AUDIO-BASED COMPUTATIONAL STYLOMETRY FOR ELECTRONIC MUSIC

Francisco Rodríguez-Algarra

MASTER THESIS UPF / 2014
Master in Sound and Music Computing

September, 2014

Master Thesis Supervisor:
Perfecto Herrera
Music Technology Group
Department of Information and Communication Technologies
Universitat Pompeu Fabra, Barcelona

Universitat Pompeu Fabra Barcelona

‘Noises have generally been thought of as indistinct, but this is not true’

Pierre Schaeffer

‘You can’t really imagine music without technology’

Brian Eno

‘I was never worried that synthesizers would replace musicians. First of all, you have to be musician in order to make music with a synthesizer’

Robert Moog

Abstract

Identifying artists and their stylistic signatures is a difficult problem, especially when only audio files and not symbolic sources are available. This is the most common situation when dealing with Electronic music, so the application of different constructs and techniques that have been proved useful when studying composers that wrote scores is needed. In addition to that, Electronic music increases the complexity of these problems, as timbre and rhythm tend to get more relevance than pitches, durations and chords, facets traditionally emphasized in musical style analysis.

The research presented in this dissertation aims at the exploration of the usage of Music Information Retrieval tools and techniques for the stylistic analysis of Electronic Music. For that purpose we have curately constructed a music collection specially addressed for the above-mentioned problems, containing more than 3000 tracks of 64 different Electronic Music artists. The collection has been analyzed with the help of different software libraries, and the extracted features cover different musical facets such as timbre, rhythm, and tonality aspects, and include different temporal scopes (short-term analysis windows, central tendency and dispersion measures for whole tracks, and section summaries).

The extracted features are tested in a traditional Artist Identification task, overperforming the previously reported results in similar studies when training a Support Vector Machines model and evaluating it with a holdout test. We also propose a categorization of the most relevant features for capturing the style of Electronic Music artists, distiguishing between "discriminative" and "descriptive" features. Both types are tested experimentaly, achieving satisfactory results. Finally, a detailed qualitative analysis of the results obtained when considering a small group of artists is performed, demonstrating the potential of the analyses that have been developed.

Acknowledgements

One year ago I was completely thrilled for having the chance to start the Master in Sound an Music Computing (SMC) at the Music Technology Group (MTG). And I must confess that at that time I could not even imagine how incredible this experience has become. Now that this year is coming to its end, I feel obliged to thank everybody that has made it possible.

First of all, I want to thank Xavier Serra for giving me the opportunity of joining the SMC and for having formed such an extraordinary team at the MTG. Perfecto Herrera deserves my eternal gratitude for being the best supervisor that one could imagine, always willing to help and giving constant advice. I also want to thank many other members of the MTG, especially Emilia Gómez, who has taken care of us like nobody else, Julián Urbano, for illuminating me with brilliant ideas, Sergi Jordà, for creating and conducting the Giant Steps project, Cárthach Ó Nuanáin, for helping me everytime I got stuck, and Agustín Martorell, whose advices and hard work have always been a great source of inspiration.

I don't want to forget to thank my partners of the SMC Master. They have now become part of my family, and I am pretty sure that without their vitality and sense of humor this year would have not been as incredible as it has finally been. I am specially grateful to Jaime Parra, an extraordinary traveling companion to whom I will never stop learning.

Last but not least, I could never thank enough everything that my parents have done for me, giving me complete support to follow my dreams.

Without all of them this amazing year would not have been possible. And, as Perfe once told me, I hope this will just be the end of the beginning.

Contents

List of Tables

List of Figures

Chapter 1 | Introduction

Electronic Music has come a long way since a few innovators started to generate sounds from electrical signals in the very first years of the 20th Century. What once was a purely experimental task, in which the line between art and science was extremely thin, has derived into a mainstream trend that can be found in top positions in most charts all over the world.

The exponential advance that in the last decades has experimented the technology that allows us to create, distribute and listen to music, and its much widespread availability, explains a large part of this explosion. In parallel, this evolution of the technology has also allowed the appearance of a young academic field that uses computational tools to analyze music. This discipline is usually called "Music Information Retrieval", or, more recently, "Music Information Research", and has developed a series of tools and techniques that have a great potential for supporting musicological studies. The particularities of Electronic Music leads us to think that those tools and techniques may be especially useful for this kind of music, as it relies in musical facets not usually analyzed with traditional musicological methods and that can be captured with the help of those tools. For that reason, we consider this dissertation as an exploratory study of the potential of Music Information Retrieval for the stylistic analysis of Electronic Music.

1.1 Motivation

The main motivation for the development of this Master Thesis has been trying to answer the following question: *"Are current State-of-the-Art Music Information Retrieval tools and techniques able to capture Electronic Music artists' own creation style?"*. In fact, as it is not that clear that there exists individual style when dealing with Electronic Music, trying to determine if there are particular aspects of the music that can be related with its creator is one of the main purposes of this work.

The research presented here is developed in the context of Music Information Retrieval (MIR), discipline that will provide us the methodological framework for addressing our task. However, it is very likely that traditional MIR procedures do not perfectly fit with the goals that we seek. In that case, a refinement of the methodology usually developed should be proposed.

We hope the results obtained while developing this Thesis will allow us to improve our

understanding of Electronic Music, as well as being useful for tasks such as the creation of automatic expert assistants in music production softwares. What we have no doubt about is that it will be a challenge and an oportunity to test the capabilities of MIR tools and techniques in a difficult scenario.

1.2 Goals

The main goals that will be addressed during the develpment of this research are the following:

- Create a reliable music collection for Electronic Music style analysis
- Explore the potential of MIR tools and techniques for the stylistic analysis of Electronic Music
- Determine which audio features best capture the creative style of some EM artists
- Test the ability of those audio features to discriminate between artists
- Analyze the relationship between the selected audio features and the perceived stylistic traits

1.3 Structure of this Dissertation

This dissertation is structured as follows. Chapter 2 is devoted to the revision of the most important concepts that should be clarified before starting the investigation, such as Electronic Music and Style, as well as a review of the MIR tasks that are more closely related with our own research. Chapter 3 is focused on the main methodological steps needed to achieve our objectives: construction of the music collection, extraction and analysis of the audio features, and the development of particular experiments. In Chapter 4 we explore in detail the results obtained for a particular group of artists and try to provide a more music-related interpretation of them, and, finally, Chapter 5 provides an overall discussion of the work performed, a summary of the achievements of this Thesis and some suggested future research paths that have been opened during the development of our investigation but we have not had the chance to follow.

Chapter 2 | Background

Before we can start to 'analyze style in Electronic Music' we should ensure that "Electronic Music" and "style" mean the same for all of us. For that reason, the main goal of this chapter is to clarify some essential concepts necessary to understand the scope of this work. First, in Section 2.1 we establish a delimitation to the type of music that will be considered as Electronic Music for our experiments and review the previous works in Music Information Retrieval (MIR) that address similar types of music. Afterwards, in Section 2.2 we attempt to shed some light on the rather obscure concept of "style", and we will review how is it usually considered in the arts, with a particular emphasis in music, as well as what is stylometry an how it is related to our problem. Finally, a quick review of the MIR literature regarding Style Modeling and Artist Identification in the context of Automatic Music Classification is presented.

2.1 Electronic Music

In this section we will try to establish a definition of what will be understood as Electronic Music throughout this dissertation. In addition to that, a brief review of the few works in which a computational analysis of such music has been addressed is also provided.

2.1.1 Electronic Music Delimitation

The boundaries of what should be considered as Electronic Music and what should not are fuzzy. For example, should a reinterpretation of a Baroque song with electronic instruments, such as Walter Carlos' (later Wendy Carlos) famous album "Switched on Bach" be included in the Electronic Music definition? What if a piece, or even the whole discography of an artist, is based solely in acoustic music sampling? And nowadays this distinction is getting even more difficult with the generalization of collaborations between renowned Electronic Music producers and mainstream pop artists. Those considerations are probably useless for the general public, but they are fundamental in works like this one in order to establish the scope of the results obtained.

In search of a formal definition for Electronic Music, the first source which we should check is the New Grove Dictionary of Music. Surprisingly enough, there is no specific entry for Electronic Music, as we are redirected to the article regarding "Electro-Acoustic Music".

While maybe too broad (probably many of nowadays mainstream pop music tracks would fit in that definition), Emmerson and Smalley [19] provide a quite satisfactory definition:

> 'Music in which electronic technology, now primarily computer-based, is used to access, generate, explore and configure sound materials, and in which loudspeakers are the prime medium of transmission'

However, a more in depth review of the entry reveals that the scope of the music that they are considering is limited to experimental/artistic electronic music, from *Elektronische Musik* and *Musique concrète* to the more recent term of *Sonic Art*. This is, in fact, the meaning that is usually given to Electro-Acoustic Music. Therefore, the problem is that they are not considering "vernacular" – in the sense of non-cult or non-artistic – Electronic Music as belonging to the defined concept, even if it would fit into the literal definition provided. They also refer to Electronic Dance Music (or simply Dance Music or EDM) as a vernacular musical genre related to Electro-Acoustic music, but its definition ('electronic music intended primarily for dancing at nightclubs and raves' [14]) explicitly excludes any piece not suitable for dancing. We find, then, a gap in the definitions that makes them not suitable enough for our purposes.

Nevertheless, two interesting points emerge from the review of EDM's entry, and, as EDM should be included in our particular definition of Electronic Music, those aspects may be easily adopted. Firstly, sampling is not only permited but also considered as one of the main treats of many EDM songs. Therefore, our definition should explicitly allow the usage of sampled recordings. Additionally, it is characterized by its rapid evolution, a fact that its highlighted by the constant creation of new subgenres and hybrid genres. This phenomena is deeply studied by McLeod [41], who finds more than 300 subgenre labels when considering CD compilations and specialized magazines only during 1998 and 1999. He claims that this cannot be attributed only to the evolution of music itself, but also, and sometimes even more notably, to cultural and economic factors. This point is highly relevant for our research, as it implies that a "ground-truth" based on Electronic Music subgenres would be even less trustful than usually. As it will be detailed in Section 2.3, Artist Identification seems to be the only Music Classification task whose "ground-truth" is stable enough for our purposes.

McLeod, apart from analyzing the reasons that cause the creation of that miriad of subgenres, provides a definition that may be of interest for us:

> 'Electronic/dance music is an umbrella term used in this article to label a heterogeneous group of musics made with computers and electronic instruments, often for the purpose of dancing. Electronic/dance is a metagenre name that is vague enough to describe the broad variety of musical styles consumed by a loosely connected network of producers and consumers. A slash is used (rather than a hyphen) as an and/or designation because not all the musics consumed by these communities are necessarily designed for dancing'

As it can be seen, McLeod explicitly states the heterogeneousity of the music covered by his definition and also allows non-danceable songs to be included. We could take profit of this and combine it with the previous ones to create our own definition of Electronic Music. In this way, for the purpose of this dissertation, "Electronic Music" is considered to be *a metagenre that includes a heterogeneous group of popular genres and subgenres in which electronic technology, now primarily computer-based, is used to access, generate, explore and configure sound materials, both synthesized and sampled, addressed to its public for listening and/or dancing.* We exclude from our scope Electronic and Electro-Acoustic Music generally associated with experimentation, such as the *"avant-garde"* movement of the 20s to the 60s of the 20th century. This implies that artists such as Karlheinz Stockhausen or Iannis Xenakis won't be covered in this dissertation.

2.1.2 Computational Analysis of Electronic Music

If we review the literature related to Electronic Music published by the Sound and Music Computing community, we realize that its creation has received much more attention than its analysis. As far as we know, only Nick Collins' submissions to the 2010 and 2012 International Society for Music Information Retrieval Conferences (ISMIR 2010 and ISMIR 2012) directly deal with the usage of computational tools for analyzing mainly popular Electronic Music. In fact, in the former [8], the genre considered (Synth Pop) would not be included in our delimitation of Electronic Music. He uses a number of web resources to construct a graph of influences around Depeche Mode. He then computes timbral similarity between audio excerpts from Depeche Mode and some of the artists and bands that influenced or were influenced by them. Finally he uses the extracted features to train a Support Vector Machines (SVM) classifier in order to determine the album in which a track was included, obtaining a 31.8 % of accuracy, the group of albums among a set of 10 groups in which the album of the track was included, reaching an accuracy of 77 %, and the year of release, obtaining a 55 % of accuracy.

Again, Collins [9] attempts to determine the main influences in the creation of two of the main genres of EDM, say Chicago House and Detroit Techno. He constructs a dataset comprising tracks of both genres, as well as six other genres either simultaneous or earlier to the first years of Chicago House and Detroit Techno. He computes a set of timbral features, as well as the inter-onset intervals and beatwise chroma differences, for each track, and obtains the predictive capacity of one genre class with respect to the rest. The results obtained are not what he expected, as Synth Pop and not Disco appears as the closest genre to Chicago House, and Punk and not Funk, the closest to Detroit Techno.

Apart from those publications, we are only aware of a few Master's Theses that deal with the computational analysis of Electronic Music. Kirss [33] addresses the classification of electronic music excerpts in a five-genre taxonomy (Ambient, Deep House, Uplifting Trance, Techno and Drum-and-Bass). He constructs an in-house dataset of 250 songs and extracts a number of timbric and rhythmic audio features from different toolboxes. He

then uses those features to train several different Machine Learning algorithms. He checks various combinations of features and algorithms, and determines that the best result is obtained when using timbric and rhythmic features from Rythm Patterns[1] and Marsyas[2] toolboxes to train a SVM. The accuracy reached with this combination is above 96 % in 10-fold cross-validation. Sesmero-Molina [57] continues this same work and tests other sets of features, training the system with the songs collected by Kirss. In this case he reaches a maximum of around 90 % of accuracy in a hold-out test with another 250-song collection when using a set of features developed by the MTG.

Finally, Melidis [43] addresses the problem of Electronic Music Artist Identification. However, as his work is highly relevant for us, we will review it in detail later in Section 2.3.5.

2.2 Style

Style is a quite vague term that may be understood in different ways depending on the context in which it is employed, so an explanation of what style means for the scope of this dissertation is needed. In this section we try to clarify the meaning that will be given to this concept in the present dissertation.

2.2.1 Defining Style

Most of us would have little difficulties in distinguishing a Van Gogh's painting from a Picasso's. Or a Gothic cathedral from a Neo-Classic one. Or a Kafka novel from a Dostoievki's one. Or, focusing on music, a Liszt piano composition from a Chopin's one. And what allows us to achieve those distinctions is what we usually call *Style*. Not only style is present in the arts, but also in most of our daily activities. Therefore, we can distinguish between different accents of a spoken language, or even different manners to drive a car. However, even though we can easily recognize different styles when we perceive them, and determine which styles do we like or not, a general understanding of how style works is a much more difficult task.

Verdonk, in his entry about style in the Encyclopedia of Language and Linguistics [64] recognizes the difficulties to provide a formal definition as it is an abstract concept. Elkins, in the same line, considers in his entry in the Grove Art Online [18] that "the term style is one of the most difficult concepts in the lexicon of art, and one of the chief areas of debate in aesthetics and art history". One of the clearest demonstrations of those difficulties is that the Oxford English Dictionary[3] finds twenty-seven variants for the term "style". Moreover, most of those variants include definitions with at least two competing senses, in the same way that Chatman [7] commented about the definitions suggested by Murry in his series of

[1] `http://www.ifs.tuwien.ac.at/mir/audiofeatureextraction.html`

[2] `http://marsyas.info/`

[3] *"style, n."* OED Online. Oxford University Press., accessed March 22, 2014, `http://www.oed.com/view/Entry/192315?rskey=LOS91a&result=1&isAdvanced=false`.

lectures entitled "The problem of style" [46]. Some of those definitions treat style as mere description, in the sense that we can express that something is done in a recognizeable way without needing to judge whether that way is good or bad. Therefore, if style is understood in that sense, the main goal of stylistic analysis is to determine what is that makes the manner in which a specific work is done different from others'. In other cases the definitions are undeniably bounded to value judgements, where "having style" is closely related to "being good" in a specific task. Understood in this way, the goal of a stylistic study would be to obtain the features that make a work "better" or "worst" than others.

We agree with Chatman when he argues that a "study of style" should avoid judgements of value, so, for the scope of this work, we will treat style as a neutral concept. In this way, we could define style as *a particular manner or technique by which something is done, created or performed*[4]. However, this definition does not tell us anything about what makes it being *particular*, and this has caused that lots of scholars from a number of different disciplines have attempted to improve the understanding of the nature of style.

Recently, Martin Siefkes [59] has approached the conceptualization of style from a semiotic point of view. This has allowed him to propose what he calls a "General Theory of Style" that is independent of the area of study. Therefore, not only artistic style is addressed, but also any other human activity can be described in terms of style, such as "style of running" or "style of driving".

2.2.2 Style in the Arts

Even though the English spelling of the word "style" has its origins in a confusion with the Greek term "stulos", a type of column, etymologically it derives from the Latin word "stilus", which referred to an ancient writing instrument made of metal, wood, or bone, with a a sharp-pointed end for scratching letters on a wax tablet and a blunt end for erasing them. The records that we have show that it was mainly used in the process of learning caligraphy. In fact, some of the definitions of style that we can find in the Oxford English Dictionary still refer to that meaning. It is very likely that this meaning was successively extended from "an instrument for writing" to "a manner of writing", firstly in a literal sense until it finally got understood as "a particular way of expressing". The first usages of this particular interpretation of the term associated style almost exclusively with rhetoric, but soon gained widespread currency outside the domain of language, particularly in art history and criticism.

Having its origins in the language-related disciplines, style has been traditionally associated with *how* something is said as opposed to *what* it is said, this is, its subject. However, as Goodman [23] lucidly points out, this distinction has little sense in disciplines such as architecture and nonobjective painting, as well as in most of the music pieces, as they have no subject at all. In those cases, their style cannot be a matter of how they say something, because they don't say anything. So the concept of "synonymy" as the usual

[4] *"style, n."* Merriam-Webster.com, accessed March 22, 2014, `http://www.merriam-webster.com/dictionary/style`.

way of determining style, this is, by trying to determine how *exactly* the same meaning can be expressed in different ways, is nonsense in a wide range of disciplines. But even if we consider that they mean in a non-conventional sense, the real relevancy of style appears when very different things are expressed in a very similar way.

One of the principal traits of the usage of the word "style" in art-related activities is that it can be referring to styles of people but, and sometimes more frequently, also to styles of periods. Styles of people are often called "individual styles", while the style of a period is referred to as a "general style", which, at a turn, may be divided in universal style (classicism, naturalism), historical or period style (Art Nouveau, Neo-Classicism) and school style [18]. For example, when we talk about styles of painting, as, for instance, cubism or surrealism, we are implying that we can identify a set of characteristics common to the work of an entire community of artists, what we call a school or a movement. On the other hand, when we talk about an artist's style we are referring to those characteristics that make unique the work of the artist as an individual even while it shares all the common characteristics with the rest of the artist of the school in which he or she can be englobed.

A key point when analyzing artistic styles is understanding how a new style is built, be a period or an individual style. A traditional approach considers that styles are built by means of a process of choice, either concious or unconcious, between all the possible alternatives that he or she faces in the creative process [64]. James S. Ackerman, in his essay "A Theory of Style" [1], claims that artists, by choosing to accept or to alter certain features of the art around them, establish the fundamental relationships that form their individual styles and, at the same time, contribute to the continuous evolution of the art movements. He also defends that those choices are, in fact, needs caused by the artists' fight against the society's pressure for stability. He states that style evolution and consolidation is inevitable due to the fact that, unlike machines, we are not able to reproduce without inventing and, at the same time, we cannot invent without reproducing. Recently, this same idea has been adopted by the Flow Machines project [49], whose main purpose is to provide computational tools for enhancing creativity by means of modeling the style of existing artists and allowing the user to manipulate and experiment with those models. They argue that this would improve the acquisition of skills that are required in order to develop a personal style.

Another key issue in the study of art style is determining which are the features of a specific work can be considered stylistic, in the sense that they are characteristic of the particuar style of an artist, school or period. According to Goodman, not all differences in ways of writing or painting or composing or performing are differences in style. If a work is in a given style, only certain among all the aspects of the subject, form, and feeling of the work are elements of that style. Ackerman defends that the characteristics that comprise a certain style should be at the same time stable, meaning that they have to be suitable to appear in other products of the same artist(s), era or, even, region, and flexible enough to be able to change between works belonging to different styles. In general, stylistic properties help to determine the who, when and where of a work, even if they need to be

combined in order to be relevant. However, Goodman points out that even if a property helps to determine the maker or period or provenance of a work, this is not enough to be considered as stylistic. He provides a series of examples, such as the label on a picture or the chemical properties of its pigments, which help to its identification but cannot be considered stylistic at all because, in his opinion, they do not contribute to the symbolic functioning of a work as such.

Determining which features should be studied in order to analyze the stylistic traits of a work is a challenging task. Goodman provides a non-exhaustive list of problems that may be faced:

- A property common to many works may be an element of style for some but stylistically irrelevant for others
- Some properties may be only usual rather than constant features of a given style
- Some may be stylistically significant not through appearing always or even often in works of a given author or period but through appearing never or almost never in other works.

Therefore, no fixed catalogue of the elementary properties of style can be compiled, at least not in general. In order to increase our chances to find the most suitable set of stylistic features we have to take a closer look to the specific discipline in which we are involved.

2.2.3 Style in Music

Unlike the majority of other artistic disciplines, music is usually considered as not being able to generate creative works directly from the concrete reality [34]. So style is slightly different in music than in literature, painting or sculpture, for example. According to Dannenberg [12] if we consider a short melody without words, for instance, we would not find any obvious objective meaning, story, or referent associated with it. Everything that we enjoy (or not) about the melody has to do with expectations, sound quality, performance nuance, and musical texture. Essentially every aspect of the melody that communicates something to the listener is an aspect of style. One could say that style is everything in music. Or, seen the other way around, everything in music is style.

Nevertheless, even if every aspect in a piece of music has the potential to contribute to its style, the term itself is so broad and vague that a deconstruction into its buiding blocks is needed. In that sense, Pascall [50] reviews the facets of music more widely studied with regard to style:

- *Form*: Even though some musicologists have considered form and style as opposed, in music form itself has a central role in style creation by joining together all the details from which a piece is composed
- *Texture*: The particular selection of voices that provide a certain sonority to a piece is undeniably a trait of its style, as it would be the color selection in a painting

- *Harmony*: Harmony has a strong impact in the perceived tone of a musical piece, affecting specially to the emotions that it conveys
- *Melody*: Being the central part of most music, at least in the Western tradition, it is not rare that melody keeps a huge part of the stylistic content of a song
- *Rhythm*: Tempo, metre and expressive rhythm changes are, among others, some of the key aspects that allow us to distinguish different musical styles

A very relevant issue in the study of musical style is the debate around if it is innate or, on the contrary, a culturally imposed convention. At this respect, Leonard B. Meyer, in his book "Style and Music: Theory, History and Ideology" [44] offers a hierarchical organization of the constraints that musicians face when performing the choices that determine a particular style:

1. Laws, or "transcultural physical or physiological constraints".
2. Rules, or "intracultural constraints".
3. Strategies, which represent the choices made within established rules. He distinguishes three types:
    - Dialects, in the sense that geographical neighbors or contemporaries will share similar strategies;
    - Idioms, that represent the personal voices of the authors;
    - Intra-opus styles, which are related with the changes caused by exposition to different styles.

Therefore, according to Meyer, style is partly imposed, not only culturally but also due to some human-wide constraints, and partly innate, represented as what he calls "idioms".

The usage of style as a means of describing music has been usual since the Classical period, when specific styles, such as "singing", "brilliant", "strict" or "learned", were employed for identifying pieces with particular rhythmic and textural features [53]. However, one of the main problems of the usage of the term "style" in music is that it is employed in many ways, even in a wider range than we saw earlier for arts in general. Dannenberg [12] lists the following possible meanings:

- *Historical periods* of music: Baroque and Classical are examples of styles defined by their historical periods. However, in some contexts a wide range of periods is known as "Classical", while in others a much more precise distinction is done (as, for example "Late Classical")
- Styles are associated with the set of characteristics generally found in the works by *composers*
- *Performers*, especially improvising performers, also have their own style
- Sometimes style is employed to describe the *texture* of a song, in the sense of "the way it sounds". Dannenberg speaks of a "tonal style", a "heavy style" or a "big band style" as examples of this meaning

- *Emotions* conveyed by music are also considered sometimes as styles. For instance, it is not rare that the same song can be interpreted in a "happy style" or a "sad style"
- Finally, style is often used to mean *genre*

Surprisingly enough, even though he includes textural styles in the list, he does not consider *formal styles* (sonata style, fugue style), *harmonic styles* (modal style, atonal style), *melodic styles* (motivic style, ornamental style) and *rhythmic styles* (ternary style, free style) [50]. Other authors also include a geographical dimension to musical style [37] that Dannenberg seems to be ignoring by focusing his attention into Western music.

