Jornada de Seguimiento de Proyectos, 2005
Programa Nacional de Tecnologías Informáticas

Content-Based Music Processing
(Procesamiento de Música Basado en su Contenido)
TIC2003-07776-C02

Xavier Serra Casals*
Institut Universitari de l'Audiovisual
Universitat Pompeu Fabra

Rafael Ramirez Melendez
Institut Universitari de l'Audiovisual
Universitat Pompeu Fabra

Abstract

From the beginning, the aim of the project has been to work on different aspects of content-based music processing. So far, we have studied and developed tools for musical content extraction, modeling, and processing. Specifically, we have been working on (1) the analysis and manipulation of low-level audio descriptors by spectral modeling analysis and synthesis techniques, (2) the extraction of high-level musical attributes (musical structure, transitions and articulation) from these low-level audio descriptors and (3) the use of artificial intelligence techniques such as case-based reasoning and inductive logic programming for content-based melody processing. We have also been working on the design and implementation of a special-purpose concurrent real-time programming language for music data processing

1 Project Objectives

Content-based processing is a research area which includes diverse aspects such as analysis, indexing, search and transformations of signals produced by am audiovisual source. In a musical context, content-based signal processing is a fruitful research area which has already become one of the main research areas in computer music. In the area of video and image, there are a number of active research groups, both in Spain and in Europe. In contrast, in the audio area there is much less activity. In fact the IUA coordinator group of this project is one of the few research groups which has contributed to the Mpeg-7 standard.

In this context, this project has proposed to work on different aspects of content-based music processing and develop tools for the extraction, modeling and processing of these contents. We investigate topics such as the analysis and manipulation of low-level audio descriptions, as well as

* Email: xserra@iua.upf.es

methods for extracting high-level musical attributes from the low-level attributes and the use of artificial intelligence techniques for musical analysis and transformation.

In the past, the two research groups involved in this Project have developed a tool for the analysis of musical descriptions. The tool is able to automatically extract a XML description from an audio signal input. We have also developed a synthesis tool which is able to interpret XML data in order to obtain a melody which is synthesized and stored in a file or directly sent o the sound card of a computer. This tool is based on previous work on a saxophone synthesizer. Currently, we are developing a tool for the expressive transformation of monophonic Jazz melodies.

The experience and results obtained in past collaborations of the IUA and the IIIA, place us in a privileged position for accomplishing the project objectives.

The objective of this Project is to work on different aspects of content-based audio processing. The project investigates and develops tools for musical content extraction, modeling and processing. In particular, it investigates:

1. The analysis and manipulation of low-level audio descriptors. We use spectral models analysis and synthesis techniques previously developed by members of the Music Technology Group in the IUA.

2. The extraction of high-level musical attributes (transitions, articulation and musical structure) from the low-level descriptors.

3. The use of musical knowledge for the analysis and transformation of sound. This musical knowledge is modeled using artificial intelligence techniques such as case-based reasoning and inductive logic programming. In the case of the case-based reasoning, we investigate new techniques for the tasks of adaptation and maintenance.

4. The design and implementation of a concurrent real-time programming language for modeling musical processes and to be used in audio processing applications.

On the one hand, The IUA has successful experience in sound manipulation techniques (analysis and synthesis) using spectral modeling. It also has wide experience in high-level programming environments in the area of audio processing as well as in concurrent programming languages design. On the other hand, the IIIA has successful experience in musical knowledge representation and case-based reasoning. The technologies from both groups complement each other and provide an effective approach for content-based music processing.

In order to describe the musical content of a sound source we rely on the spectral modeling and techniques developed by the music technology group of the IUA. These techniques allow the extraction of numerical values for temporal, dynamic, timbric and melodic descriptors of a

musical phrase. We have used some of these descriptors for content description and search in musical notes databases. The solid base that these techniques provide allows the development of a new level of descriptions whose utility can be maximized by the knowledge induced by manipulating a data set of examples using cased-based reasoning and other artificial intelligence techniques.

2 Project Progress

IUA

At the IUA, we have pursued two research lines: on the one hand, we have investigated the extraction of different symbolic attributes from monophonic audio recordings and the intelligent synthesis of audio from symbolic descriptions. On the other hand, we have explored different machine learning techniques in the context of expressive music performance modeling. In parallel to these two main research lines, we have worked on the design and implementation of a concurrent programming language based on first order logic for the specification, execution and verification of musical processes. The different research lines have generated a number of publications (see IUA publications in next section).

Regarding the extraction of different symbolic attributes from monophonic audio recordings, we have worked on algorithms and tools for the extraction of two types of attributes: note-level attributes and intra-note attributes. The former ones characterise properties of note in a melody, e.g. pitch, duration, onset and mean energy, while the later ones characterise internal properties of a note, e.g. attack, sustain, and release. We have used both types of attributes for automatically inducing expressive performance models. We have also explored different approaches to synthesising audio from symbolic descriptions, the most promising being using samples together with a predictive model determining the most suitable note (and transition) to be played ina determinate musicl context.

