

Bridging the Music Semantic Gap

by Xavier Serra, from the Pompeu Fabra University of Barcelona

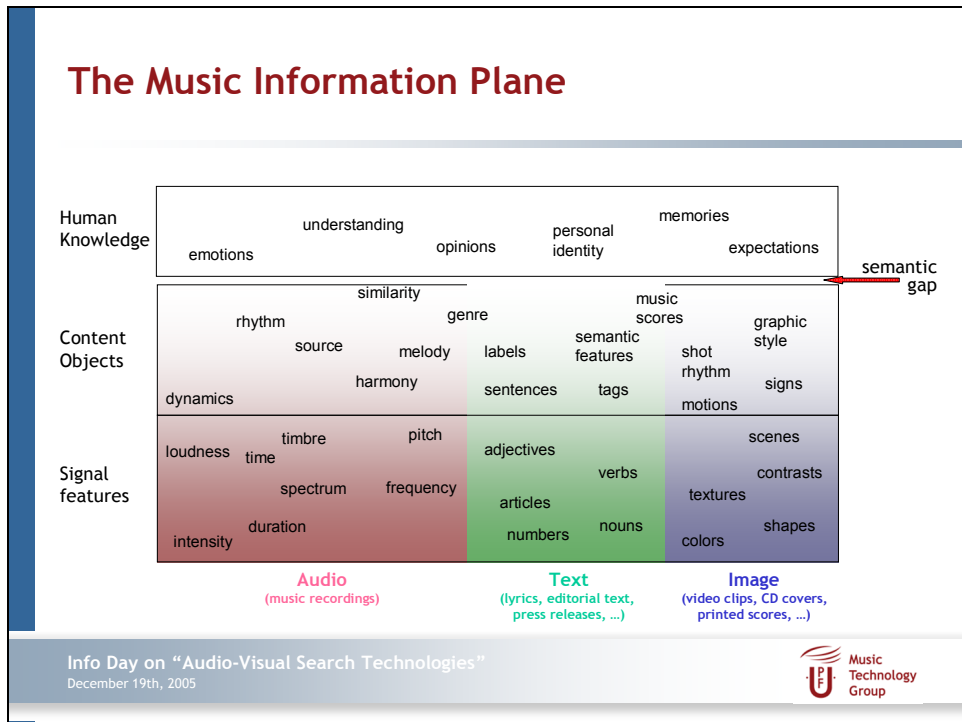
email: xserra@iua.upf.edu

(talk given at the Information Day on “Audio-Visual Search Technologies” organized by the Commission services of DG Information Society of the EC in Brussels on Dec 19th 2005)

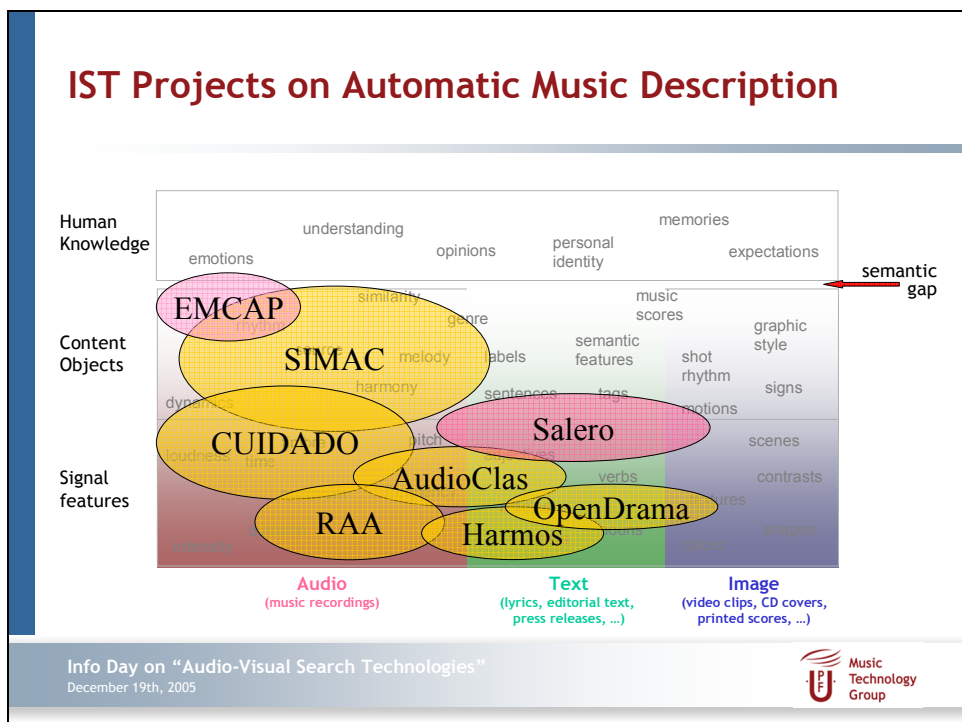
Abstract: The talk will give a brief view on how to bridge the semantic gap in the particular case of music information. This view is based on the analysis of the research that has been carried out within several IST projects and the identification of the future research developments that could push our current limits. It will start by describing the concept of the “Music Information Plane”, a plane where the relevant information about music is placed in the context of two dimensions; one is the abstraction level of the descriptors (from physical to knowledge levels) and the other includes the different media (audio, text, image). On this plane the developments of various IST projects are viewed, the placement of the semantic gap identified and finally the developments that should allow us to cross that gap are proposed.