The last of the possible meanings considered by Dannenberg, the equivalence between "style" and "genre", has been source of discussion. Moore [45] attempts to determine whether those terms are interchangeable, subordinated one to the other (be "genre" a sub-level of "style" or the opposite) or referring to independent areas, even if they both are concerned with ways of finding similarities or differences between pieces. He reviews several works in which one of those three possibilities is implied, making it clear that no single understanding is established. However, the fact that it seems to be an inclination towards the usage of "style" in musicology and "genre" in cultural studies and media, encourages Moore to think that both terms should not be considered as equivalents. In this sense, he finds at least four main differences between them:

1. Style refers to the manner of articulation of musical gestures and is best considered as imposed on them, rather than intrinsic to them. Genre refers to the identity and the context of those gestures. This distinction may be characterized in terms of "what" an art work is set out to do (genre) and "how" it is actualized (style).
2. Genre, in its emphasis on the context of gestures, pertains most usefully to the esthetic, while style, in its emphasis on their manner of articulation, pertains most usefully to the poetic.
3. In its concentration on how meaning is constitued, genre is normally explicitly thematized as socially constrained. Style, on the other hand, in its emphasis on technical features and appropiability, frequently brackets out the social or at least regards this realm as minimally determining, where it is considered to operate with a negotiable degree of autonomy.
4. Style operates at various hierarchical levels, from the global to the most local. Genre as a system also operates hierarchically, but with the distinction that "sub-genres" cover an entire genre territory that "sub-styles" do not.

From now on we will treat those terms as distinct, focusing on style (individual or groupal) rather than genre, even though a significant amount of literature employs the former when it is really referring to the latter.

2.2.4 Stylometry

Under the term "Stylometry" we are referring to the discipline that performs statistical analysis of styles. It relies on the assumption that style is mainly a probabilistic concept, in the sense that it is comprised by a series of features that can be described my means of their probability distributions. However, this assumption has been widely critisized by some of the most influential style analysts, such as Meyer Schapiro, who believed that with solely mathematical tools it was not feasible to describe the vague concepts of style [56].

Nevertheless, stylometry seeks to complement the traditional style analysis techniques by providing an alternative means of investigating works of doubtful provenance. At its heart lies an assumption that authors have an unconcious aspect to their style, an aspect which cannot be consciously manipulated but which possesses features that are quantifiable and may be distinctive. In order to be useful to perform stylometric studies, the analyzed features should be salient, structural, frequent and easily quantifiable, and relatively immune to conscious control. By measuring and counting these features, stylometrists hope to uncover the "characteristics" of an author. The two primary applications of stylometry are attributional studies and chronological problems.

Wincenty Lutoslawski [38] is considered to be the father of stylometry as a discipline during the last decade of the 19th century with the development of methods that allowed him to establish a chronology of Plato's Dialogues. However, the basic ideas of employing statistics to determine stylistic traits had been around from, at least, forty years then. Since that moment, linguistics have always been the discipline in which stylometry has been more employed. For a comprehensive review of the tools and the history of linguistics stylometry we refer the reader to the publications by David I. Holmes [26, 27].

One of the most influential stylometrists of all times, George K. Zipf, developed the statistical theory known by "Zipf's Law" [68, 69] which states that some words are used very often, while others are used rarely, following what is known as a 1/f distribution. This claim has gained interest in other areas appart from linguistics, as it seems that not only in language but also in music, for example, some features appear following this probabilistic distribution [24, 35].

Lately, stylometry has been largely influenced by computer science and artificial intelligence, so that it can be considered a problem of pattern recognition. Machine learning techniques, such as Neural Networks or Genetic Algorithms, have demonstrated to be particularly useful for this kind of tasks as they are able to recognize the underlying organization of data.

The tools and tecniques conceived for linguistic stylometry were adapted for their use in musical style analysis starting from the late fifties of the 20th century [47, 21], due to the fact that the symbolic representation of music has a strong structural resemblance with literary texts. Nowadays the usage of the computational tools for musical style modeling and recognition from symbolic sources is a common practice, as we will review in Section 2.3.2. Other disciplines, however, such as painting, cannot rely on those tools and are

forced to use sophisticated image processing techniques in order to identify stylistic traits in pieces of art [29, 28]. Similarly, audio-based identification of musical artist relies on signal processing to extract features suitable of being characteristic of the particular artist into consideration. Section 2.3.3 will be dedicated to this topic.

2.3 Automatic Music Artist Identification

In this section we will introduce the main characteristics of Automatic Music Classification, a field within Music Information Retrieval (MIR) that will provide us the bedrocks over which our research will be built. More precisely, we will review the specific tasks that are more closely related with stylistic studies of music, either from a symbolic representation or from the audio signal itself. Finally, a more in depth review of the work previously developed by Melidis is shown, as it is highly relevant for our own research.

2.3.1 Automatic Music Classification

The Automatic Classification of Music is one of the most active areas within the MIR community. Its main purpose is the usage of Machine Learning techniques for the identification of the class in which a particular music excerpt should included in, employing a series of descriptors extracted either from the audio signal or from a symbolic representation of it, such as a score or a MIDI transcription. In the following sections we will first describe the general procedure that most classifiers comply (Sec. 2.3.1.1) and afterwards a brief review of the most usual tasks in Music Classification will be performed (Sec. 2.3.1.2).

2.3.1.1 Framework

Before starting to implement any classifier it is fundamental to define carefully the classes that will be considered in the problem addressed. The taxonomy of classes that we organize should ideally be consistent, avoid overlapping of classes and complete [25]. Once the classes have been clearly defined, the next step is to construct a collection of music material that is representative of the class taxonomy considered. In order to avoid the introduction of undesirable biases in the classification results, some cautions should be taken, as, for instance, the inclusion in the dataset of the same number of observations for each class. Particular tasks may also need specific cautions. Artist Identification, for example, is often required to try to avoid “album effect” issues [32] by not using songs from the same album both in the training and the test sets.

When a sufficiently big dataset has been collected, the next step is to extract a set of features that capture specific characteristics of the music. This step is the more challenging of all, as determining which descriptors should be computed in order to separate observations belonging to different classes is not a trivial task at all. The number of available descriptors in the literature is incresingly high, and cover from the lowest level of abstraction (raw numbers extracted directly from the source, easily interpretable by a machine

but meaningless for humans) to the highest level of abstraction (semantic descriptions easily understandable by humans but extremely complex for machines). They also represent most of the facets of music: Rhythm, Pitch, Timbre, Melody, Harmony, etc. The successfulness of the classification procedure strongly depends on the features chosen, so one could argue that it is one of the most crucial parts in the process chain. However, the increasing computational capacity of nowadays machines has allowed many researchers to compute as many descriptors as possible and focus on improving other aspects of the system, such as the classifier selection, instead of making efforts in constructing better features for getting closer to the way we humans perceive music. A usual step for avoiding having to deal with a large number of features is performing some dimensionality reduction technique, such as feature selection or Principal Components Analyisis (PCA).

The way in which we take profit of the extracted features depends on the learning approach that we use. The most common one is *supervised* learning, meaning that some of the observations are used to train a machine learning algorithm which will construct a model that will allow it to perform some predictions regarding the class to which another observation belongs. It is called supervised learning because the observations in the training set include a label indicating their class, so the algorithm can infere some rules that relate a certain structure of the features with a specific class. Typical machine learning algorithms for supervised learning are NaiveBayes, k-Nearest Neighbors (k-NN), Decision Trees, Gaussian Mixture Models (GMMs), Neural Networks and Support Vector Machines (SVMs), among many more. The output of the system will be, therefore, a label indicating the class in which the new observation is more likely to belong according to the computed model. The other possible approach is *unsupervised* learning, in which no observation has a label assigned beforehand and it is the algorithm by itself who finds structures in the features of the observations and groups them together forming "classes". Strictly speaking, in this case we should not talk about a Classification problem but a Clustering one.

The final step is to determine the performance of the system. The most usual way is to compute some Figures of Merit (FoMs), such as Precision, Recall, Accuracy, F-Measure and so on, obtained from the comparison between the predicted class and a "ground-truth". Additionally, it is common to compare those FoMs with a State-of-the-Art score to determine how better or worse our system is with respect to other works addressing the same problem. However, recently, some critic voices have appeared in the MIR community arguing for a more formalized evaluation [63] or even a new approach that focuses more on the musical meaningfulness of the individual classifications than the determination of a numerical overall measure of "goodness" [61].

2.3.1.2 Tasks

According to Fu et al. [20] the main tasks in Music Classification are Genre Classification, Mood Classification, Artist Identification, Instrument Recognition and Music Annotation. Among all of them, the one that has received more attention from the community (while

being the most conflictive one) is undeniably Genre Classification. In this sense, Aucouturier and Pachet [3] show that genre is an ill-defined term, that is not founded in any intrinsic property of the music, but rather depends on cultural extrinsic habits. Nevertheless, it has demonstrated to be one of the most useful concepts to discriminate music. People usually describe music in terms of its genre, so an automatic system for identifying similarities in genre is potentially very useful.

The terms "style" and "genre", as we already mentioned in Section 2.2.3, are often treated indistinctively for referring to a group of songs or artists that share common characteristics. This could lead us to think that Genre Classification would be a good framework for testing potentially stylistic features. However, as it has been repeated a few times by now, genre is mainly culturally defined and not explicitly embedded into the music itself, so it does not provide a sufficiently solid ground-truth for our purposes. Artist Identification, on the other hand, contains an explicit ground-truth that is very rarely missing. While this could be an argument against the development of such systems – if we already know the artist of the song, why would we need to predict it? – we could take advantage of its implications and use Artist Identification as our test bed for stylistic features. In this sense, from now on we would focus on reviewing more in depth the strategies developed by the MIR community to address Artist Identification and related problems.

2.3.2 Musical Style Modeling and Recognition from Symbolic Sources

As we already mentioned in Section 2.2.4, symbolic representations of music are closely related with literary texts, so musical style analysis can take profit of similar approaches to those developed for linguistic style. In that sense, Cope [10], in one of the first works that aims at using computational tools for capturing musical style, uses a grammatical-generation system combined with what he calls *signatures* for reproducing the style of past composers. Those signatures are series of 2 to 9 notes that distinguish a particular composer but without making a certain piece unique. Nevertheless, Cope does not perform any test, such as artist identification based on the obtained signatures, in order to determine if the results are meaningful. Other works, such as [11, 51] also explore the usefulness of adapting typical strategies for linguistic style measurement to the musical analysis, but generally considering period style instead of individual style, which will be our focus of interest in this literature review.

Similarly to the goal addressed by Cope, Dubnov et al. [17] implement a system for generating new pieces according to a certain style. They compare two different methods (Incremental Parsing, based on the compression techniques of the Lempel-Ziv family, and Prediction Suffix Tree, a Probabilistic Finite Automata that uses variant order Markov chains) to perform an *unsupervised* analysis of musical sequences, capturing the relationships between rhythm, melody, harmony and polyphony that lead to a piece's particular style. The models obtained with the analysis are then used for generating new interpretations and improvisations based on the original piece. In the same way as Cope, they do not

perform any test to determine the meaningfulness of their results. However, unsupervised learning is a strategy not very frequently used in this kind of analyses that may be worth considering. A Markovian approach, but with some tuning in the management of rhythm, beat, harmony and imprecision for ensuring musicality, is also used by Pachet [48]. He implements a real-time interactive system capable of modeling the improvisational style of a performer and generating a continuation for the sequence that it had been receiving. In this case the author reports having performed tests with listeners, and claims that "in most of the cases, if not all, the music produced is undistinguishable from the user's input" but no support data is provided to confirm this affirmation.

Improvisational style analysis was also addressed in one of the first explorations of the usage of supervised machine learning techniques for style recognition. Dannenberg, Thom and Watson [13] use MIDI trumpet recordings generated while the performer follows the instructions shown at a screen in which the name of a style category is written. They then identify 13 low-level features based on the MIDI data and use them to train three different classifiers (Bayesian, Linear and a Neural Network). With a 4-class taxonomy they reach accuracies in the style recognition above 98 %, while the results obtained when increasing the number of categories to 8 ranges from 77 % with the Neural Network to 90 % with the Bayesian classifier. They then conclude that machine learning techniques are a valuable resource for style recognition.

Classical (in the broadest possible sense) composer identification is the most relevant task in which symbolic-based stylistic features have been used. And of those, contrapuntal features are known to be very useful. Backer and Kranenburg [4] measure 20 low-level counterpoint features, which they call "style markers", from over 300 scores of compositions by J. S. Bach, Händel, Telemann, Mozart and Haydn. Using simple classifiers, such as k-Nearest Neighbors, Bayesian and Decision Trees, they are able to predict with very little error the composer when dealing with a limited number of possible authors. They use their results to provide experimental evidence that a piece of music traditionally attributed to J. S. Bach is likely to be authored by J. L. Krebs. Mearns, Tidhar and Dixon [42] create a set of high-level counterpoint features that they claim are more musicologically valid than those usually extracted. They use a small dataset consisting of 66 pieces by seven composers from the Late Renaissance and Baroque periods (J. S. Bach, Buxtehude, Ruggero, Vivaldi, Monteverdi, Corelli and Frescobaldi) in Kern data format. A Naive Bayes classifier and Decision Trees are trained with the features extracted from the dataset and they obtain a classification accuracy of 66 %.

Apart form counterpoint features, other descriptors have also been considered for classical composer identification. Kaliakatsos-Papakostas, Epitopakis and Vrahatis [30] examine the potential of the Dodecaphonic Trace Vector (a Pitch Chroma Profile) for composer identification with a dataset consisting of 350 MIDI transcripts of pieces by 7 artists (J. S. Bach, Beethoven, Brahms, Chopin, Händel, Haydn and Mozart). They first use Probabilistic Neural Networks to construct a similarity matrix between the composers, and then a Feedforward Neural Network is trained to identify the author of a piece from two

possibilities. They suggest that their results might be used to construct a influence diagram between composers. Pitch Class features are also used for composer identification by Dor and Reich [15], as well as octave, note count, pitch range and pitch trigram features. In addition to that, they employ a feature discovery tool for obtaining new features not considered initially, such as note duration and pitch gradient features. They construct a collection of 1183 Kern-formatted scores by nine composers (J. S. Bach, Beethoven, Chopin, Corelli, Haydn, Joplin, Mozart, Scarlatti and Vivaldi) which they combine in different groups. Several machine learning algorithms are trained in Weka with the features extracted from those multi-composer data sets, and they obtain classification accuracies raging from around 60 % to over 95 %. The group with worst results is the one formed by string and keyboard pieces by Beethoven, Haydn and Mozart, while the one leading to the best accuracies is formed solely by keyboard pieces by Bach, Corelli and Scarlatti. Simple Logistic is revealed as the algorithm with higher performance in this task. Two-composer classifications are also explored, with most of the results over 90 % of precision.

Composer identification has also been used as a means of demonstrating that the extraction of metrics based on the Zipf's Law may be able to capture relevant aspects of music. Manaris et al. [39] propose 20 Zipf-based metrics that measure the proportion or distribution of various parameters in music, such as pitch, duration, melodic intervals and harmonic consonance, as well as 20 fractal metrics that represent the self-similarity of the distribution of the 20 simple metrics. They use the extracted features for identifying composers in five different collections of MIDI-encoded performances using various architectures of Neural Networks. In a five-composer taxonomy, formed by compositions by Scarlatti, Purcell, J. S. Bach, Chopin and Debussy, they reach an average accuracy of 94 % with a hold out test. They conclude that this high accuracy suggests that Zipf-based metrics are useful for composer identification.

Finally, it is worth noting that, as we mentioned earlier, only "individual style" recognition has been reviewed here. Other works also address style recognition, but in a sense that is more closely related with genre identification than what we consider in this dissertation.

2.3.3 Audio-Based Music Artist Identification

The identification of music artists different than the composer of a classical piece usually requires an approach based on the extraction of features from the audio signal of the song, as frequently timbral descriptors are the most relevant ones for distinguishing among artists. One of the first attempts to identify artists on a song collection was performed by Whitman, Flake and Lawrence [65], who use 8-band MFCCs as the timbral representation of the excerpt and train a Neural Network, Support Vector Machines and a combination of both. In this way they are able to correctly predict the artist of a song with a 91 % accuracy in a small dataset of 5 artists, but the results drop to 70 % accuracy when considering 10 artists.

Singer identification is the subtask of audio-based artist identification that has been

most widely addressed by the MIR community. We can distinguish two main situations in which singer identification can be performed: monophonic and polyphonic excerpts. In the first case, Bartsch and Wakefield [5], report a 95 % of accuracy when identifying a classically trained singer among 12 different possibilities by using an estimate of the spectral envelope called the Composite Transfer Function. The accuracy drops to around 70-80 % when using actual recordings of Italian arias. More recently, Tsai and Lee [62] address this task by addapting GMMs constructed from MFCC features extracted from speech excerpts with a few solo singing data for each singer. They justify the usefulness of this approach by noting the difficulty of finding a cappella samples from popular singers. They test the system in a 20-singer dataset of Mandarin pop song passages and obtain a 95 % of accuracy when 15 singing recordings per singer are used in the adaptation step, while the performance drops below 70 % of accuracy when only 5 samples are used.

Having said that, the polyphonic case has recieved more attention as it is a much more useful approach for real-life applications. A crucial point for the success of the systems addressing singer identification in polyphonic excerpts is the voice detection algorithm. For instance, a poor voice detection method doesn't allow Kim and Whitman [31] to obtain an accuracy greater than 45 % in a database containing 250 songs by 20 different singers of popular music. They extract linear frequency scale and warped features, and classify using both SVMs and GMMs. Zhang [67], on the other hand, achieves a 82 % accuracy in a small dataset of 45 songs by 8 singers extracting LPMCCs and performing a good detection of the start of the vocal part of the songs. Shen, Shepherd, Cui and Tan [58] take into consideration also non-vocal parts for improving the performance of their singer identification system. They construct GMMs from low-level features and obtain accuracies from around 76 % to above 84 % in various large datasets. Several robustness testings are performed, demonstrating the stability of the results obtained in the presence of various acoustic distortion. By using background sound reduction (or even removal) accuracies have reached values close to those obtained in the monophonic case. In this way, Fujihara et al. [22] report a 94 % accuracy for a 10-singer dataset, the highest accuracy reached up to now.

Apart from allowing the detection of timbral characteristics of a song, some studies have also employed audio signals for detecting expressive characteristics in order to classify instrument performers. The research in this field has been mainly driven by Gerhard Widmer's team at Johannes Kepler University, in Austria, where they developed tools such as "The Performance Worm", which displays performance trajectories in the tempo-loudness space in synchrony with the music [66]. From this representation, general performance alphabets can be derived and used to model the distinctive expressive styles of famous pianists. Its application to performer classification has been explored in [55] and [60], showing that a quantification of the differences between music performers is possible. Ramirez, Maestre and Serra [52] apply sound analysis techniques based on spectral models for extracting deviation patterns of parameters such as pitch, timing, amplitude and timbre of commercial monophonic saxophone Jazz performances. They train vari-

ous machine learning algorithms in a 4-artist dataset and obtain classification accuracies around 70 % when dealing with phrase-length excerpts.

2.3.4 Related MIREX Tasks

Since 2005, the Music Information Retrieval Evaluation eXchange (MIREX) is the framework in which the MIR community formally evaluates their systems and algorithms [16]. Each year, MIREX organizes various music retrieval tasks, and groups from around the world submit their systems for benchmarking. From all the tasks that are organized in MIREX, Audio Artist Identification (AAI), that took place in 2005, 2007 and 2008, and Audio Classical Composer Identification (ACCI), from 2007 to 2013, are the ones that are more closely related to the goals of this thesis.

In MIREX 2005, two different datasets were used for the evaluation of the AAI task: Magnatune, providing 1158 training excerpts and 642 testing excerpts, and USPOP, providing 1158 training excerpts and 653 testing excerpts. The system designed by Bergstra, Casagrande and Eck [6] performed best on Magnatune, with an accuracy of 77 %. It considered a large number of frame-based timbre features, such as RCEPs, MFCCs, linear predictive coefficients, low-frequency Fourier magnitudes, rolloff, linear prediction error and zero-crossing rate. For the classification step, two versions of the system were constructed, one with AdaBoost with decision stumps and other with two-level decision trees. In the case of USPOP, the system that obtained the highest accuracy (68 %) was the one designed by Mandel and Ellis [40], which used 20-dimensional MFCC features and a SVM with KL divergence based kernel for classification. In MIREX 2007 and 2008 a new dataset was used, which included 3150 excerpts of 105 different artists. In those years the performance of the systems dropped dramatically below 50 %, even those that had already achieved good results in the MIREX 2005, which suggests that they had been tuned for obtaining the best results in specific datasets.

Table 2.1: Systems with highest and second-highest accuracies in MIREX ACCI task from 2007 to 2013

Year	First	Accuracy	Second	Accuracy
2007	IMIRSEL M2K	0.5372	Mandel, Ellis	0.4700
2008	Mandel, Ellis (1)	0.5325	Mandel, Ellis (2)	0.5310
2009	Wack et al.	0.6205	Cao, Li	0.6097
2010	Wack, Laurier, Bogdanov	0.6526	Seyerlehner et al.	0.6439
2011	Hamel	0.7900	Ren, Wu, Chang	0.6883
2012	Lim et al.	0.6970	Ren, Wu, Chang	0.6865
2013	De León, Martínez	0.7031	Lim et al.	0.6970

The ACCI task was first introduced in MIREX 2007. The dataset that was constructed for that edition has been reused every year up to now, and consists of 2772 audio excerpts

corresponding to 11 "classical" composers (J. S. Bach, Beethoven, Brahms, Chopin, Dvorak, Händel, Haydn, Mendelsson, Mozart, Schubert and Vivaldi). The highest accuracies obtained for this task in each year from 2007 to the last MIREX edition so far are summarized in Table 2.1. As it can be seen, the system by Hamel in 2011 is the one that achieves the highest accuracy of all that have been submitted for this task. However, the fact that this same system was also evaluated on 2012 with much lower results forces us to be very cautious of its meaningfulness. The second highest accuracy is obtained by De León and Martínez, who had already sent the algorithm in previous years with worse results. On the other hand, systems by Ren, Wu and Chang [54], and Lim et al. [36] perform consistently in different editions. Ren et al. extracted a fixed-length feature vector (composed of some timbral features as well as modulation spectrum features) from each training clip. Then, by representing a fixed-length feature vector as a linear combination of all training feature vectors, they classify this test clip as belonging to the class with the minimal reconstruction residual. This kind of approach is usually called sparse representation based classifier (SRC) and should be also taken into consideration. Lim et al. extract timbral features and some stastistical and modulation derived from the timbral attributes, which are then filtered with a SVM feature selection ranker. Finally, a SVM with Gaussian radial basis function kernel is used for classification.

As a final comment, it is important to note that most of the algorithms submitted to those tasks were not designed in principle specifically for AAI or ACCI. With very few exceptions they were all supposed to be general classifiers or were even designed for addressing other tasks, usually genre and/or mood classification. In fact, starting from 2010 ACCI has become one of the subtasks included in the global Audio Classification (Train/Test) task, so it is very unlikely that systems specially constructed for artist/composer identification will be evaluated in MIREX. However, strategies obtaining good results in those tasks, even when are designed to address other problems, should be taken into account.

2.3.5 Electronic Music Artist Identification

The specific task of identifying the artist (say, producer) with a dataset formed solely by Electronic Music songs has been only addressed, to the best of our knowledge, by Melidis [43]. He uses an in-house dataset comprised of 5949 songs corresponding to 111 artists of different sub-genres, from which 86 are finally used. For each artist 5 albums are collected, dedicating three of them to train the system and the other two for testing. A more detailed revision of the music collection will be held in Section 3.1.1.

A series of descriptors from different toolboxes are computed from 30-seconds excerpts corresponding to the second half of the first minute of each song. From MIRToolbox[5] a number of Dynamics, Rhythm, Timbre, Pitch and Tonal descriptors are computed, while from Essentia[6] only Spectral Energy, the MFCC coefficients and the HPCPs are obtained.