Regarding the application of different machine learning techniques to the task of expressive music performnace modeling, our approach has been the following one: Instead of manually modelling expressive performance and testing the model on real musical data, we let a computer use machine learning techniques to automatically discover regularities and performance principles from real performance data. We have applied both inductive regression techniques, e.g. model trees, support vector machines, and rule-based classification techniques, e.g. inductive logic programming, in order to induce expressive music performance models. We have devised and implemented machine learning algorithms to fit our needs.

On the one hand, regression methods are considered to be black-box in the sense that it is very difficult (or impossible) to understand the predictions they produce. Black-box statistical

approaches may be good at deriving predictions from data, but formulating understandable rules from the analysis of data is something entirely different from formulating predictive models from that data. On the other hand, rule-based classification methods are good at explaining the predictions they provide but are restricted to a set of discrete classes as prediction space. Our problem at hand is one that requires the prediction precision of regression methods for generating accurate solutions (i.e. expressive performances) but at the same time it is highly desirable to be able to explain the system predictions. Having this in mind, we have not limited ourselves to anyone of the two techniques and instead we have explored both approaches applying a representative sample of both regression and classification methods.

IIIA

Similarity Measures Task:

Taking as starting point the results achieved in the Tabasco project (TIC 2000-1094-C02), we proposed new different similarity measures for comparing musical sequences (scores and performances) at different abstraction levels. We compared their behaviour in a database of jazz standards and determined the possible usage of each measure. The research on this has produced the following publications: [4] [5] [6] [7] [8] (see IIIA publications in next section). This task finalized successfully.

Adaptation Task:

The goal of the task is to propose new adaptation mechanisms for CBR systems that construct solution to problems that require a complex solution structure (such as design problems). We have proposed a formalism based on components (see [2]) that can be incorporated into the Case-Based Reasoning inference cycle. This formalism is based on our previous work on the "Constructive Adaptation" algorithm. We are still working on this task and the plan is to improve the current components and to perform experiments in the musical domain.

Case-Based Reasoning Maintenance Task:

The maintenance tool has as a goal to guarantee or increase the quality of the solutions a case-based system is proposing. In problems like our musical domain, where the representation of the problem requires complex structures using different abstraction levels, the maintenance is a critical and open issue. This task was fired just few months ago and started analysing the current state of the art. We have produced a first result (see [1]) where we show how an improvement on the "Competence Groups" approach can be applied for analysing in a more powerful way the case bases. We provide a new measure, based on analysing the quality of solutions, which allows a better characterization of the case base structure. Using this measure we have developed a visualization tool that allows a 3D navigation through the case base space. We are still working in this task and we are investigating how to develop a self-model of the reasoning cycle for automatically analyse the quality of solutions proposed and fire the appropriate maintenance mechanisms.

3 Results Indicators

During its first year, several researchers have joined the project:

- Emilia Gomez, Engineer in telecommunications and PhD student (obtained her Diploma de Estudios Avanzados (DEA) in 2004) in the computing and digital communication program of the UPF. She worked on research tasks such as signal processing and analysis.
- Esteban Maestre, Engineer in telecommunications, is currently registered as PhD student (first year) in the computing and digital communication program of the UPF. His research is devoted to audio analysis, description and synthesis using signal processing techniques.
- Amaury Hazan, Engineer in computing, is currently registered as PhD student (first year) in the computing and digital communication program of the UPF. His research is devoted to the design and implementation of artificial intelligence techniques..
- Marc Vinyes, Engineer in telecommunications and mathematician, is currently registered as PhD student (1st year) in the computing and digital communication program of the UPF. He has been working part-time on audio analysis and description.
- Maarten Grachten, Computer Scientist and PhD student (obtained his Diploma de Estudios Avanzados (DEA) in 2004). His Phd work is devoted on the research on developing a Case-Based Reasoning system for generating expressive tempo-transformations of jazz melody performances. His work proposal received the "ICCBR'03 Best Paper Award" at the International Conference on Case-Based Reasoning.
- Alejandro García, Computer Scientist and PhD student (first year). His research is be devoted on the use of Maintenance polices for improving the performance of Case-Based Reasoning systems.
- Jose María Abasolo. The Phd work of Jose María Abásolo (he obtained his Diploma de Estudios Avanzados (DEA) in 2004) is devoted on the research on a Component-Based framework for developing Case-Based Reasoning Applications. The research is specially focused on providing methodological assistance on solving analytical problems such as the use of Case-Based Reasoning techniques for generating expressive tempo-transformations.

Emilia Gomez is no longer working directly in this project, she is currently working in SIMAC, a EU project.

So far, the project has generated the following publications:

IUA

[1] Rafael Ramirez, Amaury Hazan, (2005) ``A Learning Scheme for Generating Expressive Music Performances of Jazz Standards``, Proceedings International Joint Conference on Artificial Intelligence IJCAI 2005.

[2] Rafael Ramirez, Andrew Santosa (2005), ``Formal Verification of Concurrent and Distributed Constraint-based Java Programs``, Proceedings IEEE International Conference on the Engineering of Complex Computer Systems, IEEE Press.

[3] Rafael Ramirez, Amaury Hazan (2005), ``An Approach to Expressive Music Performance Modeling``, Proceedings of 118th Audio Engineering Society Convention.

