

Three-level representation of audio and video data: from the physical level to the semantic level, via the measurement level

Roberto Bresin, KTH CSC TMH

Research studies in multimodal communication often involve the use of analysis and synthesis tools for the understanding of the content being communicated. In the Sound and Music Computing (SMC) community an establish framework for data representation is that based on a multi-layered structure made by three levels: physical level, measurement level, and semantic level. The physical level is that corresponding to the raw signal (for example audio or video) as we record it. The measurement level gives a representation of the raw signal in terms of physical variables (for example duration of notes in audio, and quantity of motion in video). The semantic level gives a high-level representation of the information extracted at the measurement level. Examples and demonstrations will be provided during the talk.