[5] `https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox`
[6] `http://essentia.upf.edu/`

He also calculates Zipf-based descriptors for MFCCs, Loudness and HPCPs as suggested by Haro et al. [24].

Various combinations of those features are used to train four different machine learning algorithms (Simple Logistic, LADTree, SMO Polykernel and NaiveBayes) obtaining a maximum accuracy in 5-fold cross-validation of 25.45 % with the SMO Polykernel classifier trained Essentia features when considering all 86 artists. In a hold-out test the accuracy achieved is approximately 10 % of accuracy, far above what a random classifier would obtain (1.16 %). If the number of artists is reduced to 20, the accuracy increases up to 39.04 % in cross-validation. The usage of Zipfian features demonstrates its usefulness, as the classification results with specific features, such as MFCCs, improve in every case.

2.4 Conclusion

A formal and well-delimited definition of what should be understood by Electronic Music and Style has been demonstrated to be a very difficult task. Nevertheless, we now have the basic knowledge necessary to start our research. Music Classification provides us most of the tools and techniques required for developing this research, and, more precisely, state-of-the-art Artist Identification systems show us which strategies are more likely to achieve good results. However, current trends in the MIR community lead towards the usage of multi-purpose classifiers instead of addressing each problem particularly. We strongly believe that, for achieving our goal, a much more targeted strategy should be followed. For that reason, further experimentation is needed in this field in order to better understand the underpinnings that conform Electronic Music artists' own stylistic traits, which, in addition, may help us to improve the predictive capacity of Artist Identification tasks specifically focused on this kind of music.

Chapter 3 | Methodology

The background review performed in the previous chapter has allowed us to better understand the framework over which we should develop the methodology for our research. In this sense, in this chapter we will report the main methodological steps performed, starting from the construction of the music collection in Section 3.1 and the extraction of audio features in Section 3.2, to finally describe the experiments that we have developed and report the overall results obtained in Section 3.3.

3.1 Music Collection

In any Music Information Retrieval task, the reliability of the music collection used for the development of the experimental analysis is a key point for ensuring the validity of the results obtained. We here present the steps that have been followed in order to construct the datasets that will be analyzed later, starting from the revision of the Electronic Music collection that Melidis constructed for the development of his Master Thesis [43]. We then report the most relevant aspects of the database that has been built to keep all the information well organized, and finally we explain the criteria that has been taken into account for selecting the tracks included in the datasets to be analyzed.

3.1.1 Revision and organization of the Music Collection

An investigation like the one developed here undeniably requires a very specific dataset that is not easy to collect. For the most of our knowledge there are no publicly available datasets focused solely on Electronic Music or, at least, that include enough EM tracks to perform our required analysis. Luckily enough we have been able to obtain the huge music collection that was constructed by Melidis for the development of his Master Thesis, that was, as we mentioned before, devoted to the Automatic Identification of Electronic Music Artists.

In his dissertation, Melidis indicated that the full music collection was comprised by 5949 tracks of 111 artists, organized in five albums per artist. Three of the albums by each artist were used for training a series of machine learning models and the remaining two were kept for the evaluation of the system. However, due to some inconsistencies regarding the amount of tracks that were included in each album, at the end "only" 86 artists were used in the experiments.

Initially, we planned to reproduce the experiments developed by Melidis as a starting point for our own investigation. Nevertheless, we soon realized that we lacked the information about which 86 artists were selected, and which tracks were considered, so we were forced to discard the idea. In addition to that, we also noticed that there were some other problems in the music collection that limited its usefulness for future investigations. More precisely, a non-exhaustive list of issues that were found includes:

- Some albums listed as being included in the collection were missing
- Some albums and tracks were not correctly labeled. For example, some albums were created independently by former members of considered groups, some albums and tracks were in fact collaborations with other artists, or even a few tracks by guests artists were also included in some albums
- Some tracks appeared in more than one album
- Some tracks were versions of other tracks also included in the collection (remixes, lives, ...)
- Some of the tracks included in the albums corresponded to Dj Mixes and not compositions of the considered artists
- Many tracks contained characteristic singing voices

Those issues, among others, needed to be solved in orther to ensure a reliable analysis. In this sense, some of the included albums had to be removed for different reasons as, for example, the impossibility of finding a source that ensured its authorship or the fact that it consisted solely on a Dj Mix. In some cases we had the chance to include new albums from our own collection in order to reach a minimum of 5 albums with 5 tracks per artist. Unfortunately, this was not always possible, so there are some artists in the collection that do not fulfill the minimum requirements that we have fixed. A few new artists were added to reduce the impact of this circumstance in the size of the final collection.

In order to avoid similar reproducibility problems in the future, we considered that the construction of a database containing all the relevant information related with the music collection would be useful not only for our own research, but also as a contribution to the community. For simplicity, we decided to use MySQL as Data Base Management System. The main goals addressed with the construction of the database were the following:

- It would be an oportunity to perform a close revision of the musical content
- It would ensure better organization and control of all the information related with the collection
- It would allow an easy access through different platforms, including eventually a web-based access for sharing the content with all the community

The stored information comprises data about the artists, the albums and the tracks

obtained from MusicBrainz[1] and the Echonest[2], as well as the Last.fm[3] Top Tags for each of the artists, information that may be useful for the construction of coherent datasets. Moreover, in the analysis step we have also stored information about the used datasets, its included excerpts and the descriptors extracted from each excerpt. All this huge amount of data is accessible through the interfaces that have been implemented in Python, Matlab and R programming languages.

The creation of the dataset has allowed us to revise individually each track and solve the aforementioned issues. In total, we have 6102 valid tracks, not considering repeated tracks, versions, remixes, collaborations between various artists, tracks by guest artists and Dj Mixes. Those tracks belong to albums by 114 artists, considering different aliases of one artist as a single artist, and albums authored by band members independently or in collaboration with other artists as not being from the same artist.

However, in the same way that was reported by Melidis, not every album in the collection contains a minimum of 5 tracks, the minimum that we have set in order to be valid for our analysis. That reduced the size of the music collection to "only" 83 artists. Even though we cannot take profit of every file stored in the collection, for our purposes this amount of artists was far above our needs. Sadly, an issue that was even be more problematic was the relatively high presence of vocal tracks. Some artists even showed a predominance of vocal over non-vocal tracks. As we mentioned in Section 2.3, human voice is a very relevant cue for artist identification, so allowing the inclusion of vocal tracks would have probably biased the results of the analysis. Filtering out any track that contained sung or spoken vocals reduced the final size of the collection to 3140 audio files corresponding to 64 different artists. A comprehensive list of all the artists and albums that are contained in this Electronic Music Artists (EMA) collection is included in Appendix A.

A relevant issue that should be taken into account is the unbalanced presence of Electronic Music (sub-)genres in the collection. This is a direct consequence of the particularities of some genres of EM, where there is a traditional preference for the releasing of Singles and EPs instead of full albums. Almost half of the artists included in the collection are labeled in Last.fm as "ambient", while very few are tagged as "house", a genre that is quite predominant in the EM market. Other dance-oriented genres, such as "trance", do not even appear in the tag list. In order to address that issue we have decided to generate an equal number of ambient and non-ambient small datasets including different sub-genres, as it will be reported in Section 3.1.2.

3.1.2 Selection of Analysis Datasets

The revision of the contents of the music collection that we have reported has reduced the number of artists and tracks to almost the half of its previous size. Nevertheless, in our opinion, the size of the collection is still too high to perform a stylistic analysis. Our goal

[1] http://musicbrainz.org/
[2] http://the.echonest.com/
[3] http://www.lastfm.com/

requires a closer look that is only feasible when a small number of elements is considered. For that reason, we have created six small datasets, with three artists each one, that include works by artists with perceptually similar styles. Each one includes exactly 5 albums per artist, and 5 tracks per album. Even though we will analyze each track completely and not extracting fixed-length excerpts, we have decided that every track should be have a minimum length of 60 seconds just in case in the future we have the chance to test our experiments with excerpts of 30 seconds length.

It is important to note that the decision to group the artist into perceptually-similar datasets doesn't mean that every track is stylistically similar to the rest that are included in the same dataset. It is not strange to find a Techno artist including Ambient tracks in his or her albums, as, for example, Jeff Mills. This kind of behaviors forces us to be very cautious when analyzing the results obtained in our experiments.

The artists included in the six specific analysis datasets that we have created and that will be used in the following chapters are listed below:

- Atmospheric Ambient: Michael Stearns, Tetsu Inoue, Vidna Obmana
- IDM Ambient: ISAN, Monolake, Tycho
- Post-Rock Ambient: Boards of Canada, The Album Leaf, The American Dollar
- Techno: Jeff Mills, Legowelt, Plastikman
- IDM: Aphex Twin, Autechre, Squarepusher
- Nu Jazz: Bonobo, Four Tet, Lemongrass

The detailed content of each of those datasets is reported in Appendix B.

We should keep in mind that the name to which we refer to each dataset is just a label that has been set using the tags that the included artists share in Last.fm. In any case we are stating that those artist are representative of genres that may be called in the same way as we are aware that genre labeling is quite conflictive.

When the artists of each dataset have been chosen between all the available ones, the selection of the specific albums (if more than 5 are elegible) and tracks is done randomly. This has been decided in that way in order to avoid subjective biases that may make up the final results of the analysis.

Once we have constructed a music collection suitable for our purposes and we have decided which specific artists are going to be analyzed, we can proceed to develop the experiments that should allow us to determine if a computational stylistic analysis of Electronic Music is feasible or not. The first step is to extract several audio features from the tracks contained in the analysis datasets by taking profit of different software libraries, as we will explain in Section 3.2. After that, in Section 3.3.1, we will test the predictive capacity of those features by performing a traditional Artist Identification task at different levels of aggregation of the datasets. Finally, two experiments specially designed for capturing the most relevant features for each artist are tested and reported in Sections 3.3.2 and 3.3.3.

3.2 Audio Features

Obtaining a set of features able to represent accurately the main characteristics of the works by the considered artists is essential in order to perform a stylistic analysis. For that purpose, we have decided to use descriptors provided by three different software libraries, as it is explained in Section 3.2.1. However, some features will be more relevant to characterize the works by some artists than others. For that reason, in Section 3.2.2 we will introduce a nomenclature that will help us to express different types of features according to their influence in the characterization of one artist's style.

3.2.1 Extracted Features

In broad terms, there are two main strategies that can be followed when deciding which features to use for developing studies similar to ours. The one that is usually employed in purely stylistic studies relies in a previous knowledge of the music to be analyzed to create tailored features able to capture representative traits of that particular music. We can find this strategy in most studies based on symbolic sources, creating features such as, for example, "countrapuntal" descriptors when analyzing baroque music. This Top-Down strategy contrasts with the one that can be found in most of the research addressed to perform music classification tasks, which usually constructs prediction models based on several general-purpose audio features and lets the algorithm to explore which of those features are able to generate the best performance.

Even though our main purpose is not to achieve the highest classification accuracy, we have decided to perform our analysis using a Bottom-Up strategy. The undeniable exploratory nature of this research leads us to consider as many features as possible as candidates for being representative of the style of some Electronic Music artists. In this way, we expect to increase our knowledge of this kind of music, allowing us to design in the future tailored features with a higher guarantee of success.

We have used three different software libraries to obtain the audio features. First of all, Essentia provides us several features that cover a wide range of musical facets, such as timbre, rhythm, tonality and dynamics. The Freesound Extractor[4] includes most of the useful descriptors from Essentia, but we have decided to also compute the features returned by the Predominant Melody algorithm[5] (Pitch and Pitch Salience), as well as Danceability and Dynamic Complexity. Matlab's library MIRToolbox[6] is another great source of audio features which we have decided to include. However, some of the descriptors, such as Spectral Centroid, returned null values for some tracks and had to be discarded. Finally, we decided to test the features that can be obtained from The Echonest Analyzer[7] via its API. It provides some low and high level audio features, and also detailed metadata that has

[4] `http://www.freesound.org/docs/api/analysis_docs.html`

[5] `http://essentia.upf.edu/documentation/reference/std_PredominantMelody.html`

[6] `https://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/MIRtoolbox1.5Guide`

[7] `http://developer.echonest.com/docs/v4/_static/AnalyzeDocumentation.pdf`

been attached to the collection. However, probably the most useful information that this source provides is a very detailed structural and rhythmic analysis, with data at different levels such as Section, Segment, Bar, Beat and Tatum. Unfortunately, the documentation of the algorithms that this extractor computes is scarce, almost like a "black-box" for us.

Additionally to the features obtained "directly" from the extractors, we have decided to compute a series of derived features that we expect to be helpful for our analysis, as they roughly capture some of the temporal evolution of the features:

- Statistical moments (Mean, Median, Variance, Skewness and Kurtosis) of:
  - Rhythmic Divisions (Duration and Confidence)
  - Structural Divisions (Duration and Confidence)
  - Low and High level features at Section / Segment level (Dynamics, Timbric, Tonal)
  - Rhythmic Structure (tatums per beat, beats per bar, ...)
- Segment and Section rate

Globally, we have available more than 500 audio features for each track, which are listed comprehensively in Appendix C. We expect that having such a big and varied amount of features covering a wide range of musical facets will allow us to represent a good part of the stylistic traits of each artist. However, we should keep in mind that some of them are repeated in different libraries, and that their values are not always coherent between them. Moreover, we should take into consideration that the list is far from being exhaustive, as some features not included in any of the three libraries, such as, for example, audio modulation descriptors, are missing.

3.2.2 Nomenclature

Traditional clasification tasks, such as Automatic Artist Identification, focus on finding models capable to distinguish between the considered classes. Therefore, only those features that increase the distance between those classes are usually taken into account. However, for the purpose of this Thesis we need to be able to identify every aspect of the audio signal that can be relevant for characterizing its own style.

In this sense, it is important to note that not every feature that may help us to distinguish between authors has to be representative of that artist. When an artist creates a piece completely different from any other previous work ever composed, be his or not, it is be very easy to use the particular characteristics of that specific work to determine its authorship. Think, for example, in John Cage's 4'33". Its uniqueness allows us to perfectly recognize it as being created by Cage. Nevertheless, this does not mean that having 4 minutes and 33 seconds of complete silence is a descriptive characteristic of the works by Cage, as no other piece is similar to that one. On the other hand, some features that are common between different artists can also be very relevant. In that way we can talk about

general styles, schools, etcetera. Moreover, if we wanted to replicate the style of a certain artist by only taking into account those features that only appear in the works by that author it is very likely that not only we don't get anything similar to his or her style, but also that the new piece does not sound musical at all as we would have not considered those attributes that capture the building blocks of the musical school to which that style belongs to.

Those considerations led us to hypothesize that at least two different kinds of features are needed to capture the most relevant traits of the style of one artist. From now on we will use the terms "discriminative" and "descriptive" feature to refer to the aforementioned cases. The schema presented in Fig. 3.1 may be helpful to provide a more formal definition of those concepts. Be A the set of feature values found in the works of a particular artist (and B the equivalent to another artist) and a (b) the subset of A (B) comprised by those features that show a restricted range of values for that artist. Then, we could consider a (b) as being "coherent"[8] throughout most of the works of an artist and we will note them as his or her "descriptive" features. Those features that do not intersect with the ones by the other artist, this is A - A $\cap$ B (B - B $\cap$ A), are the ones that will help us to distinguish the works of an artist in relation with another. Thus, we will consider them as the "discriminative" features of that artist. Finally, we can add a third category to our nomenclature by considering those features that are both descriptive and discriminative at the same time. In the scheme they are represented as a - A $\cap$ B (b - B $\cap$ A) and we will use the term "characteristic" features to refer to them.

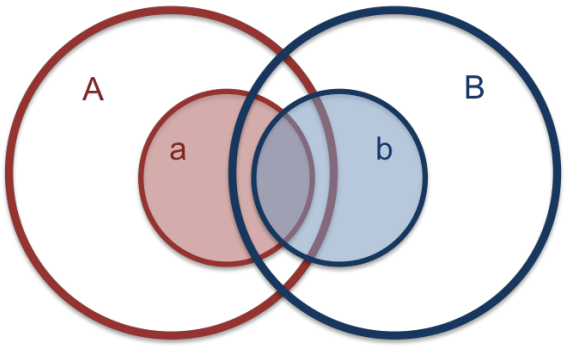

Figure 3.1: Simplified schema representing the feature space found in the works by two artists noted as A and B. The smaller coloured circles (a and b) indicate those features that can be considered as descriptive for each of the artists

In Sections 3.3.2 and 3.3.3 we will explain the procedures that we propose in order to capture the most discriminative and descriptive features for each artist. Nonetheless, before

[8]The interpretation that in this dissertation will be given to the term "coherence" should not be confused with the statistic "spectral coherence" that is used in signal processing to examine the relation between two different signals, usually input and output of a system. In our case, the meaning of "coherence" is closer to the one employed in probability theory to describe the degree of self-consistency across multiple assessments, which we can assimilate to the consistency of an audio feature across various tracks of the same artist

that we will report in Section 3.3.1 the results obtained when performing a traditional Artist Identification task.

3.3 Performed Experiments

3.3.1 Artist Identification

Being aware of the issues concerning the extracted features that we have previously mentioned, we thought it would be interesting to test the ability of those features as cues for a traditional Artist Identification task as our first experiment. At this respect, given a collection of musical tracks or excerpts, Automatic Artist Identification seeks to build a computational model able to predict the artist to which each instance belongs to. The term artist can refer to a variety of roles in the music industry, but in this case we will deal with the prediction of the creator/composer rather than the interpreter/performer. For that purpose we will use the previously extracted audio features to train a machine learning model, which will then be used to perform predictions about the corresponding artist of each track. In the following paragraphs we will briefly describe the procedure developed for this purpose and report the overall results obtained.

Data

The audio files used for this experiment are those included in the analysis datasets described in Section 3.1.2 and listed in detail in Appendix B. The experiment has been performed individually for each of the six datasets, thus, considering three different artists each time, as well as joining all of them in a single dataset ("Complete") including the 18 artists. The usage of the small datasets will allow us to analyze in detail the implications of the results obtained, as we will do in Chapter 4, while the joined datasets will provide us valuable information for determining the goodness of the performance of our own experiment compared to previously reported studies.

With respect to the audio features considered, we have decided to test the predictive capability not only of the complete set of extracted features ("All"), but also of a few smaller subsets created according to different criteria. In this sense, we have grouped them by extractor -Essentia ("ESS"), MIRToolbox ("MTB") and Echonest ("EN")- and by musical facet -Timbre ("Tim"), Rhythm ("Rhy"), Pitch/Tonality ("Pit_Ton") and Dynamics ("Dyn")-. We have also generated subsets including only High Level ("HL") features and the derived descriptors computed from the Echonest Structural Analysis ("Str").

Procedure

The procedure followed to perform this particular experiment has been designed according to the general framework for automatic music classification presented in Section 2.3.1.1.

As we mentioned before, we have performed multiple iterations considering the different combinations of datasets and feature subsets. In each of those iterations the first step is to generate the train set that will be used to construct the prediction model and the test set over which the evaluation of the system will be performed. Among all the possible choices regarding the composition of those sets, we have decided to force a worst-case scenario by completely separate the tracks included in the test set of those that were used for training. The evaluation that is performed in this way is usually called a "Holdout test", and is rarely found in reported music classification studies as it causes the obtained accuracies to be generally lower than other methods such as N-fold Cross Validation. In our case, we have selected three albums per artist for the training set, and the remaining two have been included in the test set. This implies that the 60 % of the tracks of each of the considered datasets have been used to build the prediction models, while the remaining 40 % were used for the evaluation stage. In order to minimize the possible impact of style change during the career of the studied artists in the constructed models, we have ordered chronologically the albums of each artist and assigned alternatively each one either to the train or test set.

Regarding the construction of the prediction models, studies related with music classification often report a comparison between the results obtained by using different learning algorithms in order to determine which one leads to a higher performance. Having in mind that the main goal of this project is not to report the highest possible classification accuracy but to be able to analyze the stylistic traits of some artists, we have decided to avoid training of different learning algorithms and focus on the artists and the features. In this sense, as we already saw in the background review, Support Vector Machines (SVM) is one of the learning algorithms that is more frequently used in this kind of tasks. Even though we are aware that other techniques, such as the boosting of weak learners, have obtained similar or even higher performances, we decided to employ SVM mainly to be able to compare the results obtained by our feature sets with the ones reported in similar studies. Moreover, in our opinion it provides a good balance between high prediction power and model understandability. Therefore, we have performed our experiment by training SVM models using the LibSVM implementation included in R's e1071[9] package and using them to predict the artist of each track included in the test sets.

Results

A summary of the average classification accuracies obtained when performing the experiment described previously is shown in Table 3.1. As it was mentioned before, the values reported in the table have been computed when predicting the author to which the tracks included in two albums per artist belong to, using a SVM model constructed from the audio features of the remaining three albums per artist.

The first criteria that we should take into account to determine the validity of our results

[9] `http://cran.r-project.org/web/packages/e1071/index.html`

Table 3.1: Classification accuracies obtained with a 60 %-40 % holdout test in the Artist Identification experiment for different datasets when training SVM models using different feature sets

Dataset	ALL	ESS	MTB	EN	Tim	Rhy	Pit_Ton	Dyn	HL	Str
Complete	**50.56 %**	36.11 %	36.11 %	**47.22 %**	32.22 %	38.33 %	25 %	31.67 %	21.67 %	44.44 %
Atm Amb	**70 %**	**73.33 %**	66.67 %	63.33 %	60 %	66.67 %	70 %	56.67 %	50 %	66.67 %
IDM Amb	**93.33 %**	**93.33 %**	76.67 %	90 %	70 %	76.67 %	90 %	66.67 %	76.67 %	86.67 %
PR Amb	**70 %**	63.33 %	56.67 %	63.33 %	50 %	63.33 %	66.67 %	50 %	60 %	**70 %**
Techno	**96.67 %**	83.33 %	70 %	**96.67 %**	46.67 %	73.33 %	63.33 %	73.33 %	53.33 %	86.67 %
IDM	43.33 %	36.67 %	**50 %**	40 %	**53.33 %**	43.33 %	30 %	33.33 %	46.67 %	43.33 %
Nu Jazz	66.67 %	66.67 %	**70 %**	53.33 %	60 %	**70 %**	43.33 %	46.67 %	50 %	63.33 %

is comparing them with the random baseline for the corresponding number of classes. In the case of the complete dataset, which includes works by 18 different artists, this baseline is around a 5.55 % of classification accuracy. This means that by using any of the tested feature sets we are able to get results far above mere chance. This conclusion is not always true in the case of the individual analysis datasets, as the one comprised by the "IDM" artists show for some feature sets results very close or even below the random baseline of 33.33 % for three artists. However, this fact is not too surprising, as the artists included in that dataset are by far the ones of the collection that present a higher level of experimentation in their works.

In the table we have marked in bold letters the highest and second highest accuracies for each dataset. One of the most relevant aspects that can be observed in those results is that for specific datasets the accuracy reached using a non-exhaustive feature set is above the one obtained when using all the extracted features. This happens, for example, when using only the Essentia features in the Atmospheric Ambient dataset. This fact suggests that some of the inconsistencies that we have already commented are somewhat tricking the model. In the future it will be interesting to investigate this point more closely.

The only previously reported experiment which may be useful for determining the degree of success of our own test is the one performed by Melidis using 20 artists. He does not report the accuracy reached in a holdout test for this dataset size, but he reaches 39.05 % of accuracy in 5-fold cross-validation. Taking into account that our 50.56 % of accuracy when dealing with 18 different artists has been obtained in a 60-40 holdout test, we are fairly confident that the performance of our system improves current State-of-the-Art in similar tasks. However, the fact that we are using full songs instead of 30-seconds excerpts may be distortioning the comparability of both results, so we should be very cautious in this aspect. In this sense, replicating the experiment but using fixed-size excerpts of different length (and location within the track) may be of interest.

Figure 3.2 represents the confusion matrix for this experiment, obtained when using the complete feature space when considering the 18 different artists. As it can be seen, most of the predictions are located within or close to the diagonal of the matrix. This indicates that even when the system does not successfully predict the correct artist of a track, it usually confuses it with another artist within the same dataset. However, this is

not always true. In some cases, such as within the "Techno" dataset, every confusion has been performed with artists not contained in that particular dataset. According to our knowledge of the specific dataset, it may be caused by the inclusion of a relevant amount of Ambient tracks in the albums of those artists.