The slide features a white background with a blue vertical bar on the left side. The title "Bridging the Music Semantic Gap" is centered in a dark red font. Below the title, the presenter's name "Xavier Serra" is centered in a black font, followed by "Music Technology Group - Pompeu Fabra University" and "Barcelona, Spain" in a smaller black font. The website "http://www.iua.upf.es/mtg/" is centered at the bottom. A footer bar at the very bottom contains the text "Info Day on 'Audio-Visual Search Technologies'" and "December 19th, 2005" on the left, and the "Music Technology Group" logo on the right.

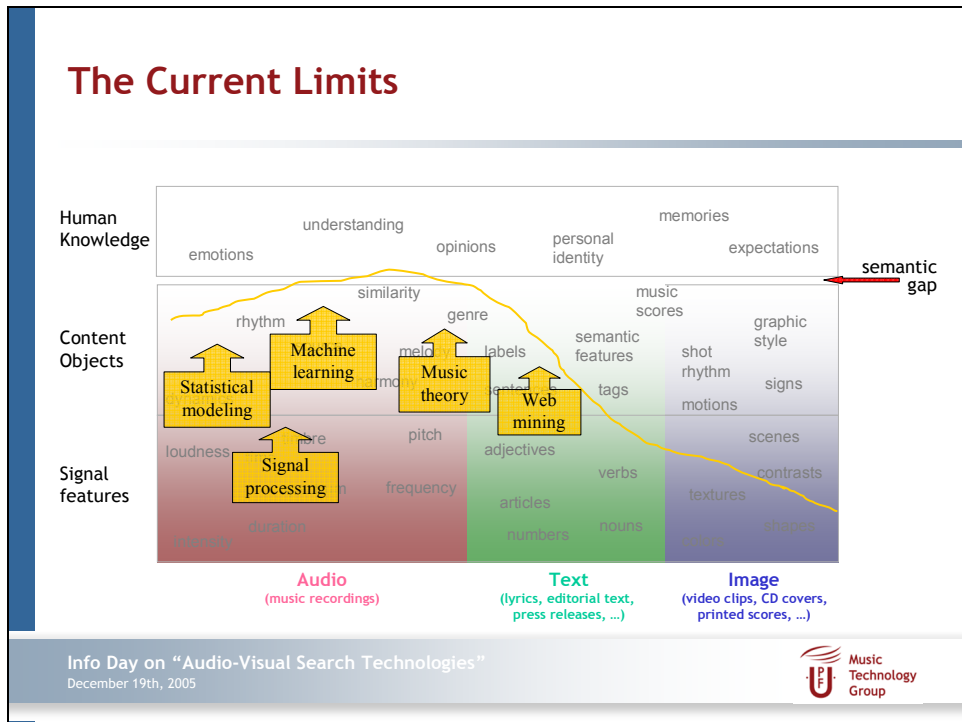
Slide 1: The basic message that I would like to transmit in this very short talk is that we might be quite close from bridging the semantic gap in the music information domain. The developments carried out in various IST projects have been very important in advancing the research in music content processing and we have some clear ideas about what to do next.