[4] Maestre, E. Gómez, E. (2005) 'Automatic characterization of dynamics and articulation of expressive monophonic recordings' Proceedings of 118th Audio Engineering Society Convention.

[5] Rafael Ramirez, Amaury Hazan, (2005), `` Modeling Expressive Music Performance in Jazz``, Proceedings International Florida Artificial Intelligence Research Society Conference, AAAI Press.

[6] Rafael Ramirez, Amaury Hazan, (2005) , `` Understanding Expressive Music Performance Using Genetic Algorithms``, Proceedings European Workshop on Evolutionary Music and Art, Springer-Verlag LNCS.

[7] Rafael Ramirez, (2004), ``Inductive Logic Programming and Music``, Proceedings International Computer Music Conference, ICMA Press.

[8] Rafael Ramirez, Amaury Hazan, (2004), ``Learning Expressive Performance Rules in Jazz``, Proceedings International Computer Music Conference, ICMA Press.

[9] Rafael Ramirez, Amaury Hazan, Emilia Gomez, Esteban Maestre, (2004), ``Understanding Expressive Transformations in Saxophone Jazz Performances Using Inductive Machine Learning``, Proceedings Sound and Music Computing International Conference.

[10] Rafael Ramirez, Amaury Hazan, (2004), ``Rule induction for expressive music performance modeling``, ECML Workshop Advances in Inductive Rule Learning.

[11] Rafael Ramirez, Juanjo Martinez, Andrew Santosa, (2004),``Constraint-based Concurrent and Distributed programming in Java``, Proceedings International Symposium on Distributed Computing and Applications to Business, Engineering and Science, Wuhan University Press.

[12] Rafael Ramirez, Amaury Hazan, Emilia Gomez, Esteban Maestre, (2004) ``A Machine Learning Approach to Expressive Performance in Jazz Standards``, Proceedings International Workshop on Multimedia Data Mining (in conjunction with ACM SIGKDD International Conference on Knowledge Discovery & Data Mining).

[13] Rafael Ramirez, (2004), ``Concurrent and Distributed Programming Using Constraint Logic Programs``, Proceedings ACM Symposium on Applied Computing, ACM Press.

[14] Rafael Ramirez, (2003), ``Inductive Logic Programming for Learning Musical Rules``, Proceedings International Conference on Applied Artificial Intelligence.

IIIA

[1] M. Grachten, F.A. García, J. L. Arcos, (2005), Navigating through Case Base Competence. 6th International Conference on Case-Based Reasoning ICCBR'05. (in press)

[2] E. Plaza (2005), Cooperative Reuse for Compositional Cases in Multi-Agent Systems. 6th International Conference on Case-Based Reasoning ICCBR'05. (in press)

[3] J. L. Arcos (2004), Improving the Quality of Solutions in Domain Evolving Environments. 7th European Conference on Case-Based Reasoning ECCBR'04. Peter Funk and Pedro A. Conzález-Calero eds. Lecture Notes in Artificial Intelligence, nº 3155. Springer-Verlag, pp. 464-475.

[4] M. Grachten, J. L. Arcos, R. López de Mántaras (2004), TempoExpress, a CBR Approach to Musical Tempo Transformations. 7th European Conference on Case-Based Reasoning ECCBR'04. Peter Funk and Pedro A. Conzález-Calero eds. Lecture Notes in Artificial Intelligence, nº 3155. Springer-Verlag, pp. 601-615 .

[5] M. Grachten, J. L. Arcos (2004), Using the Implication/Realization Model for Measuring Melodic Similarity. 16th European Conference on Artificial Intelligence ECAI'04. Ramon López de Mántaras and Lorenza Saitta eds. IOS Press, pp. 1023-1024.

[6] M. Grachten, J. L. Arcos, R. López de Mántaras (2004), Melodic Similarity: Looking for a Good Abstraction Level. 5th International Conference on Music Information Retrieval ISMIR'04. pp 210-215.

[7] M. Grachten, J. L. Arcos (2004), Music Perforrmance Generation as Time Series Prediction. ECCBR'04 Workshop on Applying CBR to Time Series Prediction. Pablo Gravás and Kaylan Moy Gupta eds. Tech.R. 142-04 Dep. Sistemas Informáticos y Programación (UCM), pp. 329-336.

[8] M. Grachten, J. L. Arcos, R. López de Mántaras (2004), Evolutionary Optimization of Music Performance Annotation. Computer Music Modeling and Retrieval CMMR'04. Lecture Notes in Computer Science. Springer-Verlag

The creation of the Start-Up MusicStrands, S.A. has motivated the collaboration with the project. We have a research contract devoted on recommendation algorithms for musical recommendation. The goal of this research was motivated from our work in the ProMusic project on similarity measures for musical sequences. We are using our research experience for developing similarity measures to be used in a collaborative filtering framework for recommending songs from user play-lists.

We have also established collaborations with academic institutions. We have visited and have visitors from both Queens University Belfast and the University of Alicante. We are currently working closely with member of these two universities on musical analysis and automatic transcription, respectively.