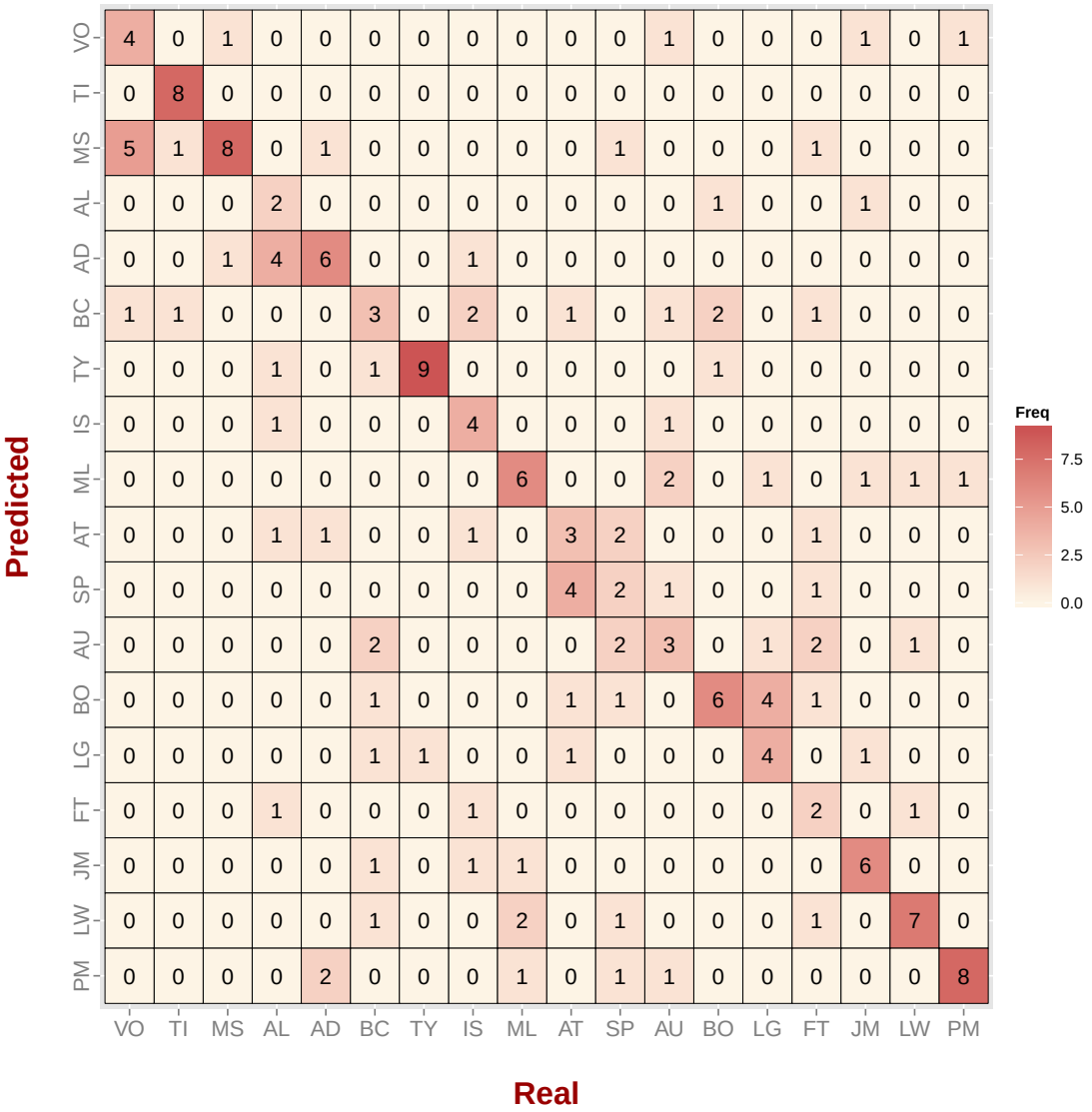

Figure 3.2: Artist Identification confusion matrix using the complete dataset of 18 artists. The included artists are: Vidna Obmana (VO), Tetsu Inoue (TI), Michael Stearns (MS), The Album Leaf (AL), The American Dollar (AD), Boards of Canada (BC), Tycho (TY), ISAN (IS), Monolake (ML), Aphex Twin (AT), Squarepusher (SP), Autechre (AU), Bonobo (BO), Lemongrass (LG), Four Tet (FT), Jeff Mills (JM), Legowelt (LW) and Plastikman (PM)

Even though it would probably be very revealing to analyze every particular case, it is not feasible. For that reason, we have decided to analyze in detail in Chapter 4 a specific case which will include a close review of the confusions made in the Artist Identification task in Section 4.1.

3.3.2 Discriminative Features

In Section 3.3.1 we have reported satisfactory results when performing an Artist Identification task. Nevertheless, we should not forget that our main goal is to determine which are the features that better represent the style of *each* artist individually or group of artists, and a traditional Artist Identification task is far from providing us that information. In this sense, as we introduced in Section 3.2.2, retrieving which are the features that help us to distinguish the works created by an artist to those that don't is a fundamental step for performing a stylistic analysis. In the following paragraphs we will describe the methodology that we have developed to address that goal.

Data

For performing this experiment we have used exactly the same audio files that we employed in the Artist Identification task. However, in this case we have been forced to slightly modify the labelings in order to fulfill the requirements of the described problem. More precisely, the main idea behind the "discriminative" features is that they should be able to capture those traits that allow us to distinguish the works by one artist to those created *by any other artist*. This implicitly requires a binary classification task, so for each analyzed artist we have grouped the works by the remaining artists included in the considered dataset into a single category that we have labeled as "Others". We have performed this operation not only for each of the three-artists analysis datasets, but also grouping the artist into two bigger datasets comprising nine Ambient and Non-Ambient artists respectively. Unlike what we did in the Artist Identification experiment, in this case we have not considered a complete dataset with all the 18 different artists because of the excessive amount of instances that would have been included in one sigle class compared with the other.

With respect to the audio features considered, we have decided to use the complete feature set and not including the evaluation of smaller subsets. As the main goal of this experiment is in fact to obtain subsets of descriptors it will have very little sense to explicitly reduce the amount of employed features.

Procedure

Exactly in the same way as in the previous experiment, once we have decided which files and features are considered, the first step of the experimental procedure is to create the training and test sets. Again, with the purpose of testing our system in a worst-case scenario, we have selected 60 % of the tracks for training the learning algorithm and we have kept the remaining 40 % for the evaluation stage. We should be aware that this procedure leads to an unbalanced amount of instances per class, as the "Others" category will include all the works by a minimum of two artists. Depending on the amount of artists considered

this unbalance may cause the prediction algorithm to perform much worst than usually, as the class with more instances will be preferred when in doubt. Moreover, the need to avoid tracks from the same album in the train and test sets may lead to, depending on the size of the specific dataset, a different weighting of artists in the train and test sets. For example, when dealing with an analysis dataset of 3 artists, two of them will be grouped as "Others". Therefore, ten albums representing that class will require to be splitted into the train and test sets. As this splitting is done chronologically, in some cases the amount of albums per artist may end being unbalanced. Those issues should be taken into account and may require further investigation to determine if they bias the obtained results.

Apart from the transformation to a binary classification task, the main point that differentiates this procedure to the one previously described is the inclusion of dimensionality reduction methods to determine which features have a higher influence in the discriminative power of the system. From all the possible candidates that would allow us to perform this task, we have decided to only consider a few feature selection filters. Mire precisely, we have tested the implementation included in R's FSelector[10] package of the following feature selection methods: Chi Squared, Information Gain, Gain Ratio and Symmetric Uncertainty. It is known that other techniques, such as Principal Components Analysis (PCA) or feature selection wrappers, create feature subsets with higher discriminative power for the same set size. However, the former case generates a transformation of the features which probably would not allow us to interpret its meaning and link the created features with perceptual properties of the music, while the result of the latter is too dependent on the learning algorithm employed and not necessarily generalizable. Those reasons led us to consider that for our purposes the main priority was to ensure the generation of understandable and algorithm-independent feature sets, even if the quantitative results were not optimal.

Once a feature subset of size N has been created over the training set according to a certain feature selection algorithm, a SVM model is trained and finally evaluated by predicting the classes over the test set. The overall procedure is iterated to cover every artist included in all the considered datasets (and groups of datasets).

Results

Tables 3.2 and 3.3 show a summary of the performance of the previously described method when generating subsets of 20 features for each artist of the considered datasets. In the first case, the discrimination was performed only to those other artists included in the same three-artists analysis dataset, while in the second we sought to distinguish them to any other artist that had been fitted in the same Ambient/Non-Ambient category as them.

We have chosen F1-Score[11] as the figure-of-merit for reporting the results of this ex-

[10] `http://cran.r-project.org/web/packages/FSelector/index.html`

[11] F1-Score is computed as $2 \cdot (precision \cdot recall)/(precision + recall)$

Table 3.2: F1-Scores obtained when trying to distinguish the works from each artist to those from the other artists included in their own analysis dataset using the 20 most discriminative features according to different feature selection methods

Dataset	Artist	No FS F1-Score	Chi Squared F1-Score	Info Gain F1-Score	Gain Ratio F1-Score	Sym Unc F1-Score
Atmospheric Ambient	Vidna Obmana	0.3333	0.7500	0.6316	0.3333	0.6316
	Tetsu Inoue	0.7500	0.8182	0.8182	0.7500	0.8182
	Michael Stearns	0.2667	0.6667	0.6667	0.2667	0.6667
	Average F1-Score	**0.4500**	**0.7450**	**0.7055**	**0.4500**	**0.7055**
IDM Ambient	Tycho	0.8889	0.9000	0.8421	0.8889	0.8421
	ISAN	0.6957	0.7619	0.8182	0.6957	0.8182
	Monolake	0.9474	0.8696	0.8696	0.9474	0.8696
	Average F1-Score	**0.8440**	**0.8438**	**0.8433**	**0.8440**	**0.8433**
Post-Rock Ambient	The Album Leaf	0.5333	0.4706	0.5556	0.5333	0.2857
	The American Dollar	0.5882	0.6667	0.6364	0.5882	0.6364
	Boards of Canada	0.6667	0.7619	0.8000	0.6667	0.6667
	Average F1-Score	**0.5961**	**0.6331**	**0.6640**	**0.5961**	**0.5296**
Techno	Jeff Mills	0.0000	0.5000	0.4211	0.0000	0.4444
	Legowelt	0.6667	0.7368	0.6364	0.6667	0.6364
	Plastikman	0.5882	0.7368	0.7619	0.5882	0.6957
	Average F1-Score	**0.4183**	**0.6579**	**0.6065**	**0.4183**	**0.5922**
IDM	Aphex Twin	0.0000	0.1333	0.2222	0.0000	0.2353
	Squarepusher	0.0000	0.4444	0.4444	0.0000	0.4444
	Autechre	0.0000	0.5217	0.4545	0.0000	0.4348
	Average F1-Score	**0.0000**	**0.3665**	**0.3737**	**0.0000**	**0.3715**
Nu Jazz	Bonobo	0.1818	0.4706	0.4706	0.1818	0.4706
	Lemongrass	0.5714	0.7368	0.8182	0.5714	0.8182
	Four Tet	0.4286	0.4545	0.4762	0.4286	0.4762
	Average F1-Score	**0.3939**	**0.5540**	**0.5883**	**0.3939**	**0.5883**
	Total Average F1-Score	**0.4504**	**0.6334**	**0.6302**	**0.4504**	**0.6505**

periment, as it captures both the precision and the recall of the preditions. In other words, it penalizes both the mistakes made by predicting a work created by the considered artist as being from another author, and those commited by labeling as done by the artist when in fact belongs to another artist. We think that, in this case, both kinds of mistakes are equally important. On the other hand, using classification accuracy would have not been representative, as the unbalanced amount of negative instances would have hidden the real performance of the experiment.

On the tables, it can be seen that selecting only the 20 most discriminative features overperforms the same task when using a non-focused feature space of above 500 different descriptors. According to those results, apart from Gain Ratio, that gets exactly the same performance as the complete feature space, the other three methods have no significant differences between them in the accuracies that they are able to reach.

One of the most surprising results that is shown in both tables is the extremely low performance that is obtained when using the full feature space. This is even more shocking if we compare them with the results reported in Section 3.3.1. The most probable explanation to this significant decrease of the performance is the unbalanced amount of instances of the positive and negative cases. It would be interesting to replicate the same experiment but this time forcing the amount of negative instances to be the same as the number of positives by randomly picking 10 excerpts among all the possible candidates. As the selection of instances would be randomized, and the results may vary a lot depending

Table 3.3: F1-Scores obtained when trying to distinguish the works from each artist to those from the other artists from the Ambient or Non-Ambient categories using the 20 most discriminative features according to different feature selection methods

Dataset	Artist	No FS F1-Score	Chi Squared F1-Score	Info Gain F1-Score	Gain Ratio F1-Score	Sym Unc F1-Score
Ambient	Vidna Obmana	0.0000	0.5333	0.4286	0.0000	0.5882
	Tetsu Inoue	0.5714	0.4211	0.3529	0.5714	0.4444
	Michael Stearns	0.0000	0.1818	0.0000	0.0000	0.1667
	Tycho	0.3333	0.5882	0.6667	0.3333	0.5000
	ISAN	0.0000	0.3333	0.3333	0.0000	0.3333
	Monolake	0.8235	0.9474	0.9474	0.8235	0.9474
	The Album Leaf	0.0000	0.0000	0.2667	0.0000	0.0000
	The American Dollar	0.3333	0.6667	0.7368	0.3333	0.5556
	Boards of Canada	0.0000	0.2857	0.2857	0.0000	0.2857
	Average F1-Score	**0.2291**	**0.4397**	**0.4465**	**0.2291**	**0.4246**
No Ambient	Jeff Mills	0.0000	0.0000	0.3077	0.0000	0.1429
	Legowelt	0.0000	0.4615	0.5217	0.0000	0.4545
	Plastikman	0.0000	0.8000	0.6250	0.0000	0.6316
	Aphex Twin	0.0000	0.0000	0.0000	0.0000	0.0000
	Squarepusher	0.0000	0.1538	0.1538	0.0000	0.1538
	Autechre	0.0000	0.0000	0.0000	0.0000	0.0000
	Bonobo	0.0000	0.1667	0.1667	0.0000	0.1667
	Lemongrass	0.0000	0.2500	0.2500	0.0000	0.2857
	Four Tet	0.0000	0.0000	0.0000	0.0000	0.0000
	Average F1-Score	**0.0000**	**0.2036**	**0.2250**	**0.0000**	**0.2040**
Total Average F1-Score		**0.1145**	**0.3216**	**0.3357**	**0.1145**	**0.3143**

on which excerpts are picked, an averaging between a certain number of repetitions would be also interesting. In this way we could determine if the observed improvement that we get when using Feature Selection also appears in a balanced dataset.

A qualitative analysis of the features that the system is returning as being the most discriminative for each artist is not performed here as it is not feasible due the number of descriptors obtained. However, in Section 4.2 we will develop a detailed review of the discriminative features of the artists included in a specific dataset, trying to determine if they link to the perceptual characteristics of each of those artists.

3.3.3 Descriptive Features

As we mentioned before, we consider that determining only the discriminative features of an artist is ignoring a big part of the characteristics that are relevant to the style of that artist. In this sense, we should also try to identify the most "descriptive" features associated with an artist (or group of artists), or, in other words, our goal is to find those features whose values are more "coherent" within the works of the artist (or group of artists).

Having said that, determining what coherence means in this context is not trivial. We may understand it as equivalent to "less disperse", in the sense that a feature would be more coherent as their values accumulate closer to the mean of its distribution. The statistical measures that would capture the coherence of a feature would be, thus, the Standard Deviation/Variance and/or the Inter-Quartile Range. Obviously, in any of those measures, the lower the value, the more coherent a feature should be considered. However, there

could be some cases in which this definition is not as reasonable as it should be. Consider, for example, the case shown in Figure 3.3. As it can be seen, the distribution labeled as "A" contains only two values located at the extremes of the range (0, 1), while distribution "B" has uniform probabilities in the same range. If we assume that both represent *continuous* variables it is reasonable to argue that distribution A is more "coherent" than distribution B. If this is the case, then the degree of dispersion of the distribution is not capturing its coherence, as A has a higher Standard Deviation than B.

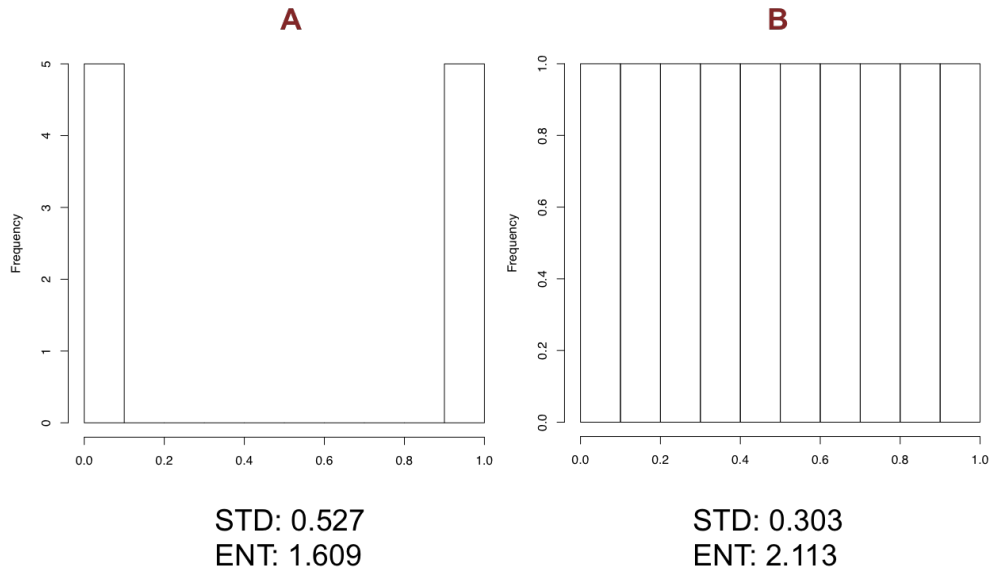

Figure 3.3: Probability density functions of a multimodal feature and an uniformly distributed variable, indicating their respective standard deviation and entropy

To deal with situations similar to that explained above, we may consider an alternative understanding of what "coherence" means. The degree of predictability, or in other words, the degree of uncertainty, has the potential to represent the coherence level of a variable. The measured entropy of the two distributions showed in Figure 3.3 is more closely related with our personal understanding of coherence. For that reason, we will also compute this measure and compare the features selected in terms of dispersion with those selected according to their entropy.

One important point that should be taken into account before selecting the features, is that each of them is computed in its own scale. Even though this seems quite obvious, it has a very relevant consequence when trying to determine those features that have less dispersion. Fig. 3.4 shows an example that may be helpful to understand the implications of this factor. If we compare directly dispersion measures of the observed values of two different features we may think that they are both similarly coherent, or even selecting them in the oposite way. In the figure the observed values of the Average Loudness have more standard deviation than the observed values of the Dissonance. However, if we force them to be in the same scale, say 0 to 1, by normalizing them using the minimum and the

maximum of all the possible values of the variable or, at least, the minimum and maximum observed values for all the analyzed data, we may get a very different result. If we compare the normalized values for the same two features we now realize that Dissonance is, by far, much more disperse than Average Loudness. As a consequence, we will always compare the dispersion measures only after having normalized the values of the features.

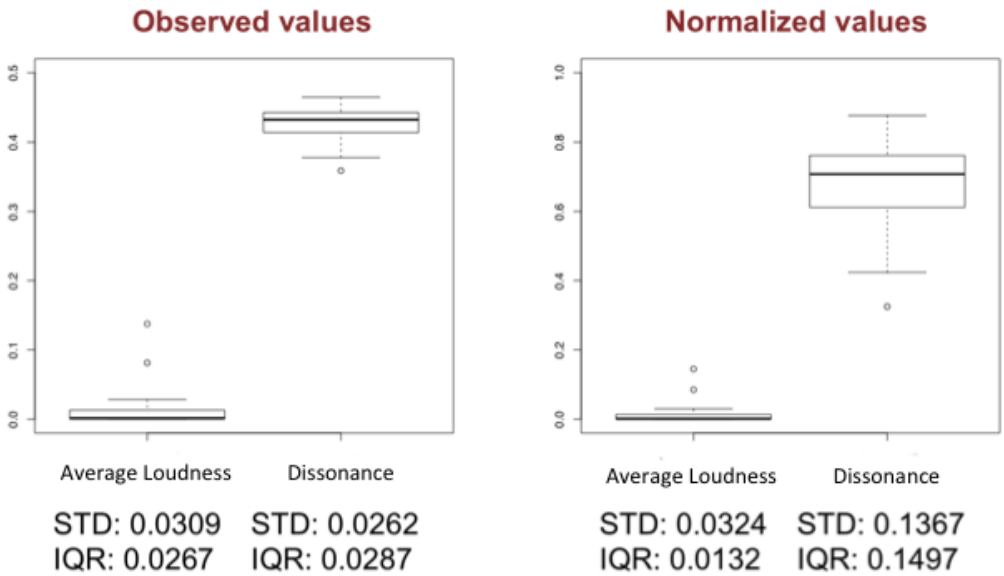

Figure 3.4: Dispersion measures computed over observed and normalized values of the Average Loudness and Dissonance descriptors

On the other hand, the normalization step is not required for the ranking of features according to its entropy, as the obtained value of this measure is independent to the scale of the variable. This is arguably an advantage as not only less computations are needed, but also we get a reference value that will not change even if we analyze new excerpts. This is not true in the case of the dispersion measures, as a new analyzed excerpt may cause the minimum and maximum observed values of some features to change. This implies that the descriptive features that are selected according to the dispersion criterion may change everytime the dataset is updated, causing some inconsistencies.

Data

For developing this experiment we have considered again all the tracks included in the different three-artists analysis datasets. However, the main difference in this sense is how they have been be grouped, as they have not only been considered at dataset level, but also each one individually, as it will be explained in the description of the Procedure. Regarding the features, in the same way that we have already reported in the previous experiment, we have also employed the complete space of extracted features. Nevertheless,

in order to select the features according to their dispersion measures (standard deviation and inter-quartile range), their values were normalized to a range between 0 and 1.

Procedure

Even though the selection of audio files and features has been performed almost exactly like the previous experiments, with the exception of the normalization step when needed, the idea behind this experiment is, in broad terms, the opposite of a traditional classification task. More precisely, our purpose is to test if the features selected by means of the different possible measures are common among the different works included in a group (which can be the works of an individual artist or those created by various artists that we consider perceptually similar). The main hypothesys is that by solely using the features that are considered as descriptive of a group, we should have more difficulties to distinguish between members of that group than with any other set of features.

As always, the first step once we have selected which tracks will be included in the experiment and which features will be considered in the analysis, the next step is defining the contents of the training and test sets. Having in mind the purpose of the experiment, we consider that the worst-case scenario is achieved when both sets contain *exactly the same data*. Taking into account that the numeric value that indicates the validity of our analysis is the difference between the classification accuracy using descriptive and non-descriptive features, facilitating the task of the prediction algorithm represents making it more difficult for the experiment to get good results.

A SVM model is then trained by using all the works contained in the considered dataset, and this model is used to predict the classes of a sub-level inside the dataset. For example, if we are considering the descriptive features for a group of artists, the model will try to predict to which individual artist belongs each track, while if we are dealing with the descriptive features of particular artists, the task will consist on trying to distinguish between albums of that artist.

Results

In order to determine the validity of the designed experiment, as we mentioned before, we have decided to perform it at two different levels. In this sense, Figure 3.5 represents the differences in the achieved classification accuracy when trying to predict the artist to which a particular excerpt belongs to by using descriptive and non-descriptive feature sets of different sizes. The accuracies obtained in the non-descriptive cases are averages of the results achieved by 10 randomly selected feature sets of size N. Table 3.4 summarizes the mean values of each size considering the six different datasets. As it can be seen, in every case reported in the figure the non-descriptive features are able to get better results than the descriptive ones. This means that the selected descriptive features are not able to capture the discriminative characteristics of each artist. With respect to the different

measure candidates, it is clear that the Inter-Quartile Range performs worse than the other two. Standard Deviation and Entropy achieve similar results, especially when we don't take into consideration subsets of very small size.

