Chapter V

Identifying Saxophonists from Their Playing Styles

Rafael Ramirez
Pompeu Fabra University, Spain

ABSTRACT

We describe a novel approach to the task of identifying performers from their playing styles. We investigate how professional Jazz saxophonists express and communicate their view of the musical and emotional content of musical pieces and how to use this information in order to automatically identify performers. We study deviations of parameters such as pitch, timing, amplitude and timbre both at an internote level and at an intranote level. Our approach to performer identification consists of establishing a performer dependent mapping of internote features to a repertoire of inflections characterized by intranote features. We present and discuss some results of our approach and comment on future trends in this exciting research area.

INTRODUCTION

A key challenge in the area of music information, given the explosion of online music and the rapidly expanding digital music collections, is the development of efficient and reliable music search and retrieval systems. One of the main deficiencies of current music search and retrieval systems is the semantic gap between the simplicity of the content descriptors that can be currently extracted automatically and the semantic richness in music information. Conventional information retrieval has been mainly based on text, and the approaches to textual information retrieval have been transferred into music information retrieval. However, music contents and text contents are of a very different nature which very often makes textual information retrieval unsatisfactory in a musical context. It has been widely recognized that music retrieval techniques should incorporate high-level semantic music information.

In this chapter we focus on the task of identifying performers from their playing style using high-level semantic descriptors extracted from
Identifying Saxophonists from Their Playing Styles

audio recordings. The identification of performers by using the expressive content in their performances raises particularly interesting questions but has nevertheless received relatively little attention in the past. Given the capabilities of current audio analysis systems, we believe expressive-content-based performer identification is a promising research topic in music information retrieval. After presenting the background to this area and briefly discussing the limitations of this approach to performer identification, we present an algorithm for identifying Jazz saxophonists using high-level semantic information obtained from real performances. This work is based on our previous work on expressive performance modeling (Ramirez & Hazan, 2005; Ramirez, Hazan, Maestre, & Serra, 2006). Finally, we discuss the results from the case study and draw some conclusions.

The data used in our investigations are audio recordings of real performances by Jazz saxophonists. The use of audio recordings, as opposed to MIDI recordings where data analysis is simplified, poses substantial difficulties for the extraction of music performance information. However, the obvious benefits of using real audio recordings widely compensate the extra effort required for the audio analysis. We use sound analysis techniques based on spectral models (Serra & Smith, 1990) for extracting high-level symbolic features from the recordings. The spectral model analysis techniques are based on decomposing the original signal into sinusoids plus a spectral residual. From the sinusoids of a monophonic signal it is possible to extract high-level semantic information such as note pitch, onset, duration, attack and energy among other information. In particular, for characterizing expressivity in saxophone we are interested in two types of features: intranote or perceptual features representing perceptual characteristics of the performance, and internote or contextual features representing information about the music context in which expressive events occur. We use the software SMSTools (SMS) which is an ideal tool for preprocessing the signal and providing a high-level description of the audio recordings. Once the relevant high-level information is extracted we apply machine-learning techniques (Mitchell, 1997) to automatically discover regularities and expressive patterns for each performer. We use these regularities and patterns in order to identify a particular performer in a given audio recording. We discuss different machine learning techniques for detecting the performer’s expressive patterns, as well as the perspectives of using sound analysis techniques on arbitrary polyphonic audio recordings.

The rest of the chapter is organized as follows: Section 2 sets the background for the research reported here. Section 3 describes how we process the audio recordings in order to extract both perceptual and contextual information. Section 4 presents our algorithm for identifying performers from their playing styles, as well as some results. Section 5 briefly discusses future trends in the context of this research, and finally, Section 6 presents some conclusions and indicates some areas of future research.

**BACKGROUND**

Music performance plays a central role in our musical culture today. Concert attendance and recording sales often reflect people’s preferences for particular interpreters. The manipulation of sound properties such as pitch, timing, amplitude and timbre by different performers is clearly distinguishable by the listeners. Expressive music performance studies the manipulation of these sound properties in an attempt to understand why, what, how and when expression is introduced to a performance. There has been much speculation as to why performances contain expression. Hypothesis include that musical expression communicates emotions (Justin, 2001) and that it clarifies musical structure (Kendall, 1990), that is, the performer shapes the music according to her own inten-