separation, sound spatialization

## 1 Active Listening

Composers of acousmatic music conduct different stages through the composition process, from sound recording (generally stereophonic) to diffusion (multiphonic). During live interpretation, they interfere decisively on spatialization and coloration of pre-recorded sonorities. For this purpose, the musicians generally use a(n un)mixing console. With two hands, this requires some skill and becomes hardly tractable with many sources or speakers.

Nowadays, the public is also eager to interact with the musical sound. Indeed, more and more commercial CDs come with several versions of the same musical piece. Some are instrumental versions (for karaoke), other are remixes. The karaoke phenomenon gets generalized from voice to instruments, in musical video games such as *Rock Band*<sup>1</sup>. But in this case, to get the interaction the user has to buy the video game, which includes the multitrack recording.

Yet, the music industry is still reluctant to release the multitrack version of musical hits. The only thing the user can get is a standard CD, thus a stereo mix, or its dematerialized version available for download. The CD is not dead: imagine a CD fully backward compatible while permitting musical interaction. . .

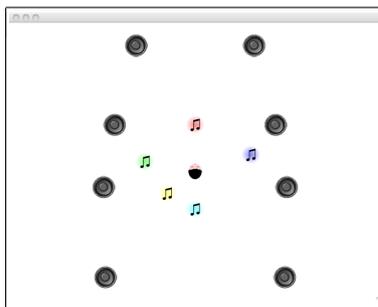
<sup>1</sup> see URL: <http://www.rockband.com>

## 2 Active CD Player

We present the proof of concept of the active audio CD, as a player that can read any active disc, decode the musical structure present in the sound signal, and use it to perform high-quality source separation. Then, the listener can see and manipulate the sound sources in the acoustic space. Our system is programmed in C++ language, the spectral processing is done using FFTW<sup>2</sup>, and the portable audio input / output are done using JACK<sup>3</sup>. It is composed of two parts.

First, a CD reader extracts the audio data of the stereo track and decodes the musical structure embedded in the audio signal. This additional information consists of the combination of active sources for each time-frequency atom. As shown in [1], this permits an informed source separation of high quality (patent pending). In our current system, we get 5 individual tracks out of the stereo mix.

Second, a sound spatializer is able to map in real time all these sources to any position in the acoustic space. Our system supports either binaural (headphones) or multi-loudspeaker configurations. As shown in [2], the spatialization is done in the spectral domain, is based on acoustics and interaural cues, and the listener can control the distance and the azimuth of each source (see Fig. 1).



**Fig. 1.** From the stereo mix stored on the CD, our player is allowing the listener (*center*) to manipulate 5 sources in the acoustic space, using here an octophonic display.

**Acknowledgments.** This research was partly supported by the French ANR (*Agence Nationale de la Recherche*) DReaM project (ANR-09-CORD-006).

## References

1. Parvaix, M., Girin, L.: Informed Source Separation of Underdetermined Instantaneous Stereo Mixtures Using Source Index Embedding. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Dallas (2010).
2. Moubia, J., Marchand, S., Mansencal, B., Rivet J.-M.: RetroSpat: a Perception-Based System for Semi-Automatic Diffusion of Acousmatic Music. In: Proceedings of the Sound and Music Computing Conference, pp. 33–40, Berlin (2008).

<sup>2</sup> see URL: <http://www.fftw.org>

<sup>3</sup> see URL: <http://jackaudio.org>