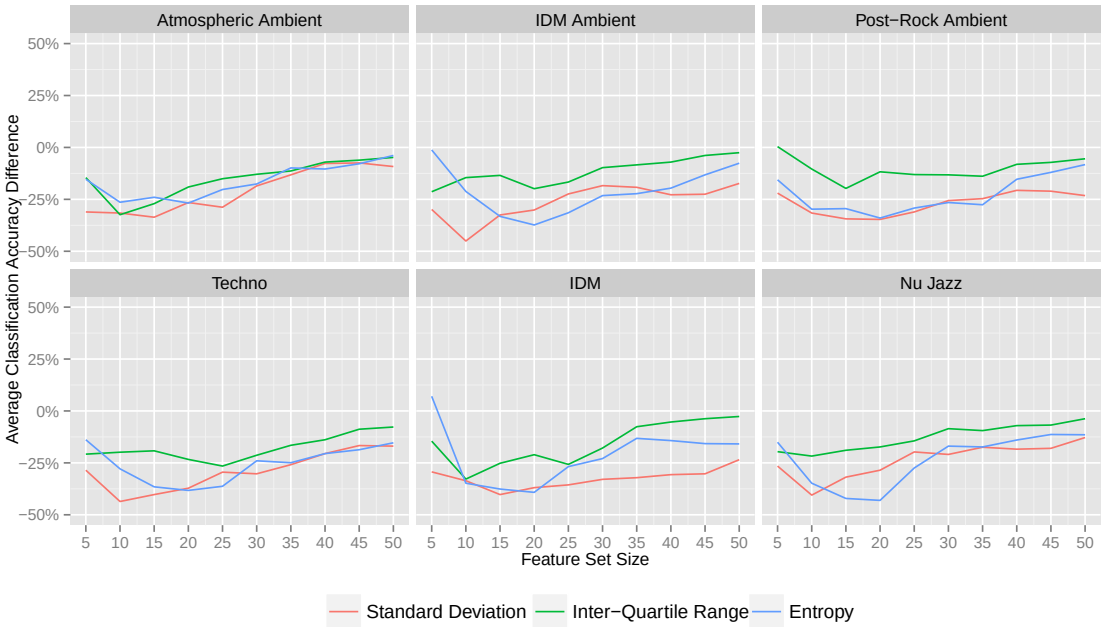

Figure 3.5: Effect of datasets' descriptive features set size (X-axis) and candidate selection method (legend) on the artist identification accuracy

Table 3.4: Mean differences in the artist identification classification accuracy for different between descriptive and non-descriptive feature sets of different sizes selected by means of different measures

Size	Standard Deviation	Inter-Quartile Range	Entropy
5	-27.89 %	-15.07 %	-9 %
10	-37.67 %	-21.96 %	-29.13 %
15	-35.49 %	-20.6 %	-33.82 %
20	-32.33 %	-18.73 %	-36.44 %
25	-27.84 %	-18.58 %	-28.58 %
30	-24.44 %	-13.93 %	-21.87 %
35	-22.09 %	-11.2 %	-19.2 %
40	-20.13 %	-8.09 %	-15.69 %
45	-19.33 %	-6.09 %	-13.13 %
50	-17.16 %	-4.49 %	-10.4 %
Mean 05-50	**-26.44 %**	**-13.87 %**	**-21.73 %**
Mean 15-50	**-24.85 %**	**-12.71 %**	**-22.39 %**

On the other hand, when focusing on the descriptive features per artist, the results are not as satisfactory as the ones previously reported. As it can be seen in Fig. 3.6, when trying to predict the album to which a track belongs to, the accuracies are not as

different for the descriptive and non-descriptive features as they were previously. This may be caused by the "Album Effect", but further investigation is needed to confirm that assumption. Nevertheless, the results continue showing some difference of performance, as it can be seen in Table 3.5.

Table 3.5: Mean differences in the album identification classification accuracy for different between descriptive and non-descriptive feature sets of different sizes selected by means of different measures

Size	Standard Deviation	Inter-Quartile Range	Entropy
5	-29 %	-15.29 %	-5.27 %
10	-24.89 %	-9.6 %	-11.11 %
15	-14.71 %	-7.89 %	-12.44 %
20	-11.22 %	-5.67 %	-11.31 %
25	-9.02 %	-4.78 %	-9.24 %
30	-7.07 %	-1.84 %	-8.27 %
35	-5.89 %	-0.96 %	-6.64 %
40	-5.11 %	-0.18 %	-4.76 %
45	-3.24 %	0.02 %	-4.04 %
50	-3.04 %	0 %	-2.38 %
Mean 05-50	**-11.32 %**	**-4.62 %**	**-7.55 %**
Mean 15-50	**-7.41 %**	**-2.66 %**	**-7.385 %**

As we have reported, both Standard Deviation and Entropy are able to perform with very similar results when using feature subsets of around 20 descriptors. For that reason, the procedural advantages that Entropy provides cause that it may probably be the best candidate for selecting the most descriptive features.

Finally, we should mention that by only analyzing the quantitative results of this experiment is not enough to determine its validity. In this sense, we should review the particular features that have been selected for each artist and, in the same way as in the previous experiment, the huge amount of combinations does not make it feasible to review all of them here. Again, in Section 4.3 a detailed analysis of the features obtained in a particular case will be developed, which we expect that will allow us to have a better insight of the goodness of this kind of analysis.

3.4 Conclusion

In this chapter we have reported the methodological steps developed for addressing our goals. The first step has been to construct a reliable music collection specially addressed for authorship attribution and stylistic analysis of Electronic Music. Later, we have extracted several audio features from different sources covering a wide range of musical facets. Finally, we have designed and developed a series of experiments, that have shown promising results despite the issues that we have reported. The extracted features are able to reach satisfactory accuracies in an Artist Identification task, and the methodologies designed to

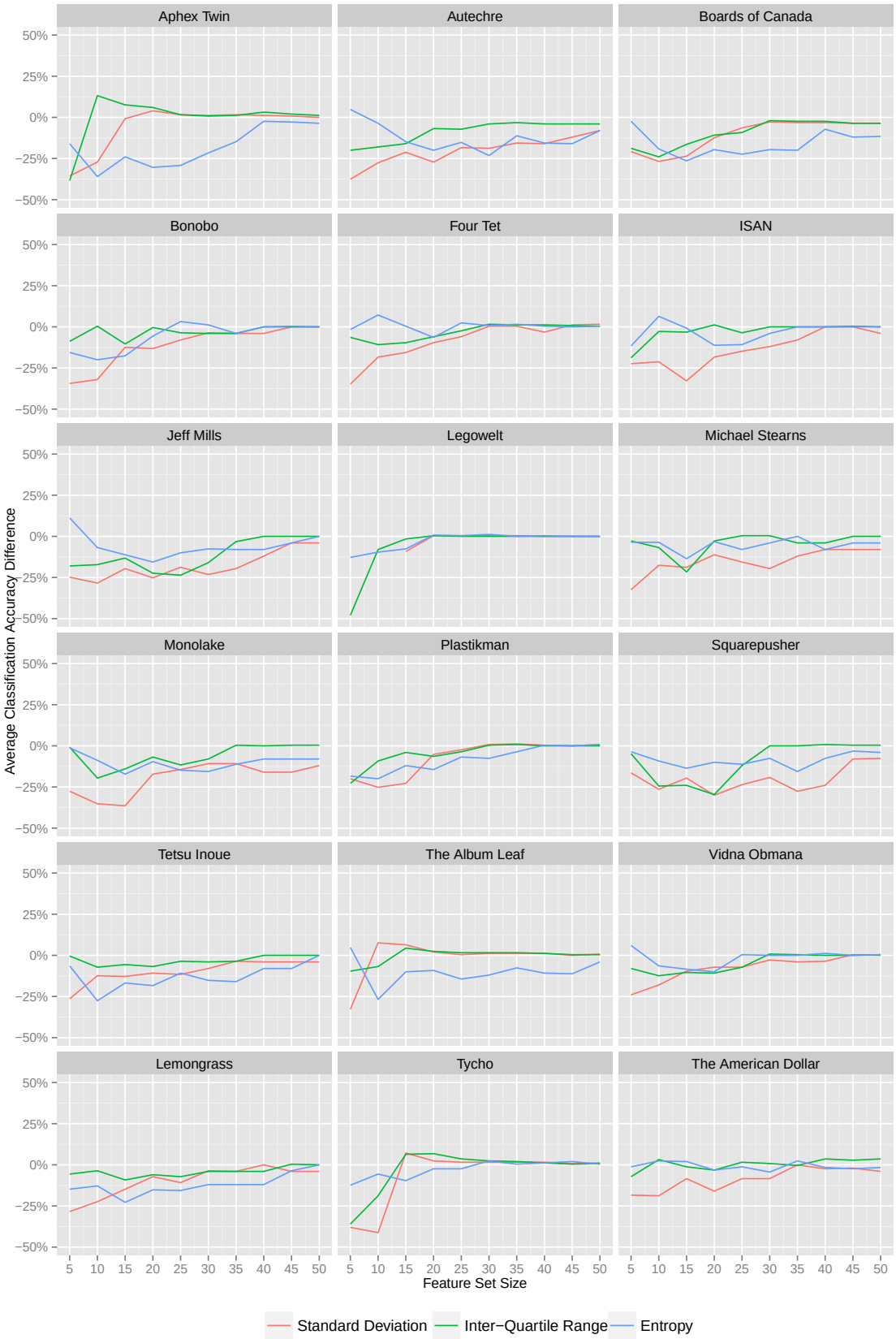

Figure 3.6: Effect of artists' descriptive features set size (X-axis) and candidate selection method (legend) on the album identification accuracy

capture the discriminative and descriptive attributes seem to be performing as expected. However, in our opinion, a closer analysis of the results obtained in one particular dataset is required in order to determine the true potential of the proposed experiments. This will be performed in the following chapter.

Chapter 4 | Study Case: Analysis of the Post-Rock Ambient Dataset

In this chapter we will try to analyze the results obtained in the previously reported experiments and link them with some musical interpretation. For that purpose we have decided to focus on one single dataset, as this will allow us to get much more detailed conclusions on the performance of our methodology. However, we should keep in mind that those conclusions cannot be generalized for the whole collection, as they will take into consideration characteristics of the music that are present on the works by the artists included in the analyzed dataset but not necessarily on the rest of Electronic Music artists.

We have decided to focus our analysis on the "Post-Rock Ambient" dataset, not only because of a previous familiarity with most of the works that are included there, but also because of the preliminary results obtained when performing the experiments. In this sense, as we will see later, the confusions that appeared during the Artist Identification task, seem to be quite balanced between the different artists, fact that does not happen in other datasets. This suggests that there is some degree of homogeneity within the dataset, even if the discrimination between the styles of the different artists is feasible.

We will first examine the results obtained in the Artist Identification experiment, trying to determine if the confusions generated by the prediction algorithm can be explained according to the content of those tracks. Moreover, we will also list the most "discriminative" and "descriptive" features for each artist, obtained using the methodology described in Chapter 3. Those features will be analyzed both globally and individually in order to check if they are capturing the traits that best represent the style of each of the considered artists. Finally, we will determine if any feature appears in both sets, forming what we called the "characteristic" features of an artist. In that case, we will try to understand the reason why those features are specially relevant in the works of that artist.

4.1 Artist Identification

As it has been already explained in Section 3.3.1, we have performed an Artist Identification experiment using three albums per artist for the construction of the Support Vector

Machines model, and the remaining two albums per artist were kept apart for the evaluation step. In the case of the "Post-Rock Ambient" dataset, the albums included in the training and test sets are shown in Table 4.1.

Table 4.1: Training and Test sets used for the Artist Identification task with the Post-Rock Ambient dataset

Training Set

Artist	**Year**	**Album**
Boards of Canada	1995	Twoism
	1998	Music Has the Right to Children
	2005	The Campfire Headphase
The Album Leaf	1999	An Orchestrated Rise to Fall
	2004	In a Safe Place
	2007	The Enchanted Hill
The American Dollar	2006	The American Dollar
	2008	A Memory Stream
	2012	Awake in the City

Test Set

Artist	**Year**	**Album**
Boards of Canada	1996	Hi-Scores
	2002	Geogaddi
The Album Leaf	2001	One Day I'll Be On Time
	2006	Into the Blue Again
The American Dollar	2007	The Technology Sleep
	2010	Atlas

4.1.1 Classification Results

Using the audio features extracted from the excerpts included in the training set, we constructed a SVM model that was afterwards used to predict the artist of the excerpts that comprised the test set. Fig. 4.1 represents the confusion matrix obtained, summarizing the amount of excerpts from each artist that were predicted as belonging to one artist or another. As it can be seen in the matrix, the majority of the excerpts were successfully classified, but some of them were predicted incorrectly. Those cases will be considered individually in Section 4.1.2.

The performed predictions reach an overall classification accuracy of 70 %, However, due to the fact that the number of instances per class is small -ten excerpts per artist-, the interval of accuracies at a 95 % of confidence is quite big. Nevertheless, the lowest value (50.6 %) is significantly better than the random baseline for a three-class classification task (33.33 %) with over a 99.99 % of confidence, so we can be pretty sure that the

extracted features are able to capture the traits that help to distinguish between the works of those artists. Overall precision, recall and F1-score, common measures to determine the goodness of classification tasks, have similar values around 70 %, fact that suggests a balance between the predictions for the different classes. Finally, the kappa value is around 0.55, which indicates a moderate agreement between the real and predicted labels, reinforcing the idea that the results were obtained due to the particular characteristics of the works of each artist and not by mere chance.

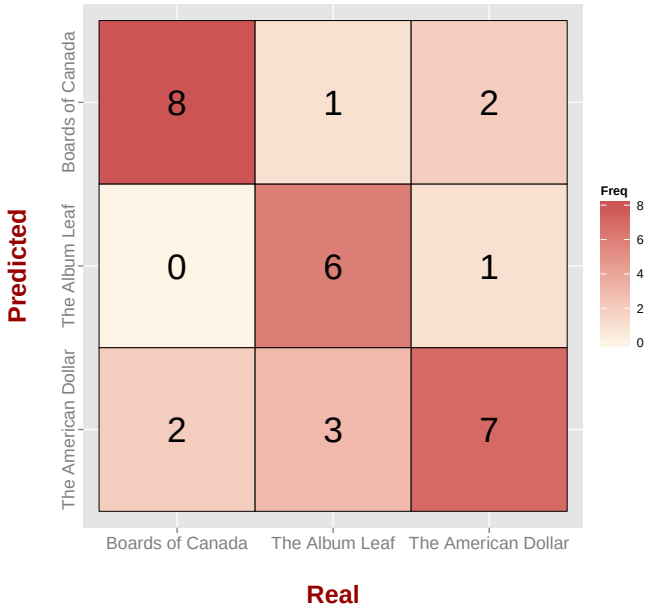

Figure 4.1: Artist Identification confusion matrix of the "Post-Rock Ambient" dataset

A more detailed performance analysis can be found in Table 4.2, where several statistics are listed for each of the three considered artists. In our specific analysis, the most relevant conclusion that can be retrieved from the table is that "The American Dollar" is the class with the lowest precision value. This means that works from other artists are frequently labeled as belonging to this one, as it is also captured by the detection prevalence measure. This, combined with the fact that the recall for that class is decently high, suggests that the works by "The American Dollar" are more representative of the overall characteristics of the dataset than those created by "Boards of Canada" or "The Album Leaf". In other words, those results seem to indicate that the music created by "The American Dollar" is less personal, in the sense that it seems to follow rules that are generic in the sub-genre to which they belong.

4.1.2 Confusions

By examining individually each of the excerpts in which the predictor failed to determine the correct artist we may increase our understanding of whether the system is able to capture the representative stylistic traits of the artists or not.

Table 4.2: By-Class performance statistics of the Artist Identification task using the "Post-Rock Ambient" dataset

	Boards of Canada	The Album Leaf	The American Dollar
Sensitivity (Recall)	0.800	0.600	0.700
Specificity	0.850	0.950	0.750
Pos Pred Value (Precision)	0.727	0.857	0.583
Neg Pred Value	0.895	0.826	0.833
Prevalence	0.333	0.333	0.333
Detection Rate	0.267	0.200	0.233
Detection Prevalence	0.367	0.233	0.400
Balanced Accuracy	0.825	0.775	0.725
F1 Score	0.762	0.706	0.636

Boards of Canada - Dawn Chorus : Labeled as "The American Dollar"

One of the main traits of this track that may distinguish it from the usual style of "Boards of Canada" is the fact that it seems to be much more dense in terms of timbre. Even though it does not sound like a typical piece by "The American Dollar", this aspect makes it a somewhat reasonable choice. On the other hand, the clear rhythmic pattern, the presence of high notes, and the dirtyness of some sounds that are present in the track, makes it difficult to understand why the system has not succeeded in predicting the real artist.

Boards of Canada - The Beach at Redpoint : Labeled as "The American Dollar"

Similarly to what happened with "Dawn Chorus", the timbric density is most likely the reason why "The American Dollar" is predicted instead of "Boards of Canada" for this track. However, in our opinion, we don't think that a trained listener would have done this mistake.

The Album Leaf - Shine : Labeled as "The American Dollar"

This track, even though it keeps some of the most characteristic traits of the typical "The Album Leaf" style, could be easily confused with a creation made by "The American Dollar". The selection of the instrumentation, the rhythmic patterns, the progressive introduction of timbres, and a very prevalent pad are traits that are usually found in the works by "The American Dollar".

The Album Leaf - Wishful Thinking : Labeled as "The American Dollar"

The most relevant characteristic that a listener perceives in this track is that it sounds completely acoustic. It is for sure more close to New Age than to Electronic Music. However, this does not explain why the system considers it to be created by "The American Dollar". The only suitable explanation may be that the string pad progressively fattens the sound, but it's very unlikely that a human listener may have done this mistake because of that.

The Album Leaf - The Audio Pool : Labeled as "Boards of Canada"

Percussion has undeniably more importance in this track than it normally does in the works by "The Album Leaf". For that reason, we think that it may be possible for a trained listened to confuse it with "Boards of Canada", as this artist has, probably, the most aggressive tracks in the dataset.

The Album Leaf - Asleep : Labeled as "The American Dollar"

Similarly to what happened with "Shine", the instrumentation used in this track, as well as the type of percussion and the way in which the timbral structure is built clearly recalls the style of "The American Dollar". In this case, thus, we consider that the mistake is quite reasonable.

The American Dollar - Frontier Melt : Labeled as "The Album Leaf"

The fact that the system has labeled wrongly this track is quite surprising, as it seems to be a prototypical piece of the style of "The American Dollar". It is very unlikely that a trained listener could have not predicted correctly the artist.

The American Dollar - Palestine : Labeled as "Boards of Canada"

It is easy to perceive that this track is very different to the typical works created by "The American Dollar". The amount of timbral layers is much lesser, with no percussion at all, and the tempo is slower, creating a deep sensation of calm. And, in the same way that we previously said that "Boards of Canada" had the most aggressive pieces of the dataset, it is also true that many of the most relaxed tracks are also created by them, so the confusion does not seem unreasonable at all.

The American Dollar - Raided by Waves : Labeled as "Boards of Canada"

In this track we get exactly the same impression that we did when listening to "Frontier Melt", as the most characteristic stylistic traits of "The American Dollar" seem to be very easy to identify in it.

As we have seen, very few of the mistaken labels of this experiment can be explained solely using the most evident stylistic traits of each of the considered artists. And this would be even more difficult as the number of artists increased. For that reason, we think that analyzing in detail which are the features that better help to discriminate between the works of the different artists is very important in order to get closer to the stylistic analysis of those artists.

4.2 Discriminative Features

As we have already explained in Section 3.3.2, we consider that one of the most important steps when performing a computational stylistic analysis of the works of some artists is determining which are the audio features that allow us to better discriminate between them. Of all the different feature selection techniques that were considered previously, we

have decided to only analyze in detail the output obtained by means of the Information Gain method. The main reason for this delimitation is to allow a much more detailed review of the specific audio features that were selected, while ensuring that they provide a high discriminative power.

The lists of the twenty most discriminative features for each artist according to the Information Gain feature selection method can be found in tables 4.3, 4.4 and 4.5. In the following paragraphs we will comment each of them individually.

The Album Leaf

Probably the most relevant conclusion that can be extracted from the analysis of the audio features listed in Table 4.3 as being the most discriminative in the works by "The Album Leaf", is that most of them belong to the "Pitch and Tonal" category. This could be interpreted in the sense that the importance of melodic and harmonic content in the works by "The Album Leaf" is much higher than in the tracks by the other considered artists. This conclusion is in concordance with the impression that the listening of the pieces of the dataset provides.

Table 4.3: Most discriminative Audio Features for "The Album Leaf" within the "Post-Rock Ambient" Dataset using Information Gain as selection method

Library	Audio Feature	Coefficient	Statistic
Echonest	Segment Pitch	1	Kurtosis
Echonest	Segment Pitch	4	Mean
Echonest	Segment Pitch	4	Median
Echonest	Segment Pitch	4	Kurtosis
MIRToolbox	Key Strength	2	
MIRToolbox	Key Strength	3	
MIRToolbox	Key Strength	8	
MIRToolbox	Key Strength	9	
MIRToolbox	Chroma	11	
MIRToolbox	Tonal Centroid	1	
Essentia	HPCP	16	Mean
Essentia	HPCP	18	Variance
Essentia	HPCP	26	Mean
Essentia	HPCP	26	Variance
Essentia	Pitch Salience		Variance
Essentia	Zero Crossing Rate		Variance
MIRToolbox	MFCC	3	
MIRToolbox	Irregularity		
Echonest	Segment Loudness Max Position		Mean
Essentia	Beats Loudness Band Ratio	2	Mean

Echonest's Segment Pitch seems to be one of the most indicative descriptors to represent this artist's music. More precisely, it seems that a stylistic trait of "The Album Leaf" is to, intentionally or not, avoid the usage of D# on its works. This

can be observed in Fig. 4.2, where the distribution of the different statistics related to this audio feature for "The Album Leaf" are compared to those obtained for the other artists. However, while the distributions of both the mean and the median of this coefficient are clearly different to those from the other artists, it is not so clear for its skewness, even though it is also selected according to the Information Gain that it provides. The same happens to the kurtosis of the fist coefficient -the one corresponding to C note-.

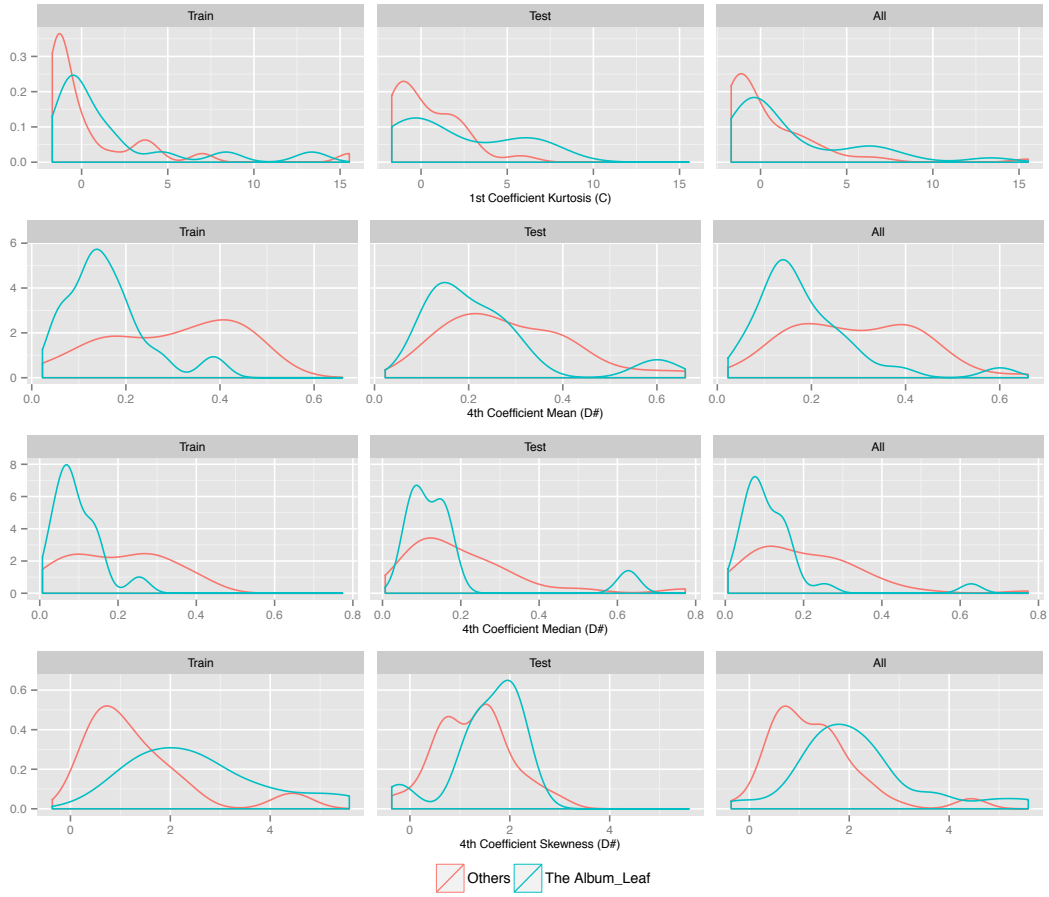

Figure 4.2: Discriminative Features of "The Album Leaf" [1/4]: Probability density functions of the Kurtosis of the 1st coefficient of the Segment Pitch descriptor and the Mean, Median and Kurtosis of the 4th coefficient of the same descriptor of the works by "The Album Leaf" according to the Echonest Analyzer, compared with the other two artists of the "Post-Rock Ambient" dataset

The tendency to avoid the inclusion of D# notes can also be observed by means of the keys that are detected or rejected when computing MIRToolbox's Key Strength. As we can see in Fig. 4.3, G# is a tonality that is clearly rejected much more frequently in the works by "The Album Leaf" than in the other two artists, and the note D# corresponds to the dominant of that tonality. On the other hand, D key appears with greater probability. This seems to have sense, as the minor second is rarely found in tonal music. C# and G are the other two keys the strength of which is considered to

be discriminative by the system, the former for being rejected usually, while the latter is supposed to be used quite often. Nevertheless, this seems to happen only within the training set, so this conclusion cannot be generalized to the complete collection of works by this artist.