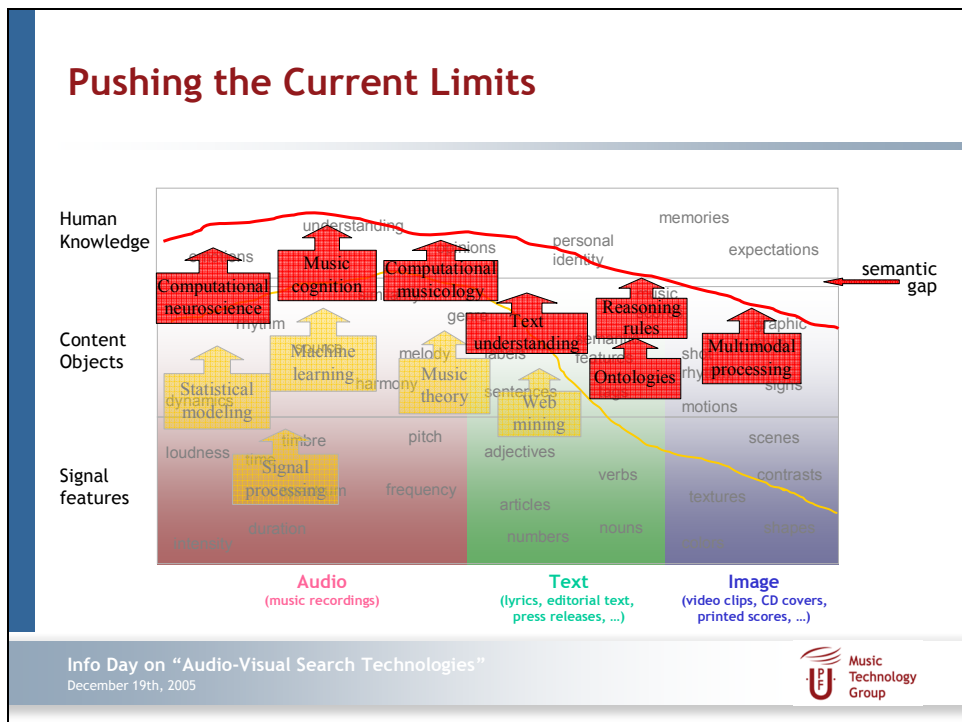
Slide 2: This is the "Music Information Plane". The vertical axis represents the different abstraction levels of the musical descriptors: Signal Features, Content Objects and Human Knowledge. The "Semantic Gap" is located right before the Human Knowledge level. The horizontal axis divides the different musical media types: audio recordings, textual information, and images; these media divisions are very relevant for signal level descriptors but they fade away as we reach the knowledge level descriptors.



Slide 3: In the last few years a number of IST projects have focused in the automatic description of music. The projects shown in the slide are the ones in which our group has been part of. Good progress has been done in the extraction of relevant signal features and content descriptors from music recordings and text information. SIMAC is close to be finished and has been very successful at pushing our music descriptors towards high semantic levels. EmCap and Salero are two projects that just have started that should push our descriptors a bit higher.



Slide 4: For many years Signal Processing has been the main discipline used to automatically generate music descriptors. More recently Statistical Modeling, Machine Learning, Music Theory and Web Mining technologies have also been used to push up the semantic level of our descriptors.



Slide 5: The current approaches to automatic music description, which are mainly bottom-up, will not allow us to bridge the semantic gap. We need an important shift in our approach. The music description problem will not be solved by just focusing on the audio signals; a Multimodal Processing approach is needed. We also need top-down approaches based on Ontologies, Reasoning Rules, Music Cognition, ... and we also need to improve our understanding of the auditory cortex and its interrelation with the rest of the brain.

Bridging the Music Semantic Gap

Xavier Serra

Music Technology Group - Pompeu Fabra University

Barcelona, Spain

<http://www.iaa.upf.es/mtg/>

Info Day on "Audio-Visual Search Technologies"
December 19th, 2005



Slide 7: We are now viewing an explosion of the practical applications coming out from this research: Music Identification systems, Music Recommenders, Play-list Generators, Search Engines, Music Discovery and Personalization systems, ... This is just the beginning. If we succeed in bridging the semantic gap there will be no limit to the applications that could be develop. We might be closer in bridging the semantic gap in music than in any other knowledge domain. Music was a key factor in taking Internet from its text-centered origins to being a complete multimedia environment. Music might do the same for the semantic-web.

Links:

1. Music Technology Group (MTG): <http://www.iaa.upf.es/mtg/>
2. Semantic Interaction with Music Audio Contents (SIMAC): <http://www.semanticaudio.org/>
3. Emergent music cognition in natural and artificial interactive systems (EmCAP): <http://emcap.iaa.upf.es/>