Tutorial on Digital Music Representations

Geoffroy Peeters Gérard Assayag
Ircam – Centre Pompidou / CNRS UMR 9912
1, place Igor-Stravinsky
F-75004 Paris, France
+33 (0)1 44 78 14 22 +33 (0)1 44 78 48 58
geoffroy.peters@ircam.fr gerard.assayag@ircam.fr

1. OBJECTIVES

This tutorial will present state of the art representation schemes (formats, protocols, algorithms, environments) used in computer storage, search, transfer and production of music in the broad sense of the word. Established and proposed standards for file formats, database metadata, protocols, markup languages will be presented and discussed as well as emerging ideas on dynamic representation schemes in relation with interactivity. Representations will be discussed according to their relative efficiency/expressivity and their adaptation to different kind of applications.

Music representations will be discussed in the context of their efficiency for different application objectives: indexing and information retrieval, computer-assisted composition and performance, music stream generation for games and web applications, multimedia frameworks.

The digital rights management aspects will be invoked when relevant but not intensively discussed (this could be the topic of another tutorial).

The tutorial will start with the presentation of a few conceptual issues: the sound substance of music, its structural relationships, the notion of content in musical media.

These three poles will be crossed with an important dichotomy between static and dynamic representation schemes.

This conceptual framework will help situate the numerous representation schemes that will be described.

1.1 Intended audience and expected level

The tutorial is aimed at the non-specialist having a general basic knowledge of computers and sound/music.

1.2 Course Material

Handout material will be distributed.

1.3 Instructors' biographies

Dr. Geoffroy Peeters received his diploma in electrical engineering from the Catholic University of Louvain (Belgium) and his M.Sc. and Ph.D. (with the "félicitations du jury" distinction) from University of Paris VI (France) in acoustics, signal processing and computer science applied to music. During his Ph.D., he developed SINOLA, a new technique for the processing of speech and instrument's sounds based on both temporal and spectral modeling. Since 1999, he has been working on the development and the editing of the new ISO MPEG-7 Audio standard. Since 2001, he has been working at Ircam for the CUIDADO I.S.T. European project on music information retrieval. He is acting as coordinator and researcher in the field of audio feature extraction. His main areas of research are instrument timbre description, sound recognition and music audio structuring.

Gérard Assayag is currently head of the Music Representation Research Group at Ircam. Born in 1960, he studied computer science, music and linguistics. He won a national contest launched in 1980 by the French Informatics Agency on "Art and the Computer". In the mid-eighties, he wrote the first Ircam environment for score-oriented Computer Assisted Composition. He invented and wrote with Carlos Agon the OpenMusic environment which is currently used by hundreds of composers/musicologists throughout the world and is taught in several music institutions. He is currently coordinator of ATIAM (a Masters/Ph.D. course in Acoustics, Signal Processing, Computer Science Applied to Music co-organized by Ircam and four French universities). His research results in computer music modeling are regularly published in proceedings, books and journals. Gérard Assayag is also founding member of the SFIM (the French Society of Music Informatics) and member of the FWO Research Society on Foundations of Music Research. His research interests are focused on music representation issues and include computer language paradigms, machine learning, logic, constraint and visual programming, computational musicology, music structure modeling and computer-assisted composition.

2. OUTLINE

2.1 Introduction

How is music represented in a computer? This question is becoming critical as online distribution of music in its various forms is becoming established as a major trend, and an ever increasing number of people searches, downloads, reads, listens, analyzes, composes and processes it.

Of course, the problem has been addressed since the very beginning of computer music in the fifties, when Hiller, Xenakis and others encoded basic music parameters such as pitch, duration and dynamics into numbers and implemented composition rules that were acting upon theses numbers.

After this initial phase, involving simple encoding of music symbolic atoms that were provided as input to — and recovered as output from — rather sophisticated algorithms, the domain was overwhelmingly invested by the sound signal generation and processing technology. This is understandable because the scientific background was deeper and the industrial issues were larger in that field. Thus the research and development manpower devoted to digital sound was just incomparable to the one devoted to representation of music as a system of forms, structures, cultural and esthetic content.

Composers, musicologists, AI researchers, have always known that music is not only sound, and cannot be represented and processed only through the sound signal. This obvious fact was somewhat forgotten for the reasons we have outlined above. It is coming back to life as new industrial issues, new cultural practices and new consumer behaviors are emerging. Music phenomena have to be understood in their deep structure in order to take up the challenge raised by this mutation.

2.2 Concepts

2.2.1 Why music is not only sound.

2.2.2 From the sound surface to the formal structure.

2.2.3 From the score surface to the formal structure.

2.2.4 Structure levels in music and adapted coding paradigms.

2.2.5 Why structure is not exactly content (content is described, structure is represented)

2.2.6 Structuring information about content.

2.2.7 Interoperable framework versus metadata scheme

2.3 Signal representation basics

2.3.1 Time domain

2.3.2 Time-frequency domain

2.3.3 Granular representations

2.3.4 Physical Modelling representations

2.4 Static representation formats and standards

2.4.1 Surface level formats

- Sound
  - Simple formats
    - Time domain (AIFF, SD2F, WAV…)
    - Time frequency (SDIF…)
  - Compressed formats
    - MPEG1 (MP3), MPEG4, ATRAC…
- Gesture/instrumentation
  - MIDI
  - Midifile
- Score
  - SMDL, NIFF , GUIDO, ETF…

2.4.2 Content description formats

- Universal multimedia access (MPEG7, MPEG21)
- Hytime

2.4.3 Structure based formats

- OpenMusic container scheme
- Humdrum
- Common Music
- …

2.5 Dynamic representations schemes

2.5.1 Statistical representations

- IPG (Incremental Parsing and Generating)
- PST (Predictive Suffix Tree)

2.5.2 Automata theory representations

- Factor Oracle
- Superposition automata

2.5.3 Constraint and Logic programming representations

- PWConstraints
- Situation
- OMConstraints

2.5.4 Functional representation

- Elody
- Nyquist
- OpenMusic maquettes