Essentia also provides us a series of tonality-related descriptors that the system considers to be discriminative of "The Album Leaf", as it is shown in Fig. 4.4. Some of the Harmonic Pitch Class Profiles coefficients seem to be particularly relevant. The 26th seems to have a lesser weight in the works by this artists than in the rest, and with more homogeneity throughout each track. In the case of the 16th coefficient, it looks like it presents a bi-modal distribution, while the 18th coefficient is much more homogeneous in the works by "The Album Leaf". The last Essentia descriptor related to this musical facet that is listed as belonging to the 20 more discriminative ones is the Pich Salience Variance, which is clearly more concentrated around a small interval in the works by the considered artist.

Even though most of the audio features that are listed as being discriminative for the works by "The Album Leaf" are related to pitch and tonality, there are some others that should be included in other categories. Those ones are represented in Fig. 4.5. However, from those five it seems that only MIRToolbox's Irregularity has an evident difference in the probability distributions for both training and test sets. According to the MIRToolbox manual, this descriptor captures the "*degree of variation of the successive peaks of the spectrum*". So, by looking at the graph, we could deduce that the works by "The Album Leaf" present a lesser degree of variation in the peaks of the spectrum, circumstance that could be associated with softer sounds.

Finally, it is worth mentioning that when using another selection of works for this artist, some of the discriminative features that we have mentioned were different. This happened when we developed the same experiment without filtering out songs with vocals. Nevertheless, the overall flavor is kept, as the proportion of tonal descriptors did not change. In fact, the relevance of D# was even higher, as both variance and kurtosis of the 4th Coefficient of the Echonest's Segment Pitch descriptor and the 4th of MIRToolbox's Key Strength were also included. So we may conclude that although not mantaining exactly the same descriptors, this experiment is able to capture the general stylistic traits of "The Album Leaf".

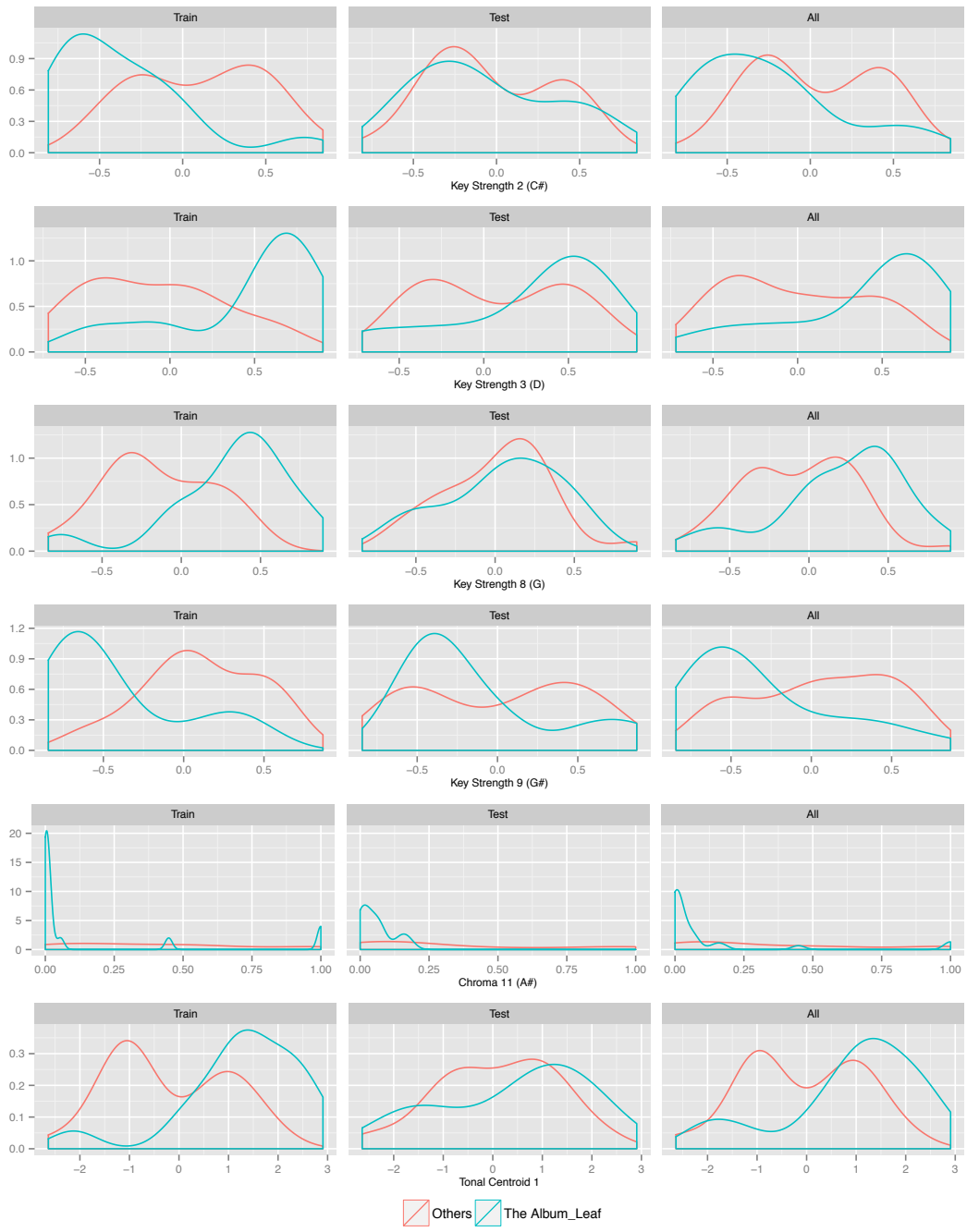

Figure 4.3: Discriminative Features of "The Album Leaf" [2/4]: Probability density functions of the 2nd, 3rd 8th and 9th coefficients of the Key Strength descriptor (C#, D, G, G#), the 11th coefficient of the Chroma vector and the 1st coefficient of the Tonal Centroid descriptor of the works by "The Album Leaf" according to the MIRToolbox library, compared with the other two artists of the "Post-Rock Ambient" dataset

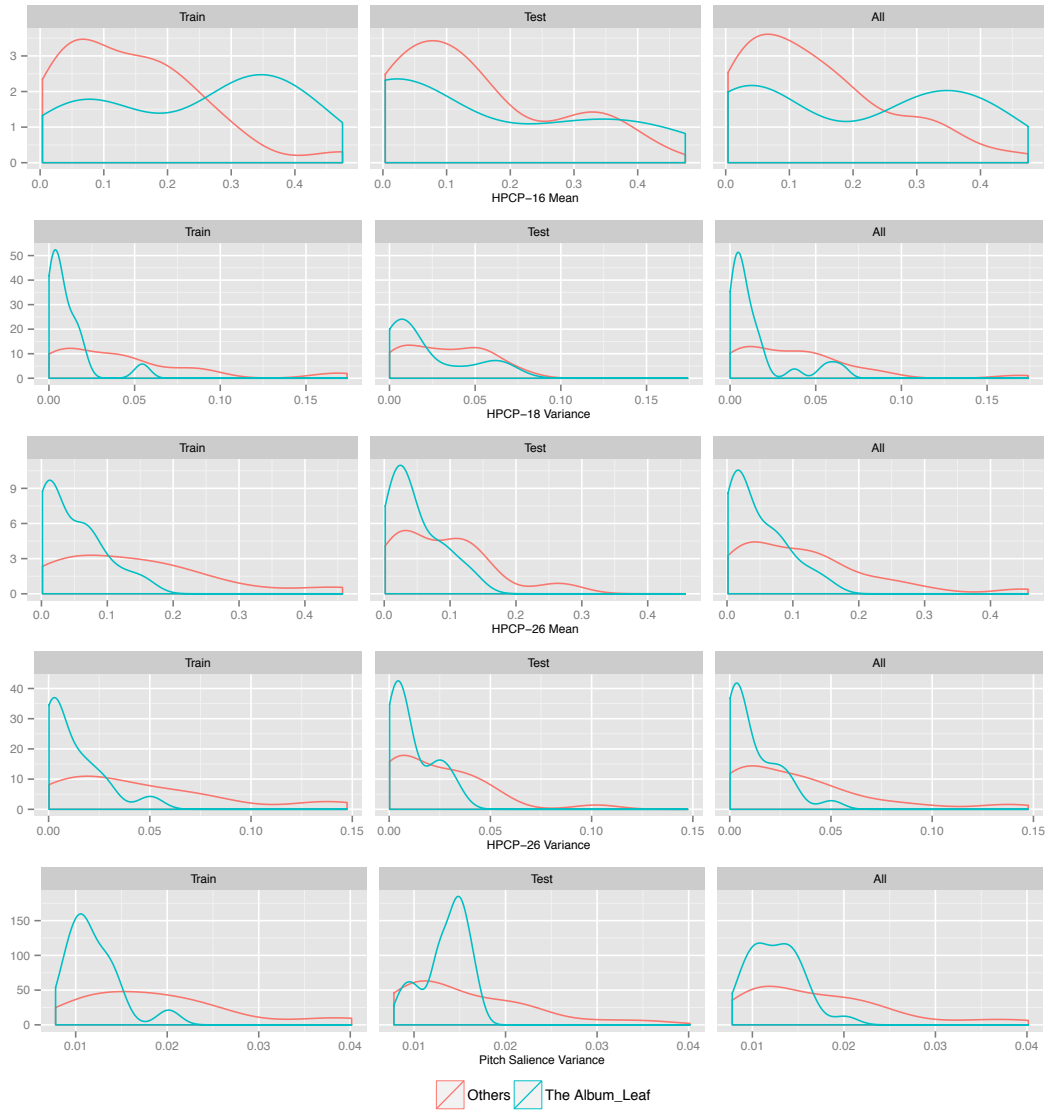

Figure 4.4: Discriminative Features of "The Album Leaf" [3/4]: Probability density functions of the Mean of the 16th and 26th HPCP, the Variance of the 18th and the 26th HPCP and the Pitch Salience Variance of the works by "The Album Leaf" according to the Essentia library, compared with the other two artists of the "Post-Rock Ambient" dataset

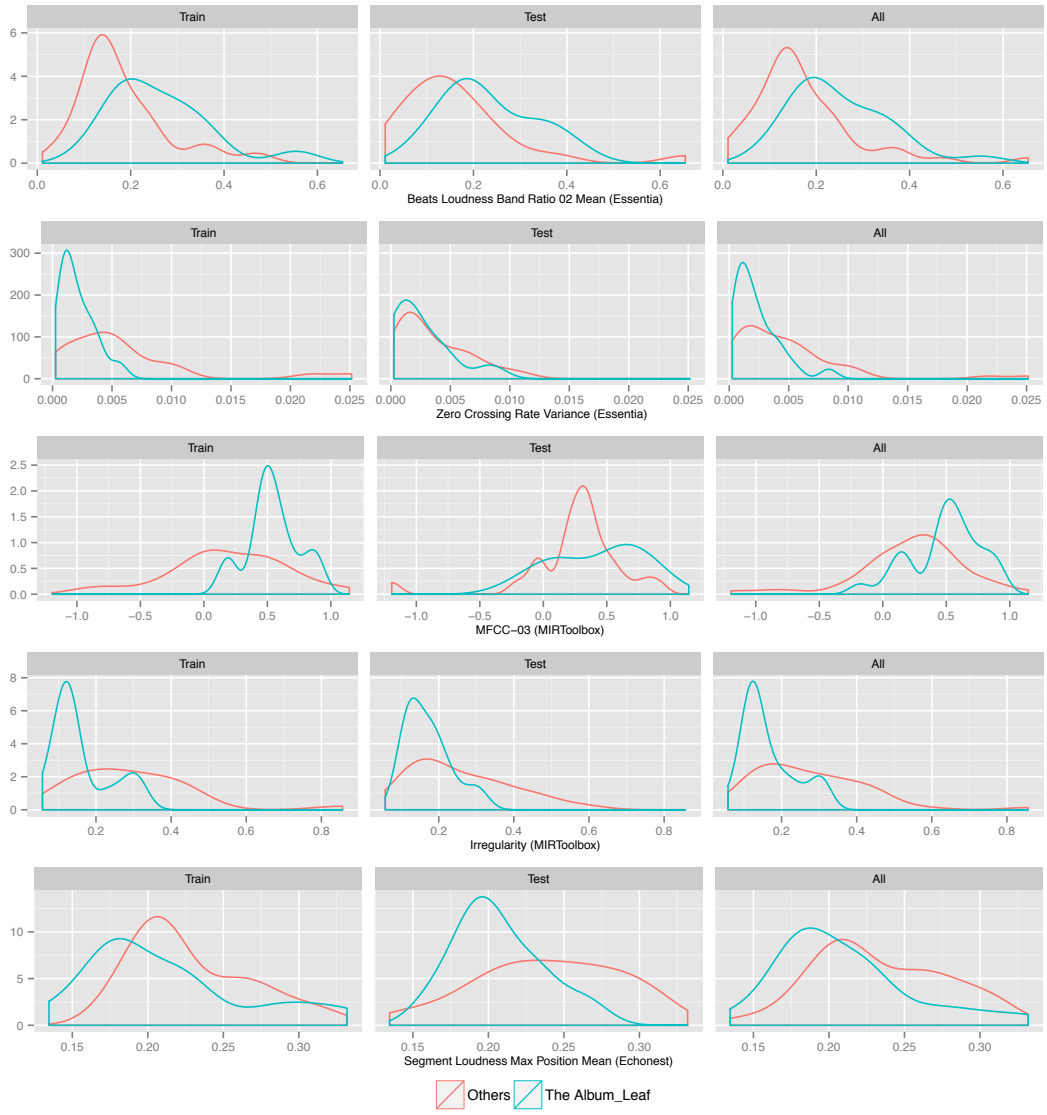

Figure 4.5: Discriminative Features of "The Album Leaf" [4/4]: Probability density functions of some non-pitch-and-tonality-related descriptors (Essentia's Zero Crossing Rate Variance and Beats Loudness Band Ratio 02 Mean, MIRToolbox's 3rd coefficient of the MFCC and Irregularity, and Echonest's Segment Loudness Max Position Mean) of the works by "The Album Leaf", compared with the other two artists of the "Post-Rock Ambient" dataset

The American Dollar

In the case of "The American Dollar", as opposed to what we saw before, there is no specific musical facet that is as dominant to the rest as it was tonality for "The Album Leaf", as it is shown in Table 4.4. In fact, for this artist, the only discriminative audio feature that is tonality-related is Essentia's Dissonance. The rest represent timbre, rhythm and structure aspects of the music, with an astonishingly high amount of descriptors directly or indirectly extracted from the Echonest Analyzer.

Table 4.4: Most discriminative Audio Features for "The American Dollar" within the "Post-Rock Ambient" Dataset using Information Gain as selection method

Library	Audio Feature	Coefficient	Statistic
Echonest	Segment Timbre	3	Mean
Echonest	Segment Timbre	3	Median
Echonest	Segment Timbre	4	Variance
Echonest	Segment Timbre	6	Variance
Echonest	Segment Timbre	11	Variance
Echonest	Section Duration		Mean
Echonest	Section Rate		
Echonest	Segments per Section		Variance
Echonest	Segment Confidence		Mean
Echonest	Segment Confidence		Median
Echonest	Segment Confidence		Variance
Echonest	Beats Confidence		Mean
Echonest	Beats Confidence		Median
Echonest	Tatums Confidence		Mean
Echonest	Tatums Confidence		Median
Echonest	Segment Loudness Start		Mean
Echonest	Segment Loudness Start		Median
Essentia	Spectral Complexity		Mean
Essentia	Dissonance		Mean
Essentia	MFCC	3	Mean

The timbral coefficients computed by the Echonest Analyzer seem to be particularly relevant for distinguishing the works by "The American Dollar" to those by the other artists of this dataset. This makes sense if we take into consideration that the instrumentation used by this artist is quite homogeneus in most of his pieces, making it probably the easiest cue for humans to identificate its works. As it is represented in Fig. 4.6, the probability densities of the Mean and Median of the 3rd coefficient are very distinctive in the train set. However, the test set does not mantain such degree of distinctiveness in those two descriptors. On the other hand, the Variances of the 4th, 6th and 11th coefficients certainly are able to keep a high disctintiveness capability both in the training and test sets, having probability density functions much more concentrated arond lower values. This seems to indicate that throughout

each the works by this artist, timbres change less than in the works of "The Album Leaf" and "Boards of Canada". Unfortunately, very little documentation is provided by the developers of the API, so we cannot fully link each of the coefficients to a more precise musical property.

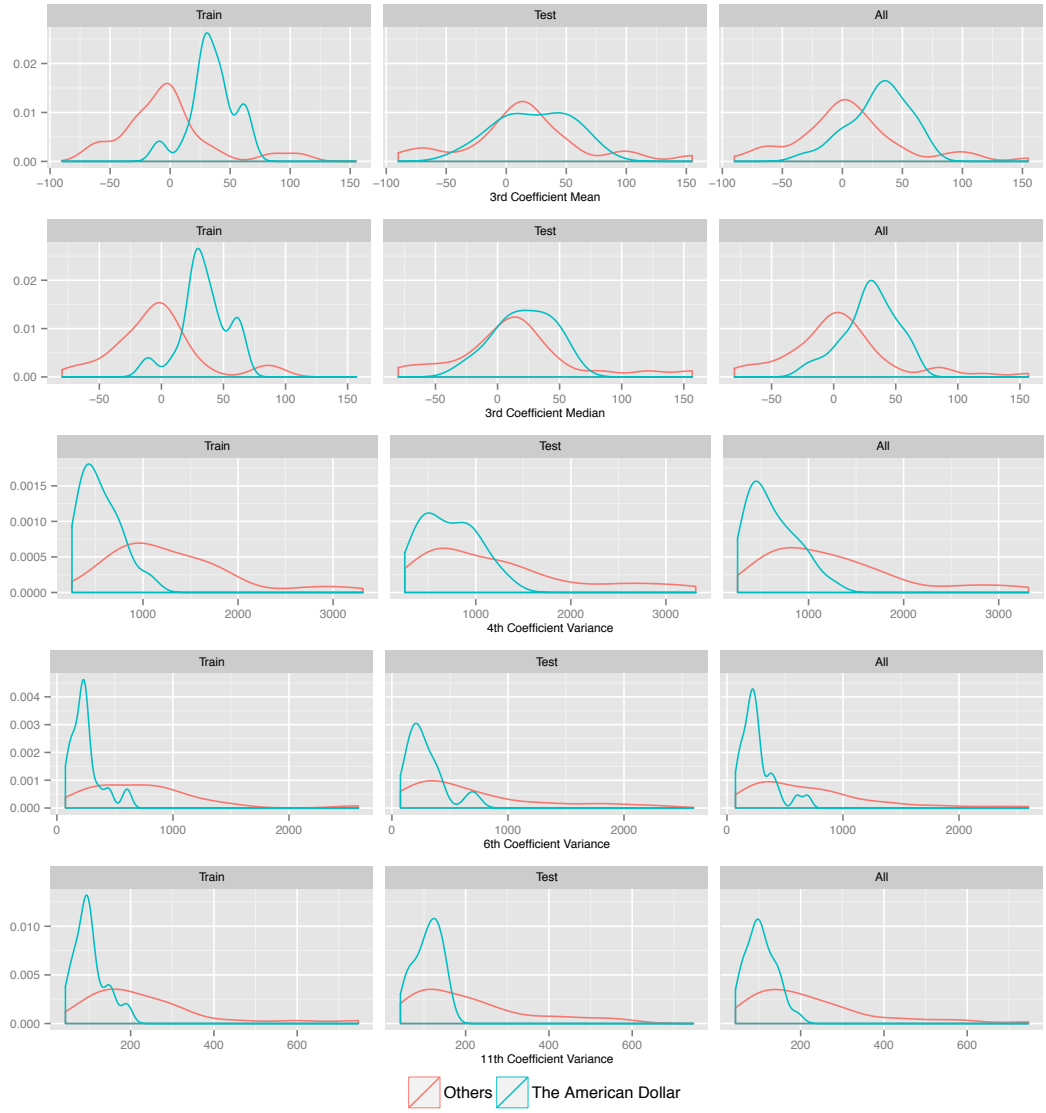

Figure 4.6: Discriminative Features of "The American Dollar" [1/4]: Probability density functions of some statistics of the 3rd, 4th, 6th and 11th coefficient of the Segment Timbre descriptor of the works by "The American Dollar" according to the Echonest Analyzer, compared with the other two artists of the "Post-Rock Ambient" dataset

Echonest's structural analysis, as well as the derived descriptors that we have computed from that analysis, has an undeniably high discriminative power for "The American Dollar", as it can be seen in Fig. 4.7. The probability density functions seem to coincide with our perceptual impression that this artist tends to progressively modify the textures in its track by adding instrumental layers one over the other. This leads to shorter segments and sections that change at a more or less constant

rate. However, those changes are more subtle than in the works of the other artists, circumstance that causes the confidence of each of the segments to be clearly lower.

This lower confidence can also be found in the rhythmic analysis, as it is shown in Fig. 4.8. Even though it doesn't seem to be particularly difficult for a human listener to follow the beat in this works, the extractor considers it a hard task, with confidences around 20 %. As always, the impossibility to access to a complete description of how this rhythmic divisions have been determined does not allow us to fully understand the reason why this happens. Having said that, we find it reasonable that the confidences are lower than for "The Album Leaf" and "Boards of Canada", as the percussive onsets are in some way hidden behind other instrumental layers.

We should also point out that, as the reader may have already noticed by observing the graphs in ig. 4.8, the probability density functions of the Beats and Tatums Confidence Mean and Median look exactly the same. In fact, their values are exactly the same for every single work analyzed so far. The most probable explanation for this situation is that the confidence of each individual beat is computed as the mean of the confidences of the tatums that are included in it. In this way, the averaging of the confidences will always be identical. Consequently, computing them both is redundant, fact that should be taken into account in future analyses.

Finally, in Fig. 4.9 we can find the probability density functions of the remaining features that have been selected as discriminative. As we can see, in average, the segments of the tracks by "The American Dollar" start with a higher loudness. Taking into account that the segmentation performed for those works has led to shorter fragments, probably smaller than complete musical phrases, it has sense that they start with much more energy. Furthermore, both Spectral Complexity and Dissonance are likely to be related to the higher amount of instrumental layers that this artist seems to usually employ. Last but not least, in the same way that happened in the analysis of "The Album Leaf", the 3rd coefficient of the MFCCs has been selected as discriminative. However, in this case it has been the one computed by Essentia and not MIRToolbox, as it happened before. It seems that not only are the coefficients provided by each library computed in different scale, but also that they have no correlation at all. It would be interesting to analyze this quite surprising fact in the future.

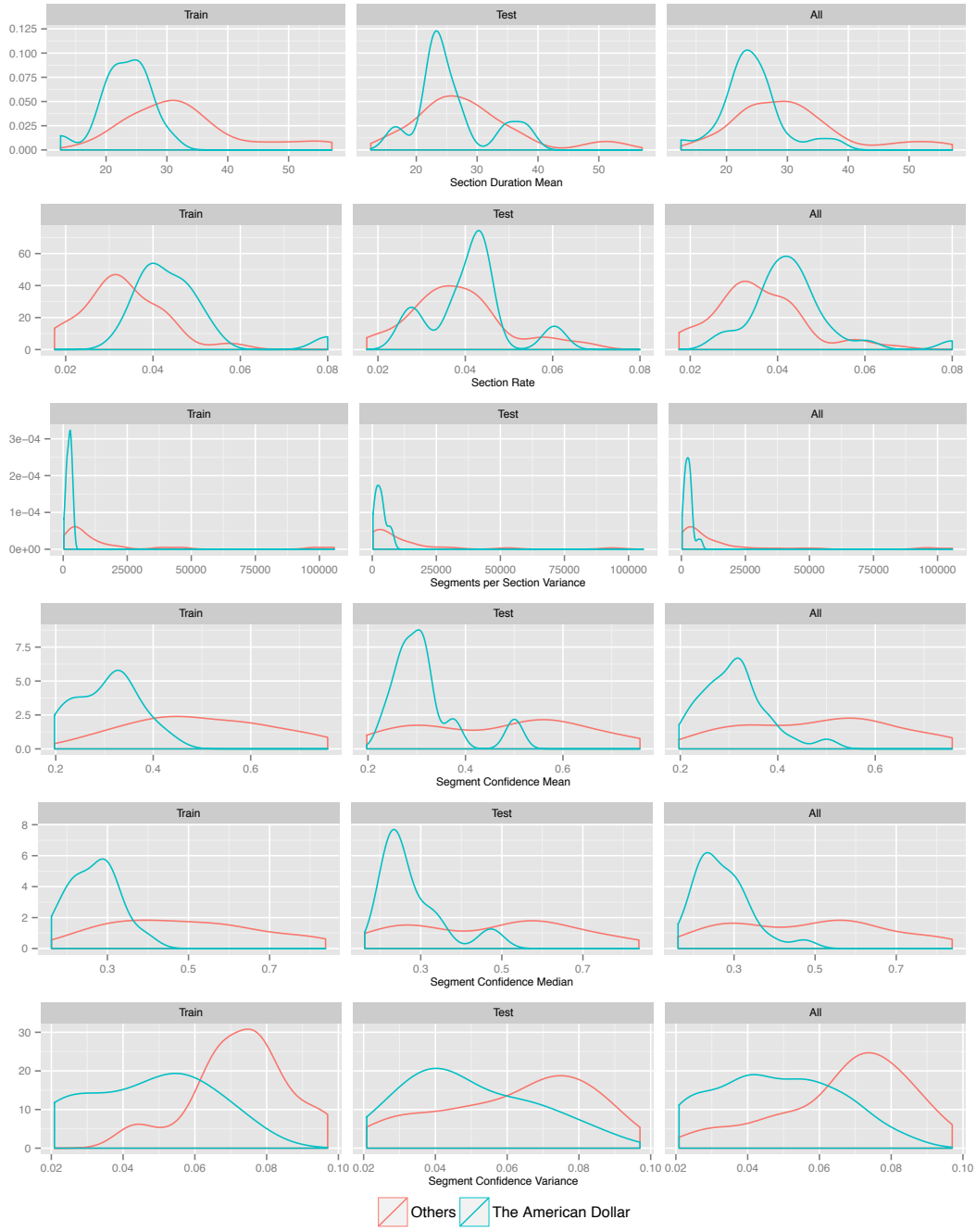

Figure 4.7: Discriminative Features of "The American Dollar" [2/4]: Probability density functions of some structural analysis features of the works by "The American Dollar" according to the Echonest Analyzer, compared with the other two artists of the "Post-Rock Ambient" dataset

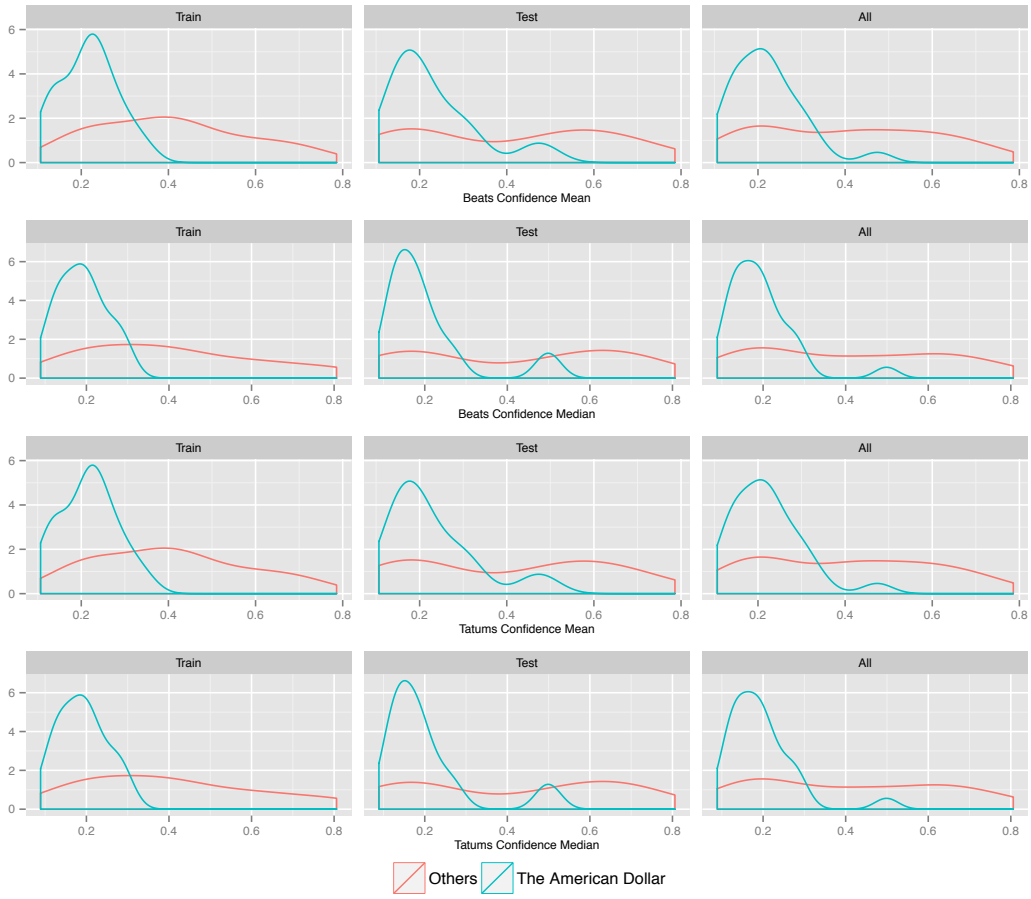

Figure 4.8: Discriminative Features of "The American Dollar" [3/4]: Probability density functions of some rhythmic structure analysis features of the works by "The American Dollar" according to the Echonest Analyzer, compared with the other two artists of the "Post-Rock Ambient" dataset

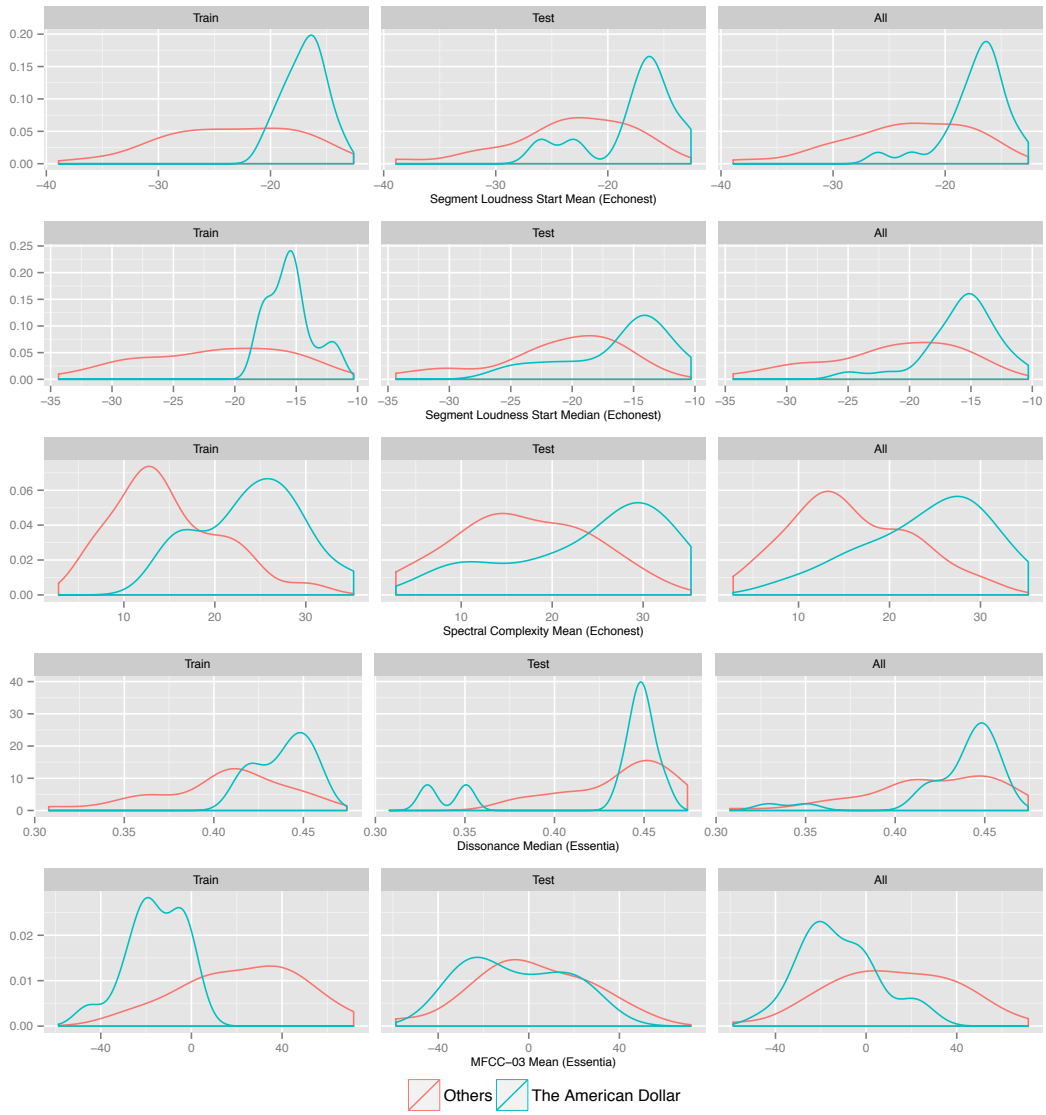

Figure 4.9: Discriminative Features of "The American Dollar" [4/4]:Probability density functions of some audio descriptors (Echonest's Segment Loundess Start Mean and Median, and Essentia's Spectral Complexity Mean, Dissonance Median and the 3rd Coefficient of the MFCCs) of the works by "The American Dollar", compared with the other two artists of the "Post-Rock Ambient" dataset

Boards of Canada

Similarly to what happened with "The American Dollar", the selection of the most discriminative features for "Boards of Canada" according to their Information Gain leads to a feature subset that is not dominated by any particular musical facet. As it is summarized in Table 4.5, timbric, rhythmic, structural, and even tonal features appear in the list as being the most useful to distinguish the works by this artist to those created by the two others included in the considered dataset.

Table 4.5: Most discriminative Audio Features for "Boards of Canada" within the "Post-Rock Ambient" Dataset using Information Gain as selection method

Library	Audio Feature	Coefficient	Statistic
MIRToolbox	Roughness		Mean
MIRToolbox	Roughness		Median
MIRToolbox	Roughness		Period Freq
Essentia	Spectral Contrast	3	Mean
Essentia	Spectral Contrast	5	Mean
MIRToolbox	Irregularity		
Echonest	Danceability		
Essentia	Danceability		
Echonest	Beats Confidence		Mean
Echonest	Beats Confidence		Median
Echonest	Beats Confidence		Skewness
Echonest	Tatums Confidence		Skewness
Essentia	Beats Loudness Band Ratio	6	Variance
Echonest	Segment Confidence		Mean
Echonest	Segment Confidence		Median
Echonest	Segment Confidence		Skewness
Echonest	Segment Loudness Start		Median
Essentia	Pitch Mean		Mean
Essentia	HPCP	26	Variance
MIRToolbox	HCDF		Median

Figure 4.10 shows the probability density functions of the audio features that are related to the timbral aspect of the tracks. As it can be seen, MIRToolbox's Roughness seems to be a very relevant feature to distinguish the works by this artist, particularly in the training set. However, it is quite surprising that the average values of this descriptor are clearly lower in "Boards of Canada" tracks compared to the ones that are obtained for the other two artists. This seems to be counterintuitive, as a decent number of tracks by the considered artist contains clearly noisy sounds. For example, the track ''Basefree" included in the album "Twoism" includes a lot of harshy elements in it. Surprisingly enough the mean and median of the Roughness descriptor computed for this piece are far below the average values obtained in the works by "The Album Leaf", an artist which can hardly be considered noisy. Some

tests with basic sounds (sine wave, sawtooth, modulated sine wave and pink noise) have been developed in order to determine if the descriptor was returning closer-to-0 values when roughness was low, and high values when it was high (or, otherwise, had the range inverted), as it happened. For that reason, we cannot provide a fully convincing explanation for this performance. Having said that, our impression is that the small amount of different timbres that are usually found in the tracks of this artist may have something to do with it.

A similar explanation may also be considered for the 3rd and 5th coefficients of the Spectral Contrast. In both cases the boundary that separates the distributions corresponding to "Boards of Canada" and the other two artists is quite clear. The values for the considered artist are lower, but larger in magnitude as we are dealing with negative numbers. Taking that into consideration, and according to the interpretation that is suggested by Akkermans, Serrà and Herrera [2], we could state that those tracks are noisier than the ones by "The Album Leaf" and "The American Dollar" in the 3rd and 5th subbands in which the descriptor divides the spectrum. The higher values that the MIRToolbox's Irregularity is returning do nothing but reaffirm this statement.

With respect to the features related to rhythm, which are shown in Fig. 4.11, and the ones related to structure and tonality, represented in Fig. 4.12, a trait that seems to be common in almost all of the probability density functions is that the curve corresponding to the works by "Boards of Canada" is clearly flatter. Taking into consideration that the comparison is done against the combination of the works by two different artists, this strongly suggest that the variability that can be found in the works by "Boards of Canada" is much higher. As a consequence, we could interpret the graphs in the sense that this artist has a much less defined style than the rest. This fits with the impression that we get when we listen to its music.

As a final remark, it should be noted that even though the means and medians of the confidence of beats and tatums according to the Echonest Analyzer are exactly the same, as we have already mentioned, only the former appears in the list as belonging to the discriminative subset. This only happens because of the size of the subset that we are analyzing, as with a feature set size increased to 22 elements, both beats and tatums confidences are included.

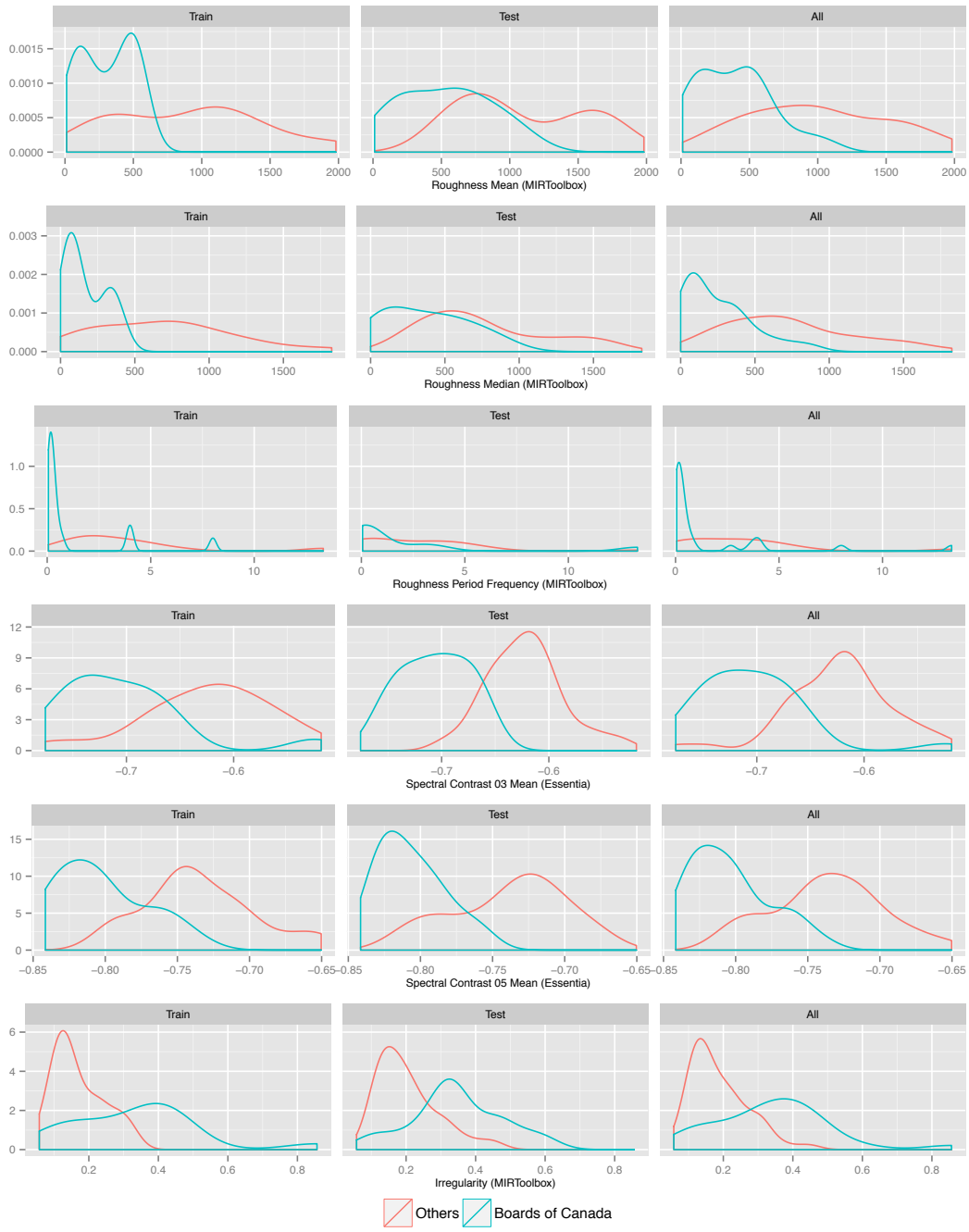

Figure 4.10: Discriminative Features of "Boards of Canada" [1/3]: Probability density functions of some timbre-related audio descriptors (MIRToolbox's Roughness Mean, Median and Period Frequency, Essentia's Spectral Contrast Mean of the 3rd and 6th bands, and MIRToolbox's Irregularity) of the works by "Boards of Canada", compared with the other two artists of the "Post-Rock Ambient" dataset

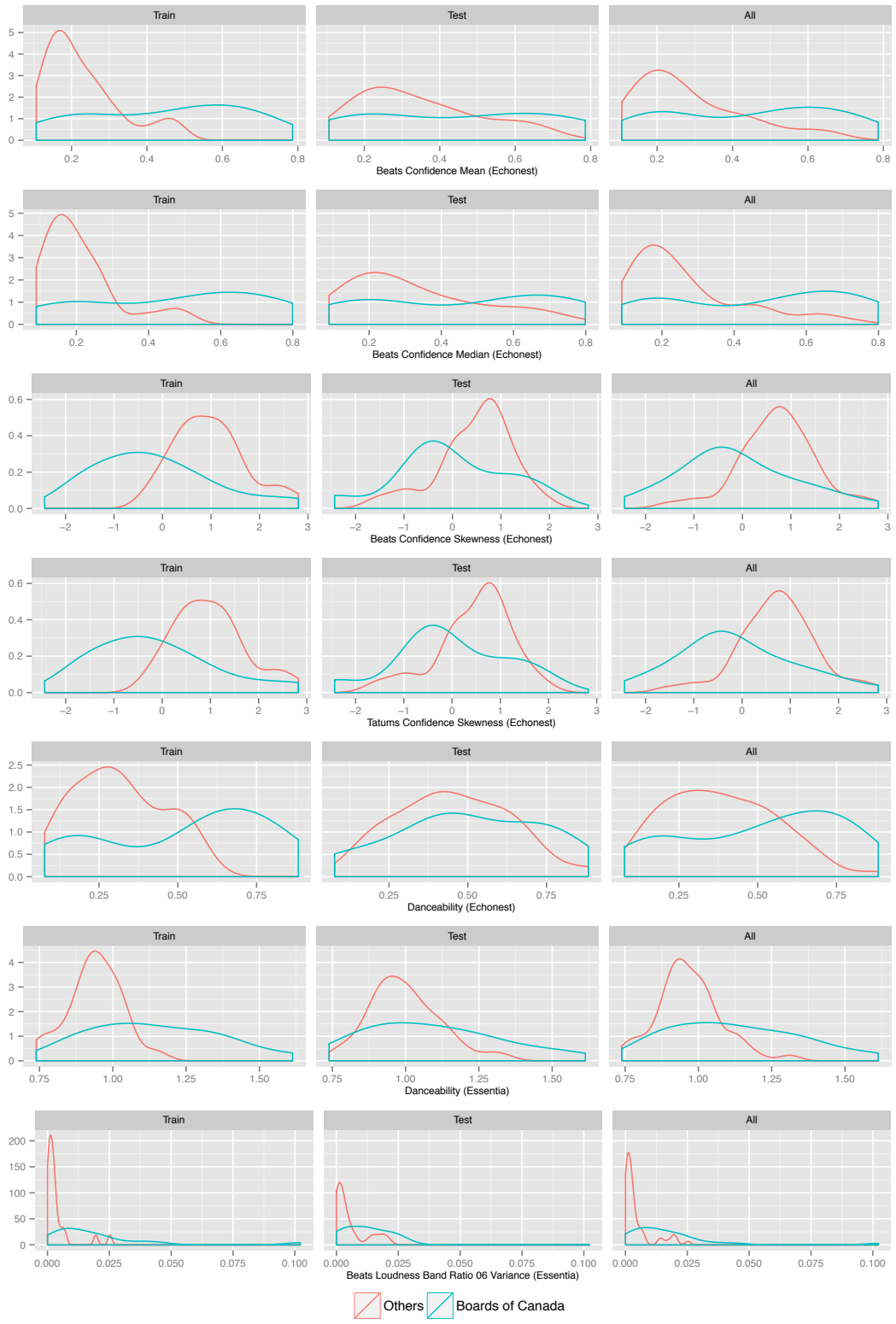

Figure 4.11: Discriminative Features of "Boards of Canada" [2/3]: Probability density functions of some rhythm-related audio descriptors (Echonest's Mean, Median and Skewness of the Beats Confidence measure, and Skewness of the Tatums Confidence measure, Echonest's and Essentia's Danceability, and Essentia's Variance of the 6th coefficient of the Beats Loudness Band Ratio vector) of the works by "Boards of Canada", compared with the other two artists of the "Post-Rock Ambient" dataset

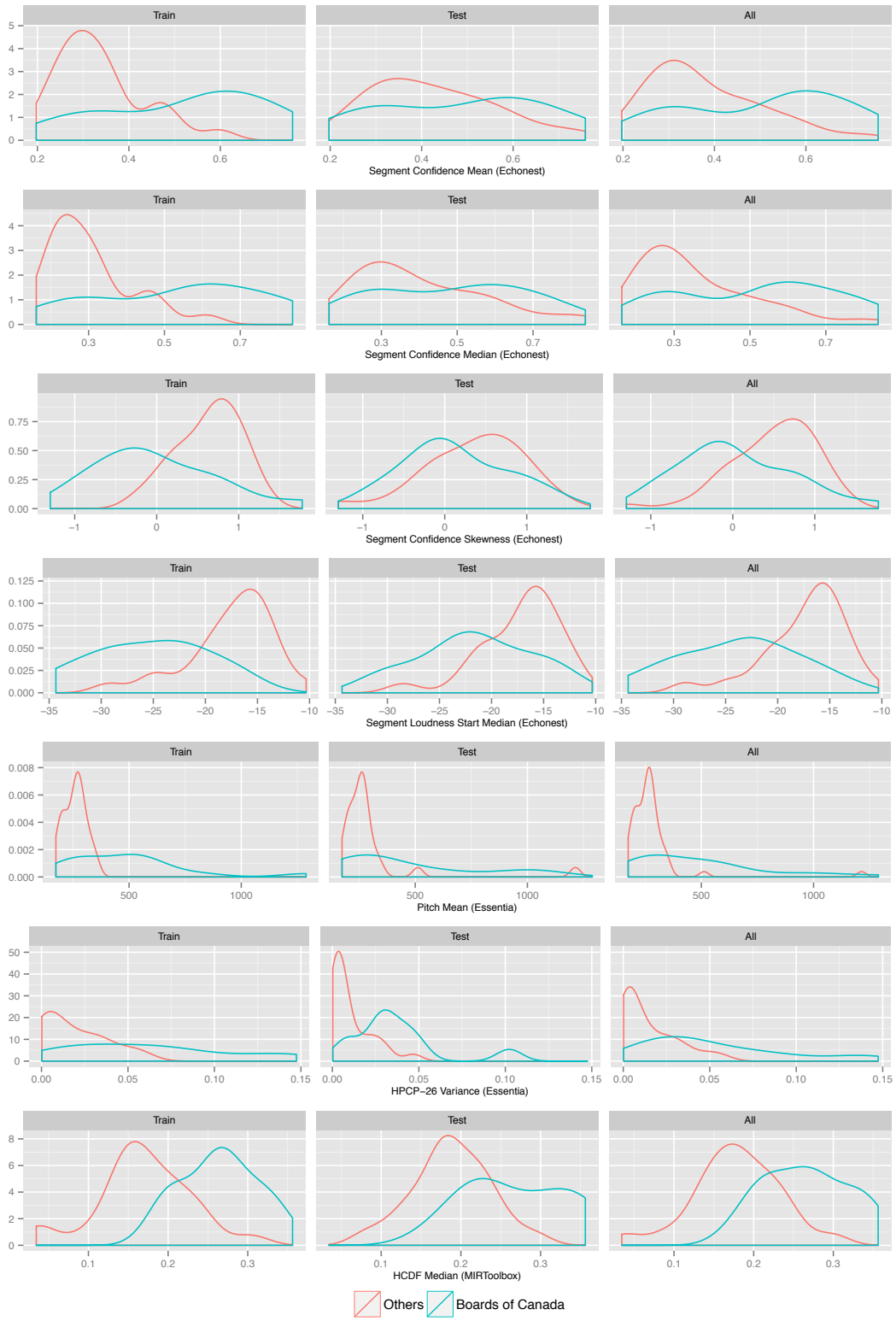

Figure 4.12: Discriminative Features of "Boards of Canada" [3/3]: Probability density functions of some non-timbre-or-rhythm-related audio descriptors (Mean, Median and Skewness of Echonest's Segment Confidence, and Segment Loudness Start Median, Essentia's Pitch Mean and Variance of the 26th HPCP, and MIRToolbox's Median of the Harmonic Change Detection Function) of the works by "Boards of Canada", compared with the other two artists of the "Post-Rock Ambient" dataset

4.3 Descriptive and Characteristic Features

The analysis of the features that better help us to discriminate between artist is useful not only to improve the results in classification tasks, but also to determine relevant stylistic traits of each artist, as we have already seen. Nevertheless, as we have already mentioned a few times now, we strongly believe that limiting the analysis to that kind of features is not enough to completely capture the style that one artist imprints to his or her creations. For that reason we have obtained the most descriptive features for the three artists included in the "Post-Rock Ambient" dataset using the methodology that was explained in Section 3.3.3.

Among the different possibilities that we considered as candidates to our measure of "coherence", we have decided to only analyze the features that are obtained using entropy as selection metric. Even though it does not always provide the largest differences of performance, it allows us to avoid problems with discrete or even constant variables, as well as not needing any previous normalization of the data, aspects that should be considered when using standard deviation or inter-quartile range.

Table 4.6 summarizes the subsets of features that were obtained when selecting the 20 features with lesser Entropy for each of the three artists included in the dataset. The first thing that may be noticed from the list is that there are several features that are considered as descriptive for more than one artist. In fact, six of them appear in all three artists' subsets. Moreover, all those six common features are related with the Echonest's rhythmic analysis. More precisely, they are all related with the shape of the distribution of the time between rhythmic divisions. This seems to be quite reasonable, as electronic music usually keeps a very stable tempo throughout all the duration of a track. Tracks of other genres, conversely, being performed live by humans and not by machines, often show fluctuations in their tempo, either for expressive reasons or simply for natural imprecisions.

Another interesting point that may be worth commenting is that there seems to be coherence in the musical facets represented by the descriptive features of each artist with respect to the ones that were selected as discriminative. Apart from the rhythmic descriptors, which appear in the three artists as we have already mentioned, tonality is the most repeated facet for "The Album Leaf", while in "The American Dollar" timbre is more relevant than in the other artists. "Boards of Canada", on the other hand, has a more diverse selection of features, with a presence of rhythmic features even higher than the rest. This leads us to be quite confident in the usefulness of this type of analysis.

Having said that, we must mention that if we take a look at each of the features individually, the conclusions may be less optimistic. Figures 4.13, 4.14 and 4.15 compare the probability density functions of the selected feature subsets of each of the three artists with the curve obtained by using the tracks by all the other analyzed artists, not only the ones included in the "Post-Rock Ambient" dataset. It is not difficult to notice that in most cases the differences between the distributions are very small. This may be due to the fact that they are traits common in any style of Electronic Music, as the regular tempo, or a

signal that the descriptor is almost irrelevant to indicate artist-specific or genre-specific differences.

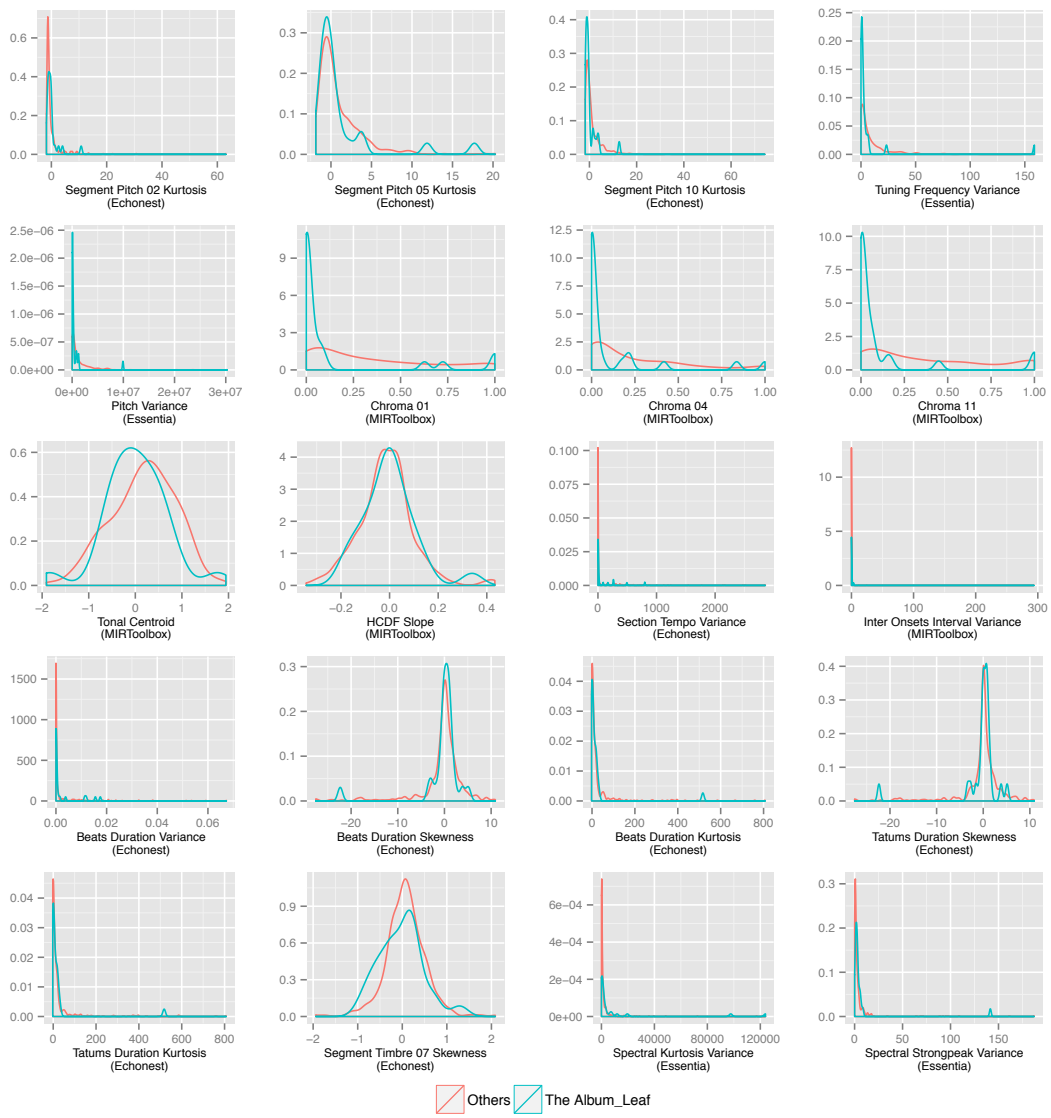

Figure 4.13: Probability density functions of the 20 most descriptive features of the works by "The Album Leaf" using Entropy as selection metric, compared with the remaining 15 artists of the Complete dataset

The most noticeable exceptions to that performance are the MIRToolbox's Chroma coefficients in "The Album Leaf". In the three cases (first, fourth and eleventh coefficient) the distribution is clearly different than the one obtained for the full collection. This does nothing but reinforce the idea that this artist's style is much more focused on tonality than what is usually found in Electronic Music. In fact, one of the very few features that appear both in the descriptive and discriminative subsets of one of the artists, what we have called "characteristic features", is precisely the 11th coefficient of the MIRToolbox's Chroma. The other one is the Echonest's Segment Confidence Skewness, that is selected as discriminative and descriptive for "Boards of Canada". Nevertheless, in this case we don't

think that there is an stylistic explanation for that.

4.4 Conclusion

The closer analysis that we have reported in this chapter has confirmed the potential of the proposed experiments. The qualitative results that we have studied seem to be most of the times coherent with the perceptual impressions that a human listener may have. In this sense, we have confirmed that in the works by "The Album Leaf" tonal characteristics are much more prevalent than in the rest of the considered artists, both in discriminative and descriptive terms. In the case of "The American Dollar", timbral and structural descriptors are the ones that better characterize its style, while "Boards of Canada" presents a higher variety in the attributes present in their works. In general, the steadiness of the rhythm is a characteristic that can be found in all three artists, and, even though it does not allow to distinguish between their works, it is an extremely relevant attribute that should be taken into account when describing the style of any of those artists. However, we think that there is still a large margin of improvement in order to solve the issues that we have detected, specially in the case of the experiment addressed to capture the "descriptive" features of each artist.

Table 4.6: Summary of the 20 most descriptive audio features according to their Entropy for each of the artists included in the "Post-Rock Ambient" dataset: "The Album Leaf" (TAL), "The American Dollar" (TAD) and "Boards of Canada" (BOC)

Library	Audio Feature	Coefficient	Statistic	TAL	TAD	BOC
EN	Section Tempo		Variance	X	X	X
ESS	Inter Beats Interval		Skewness		X	
ESS	Inter Onsets Interval		Variance		X	X
ESS	Inter Onsets Interval		Kurtosis			X
MTB	Inter Onsets Interval		Variance	X		X
EN	Beats Duration		Variance	X	X	X
EN	Beats Duration		Skewness	X	X	X
EN	Beats Duration		Kurtosis	X	X	X
EN	Tatums Duration		Variance			X
EN	Tatums Duration		Skewness	X	X	X
EN	Tatums Duration		Kurtosis	X	X	X
EN	Segment Pitch	2	Kurtosis	X	X	
EN	Segment Pitch	4	Kurtosis			X
EN	Segment Pitch	5	Kurtosis	X		
EN	Segment Pitch	6	Kurtosis		X	
EN	Segment Pitch	10	Kurtosis	X		
ESS	Tuning Frequency		Variance	X		
ESS	Pitch		Variance	X		
MTB	Mode		Mode		X	X
MTB	Chroma	1		X		
MTB	Chroma	4		X		
MTB	Chroma	11		X		
MTB	Tonal Centroid	3				X
MTB	Tonal Centroid	4		X		
MTB	Tonal Centroid	5			X	
MTB	Key Strength	1			X	
MTB	Key Strength	5				X
MTB	Key Strength	12				X
MTB	HCDF		Slope	X		
ESS	Pitch After Max to Before Max Energy Ratio					X
ESS	MFCC	5	Mean		X	
MTB	MIRToolbox	6				X
EN	Segment Timbre	7	Skewness	X	X	
EN	Segment Timbre	8	Median		X	
EN	Segment Timbre	8	Mean		X	
EN	Segment Timbre	9	Kurtosis			X
EN	Segment Timbre	10	Mean		X	
EN	Segment Timbre	10	Skewness			X
ESS	Spectral Kurtosis		Variance	X		
ESS	Spectral Strongpeak		Variance	X		
EN	Section Loudness		Kurtosis		X	
EN	Section Confidence		Skewness		X	
EN	Segment Confidence		Skewness			X

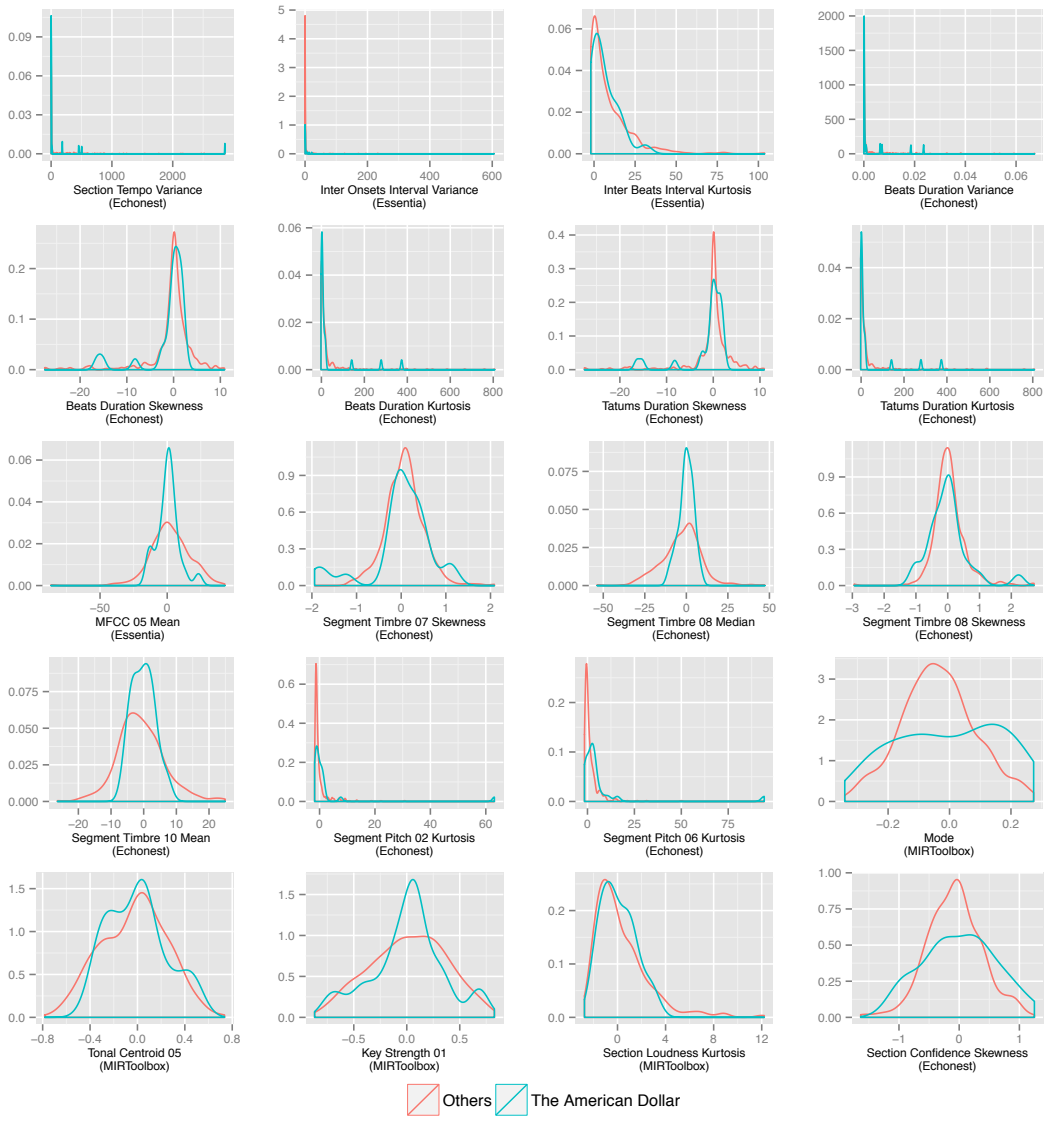

Figure 4.14: Probability density functions of the 20 most descriptive features of the works by "The American Dollar" using Entropy as selection metric, compared with the remaining 15 artists of the Complete dataset

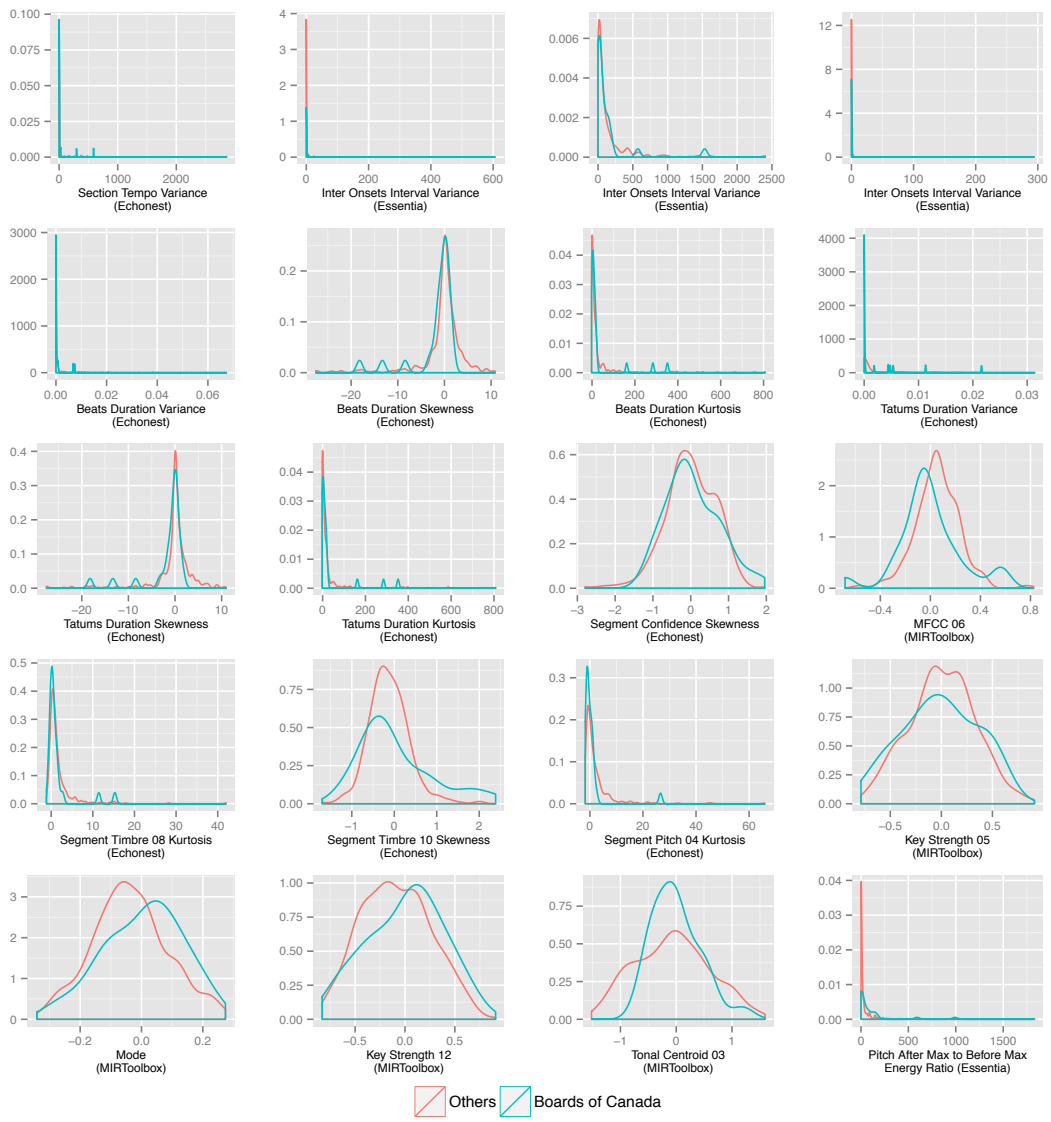

Figure 4.15: Probability density functions of the 20 most descriptive features of the works by "Boards of Canada" using Entropy as selection metric, compared with the remaining 15 artists of the Complete dataset

Chapter 5 | Conclusion

In order to conclude this dissertation, in this final chapter we will discuss the main issues that have appeared during the development of the thesis, as well as summarizing the most relevant contributions that have been achieved and propose a few lines of investigation that may be interesting to take starting from where the present research has reached.

5.1 Discussion

The development of this Master Thesis has faced several struggles from its very beginning, as the developed topic is built over two pillars which are conceptually fuzzy and have received little attention by the MIR community. Both *Electronic Music* and *Style* are difficult to define formally, so our work has required to rely on certain assumptions that may not be shared by some other researchers. For example, we decided to discard works by artists such as Karlheinz Stockhausen or Iannis Xenakis despite their undeniable importance in Electronic Music history due to our assumption that their works are mainly experimental and, thus, it is doubtful that they may be faithful to a given style. Even though we are convinced that this is a very reasonable assumption, we are aware that some readers may disagree with it, while others may even consider that some of the artists finally included in the collection should also be considered as experimental and should have been discarded.

It is important to mention that the selection of artists has been extremely limited by the need to fulfill the requirements to prevent the "Album Effect" to affect the results of the Artist Identification task. This has caused that the composition of the collection may be too biased, favoring certain styles of Electronic Music over others that arguably are at least as important as them in the Electronic Music space. For instance, dance-oriented genres, such as "House" or "Trance" have little or no presence in the collection. Moreover, many highly influential artists in Electronic Music history were not suitable for being included in the collection as they have very rarely published full albums. Singles and EPs are much more common for certain (sub-)genres and artists, causing the created collection not to be as representative as it should be. As an example, "Techno" pioneers, such as Juan Atkins, Derrick May and Kevin Saunderson, were not suitable for being included in the collection for this reason.

Another relevant point that has conditioned the definitive contents of the music collection is the assumption that vocal traits may bias the results of the experiments. In

our opinion, vocal cues help the system to increase the artist prediction accuracy without forcing it to successfully capture the main stylistic traits of the *creator* of the piece, which, as we have already repeated many times, has been the main goal of this thesis. For that reason, we decided to discard any track containing sung or spoken vocals in it. That reduced even more the number of possible candidates to be included in the collection, and probably increased the bias towards certain styles.

A path that was not taken and that may have been helpful for addressing the aforementioned issues, would have been to focus the analysis on small excerpts of fixed length instead of full tracks. This would have also allowed us to look for some musical signatures that can only be detected at smaller scales, such as particular chord progressions. However, we decided to priorize the analysis at larger scales in order to be able to compare the overall characteristics of the works. Even though we think that using fixed-sized excerpts may not be the best solution, is without any doubt an option that should be explored in the future.

With regard to the audio features that we have extracted, once we have ended our analyses, we think that the amount of descriptors employed may be excessive. Even though we have been able to reach satisfactory quantitative and qualitative results in the performed experiments, we have the feeling that the interpretation of the outcomes of the analyses would have been much easier and relevant if we had used a smaller amount of higher-level descriptors. Furthermore, such amount of features has caused some of them to be redundant, and even we have noticed the appearance of some inconsistencies between them, as, for instance, in the MFCCs computed from different libraries, as we already mentioned in Section 3.2. The lack of publicly accessible documentation for several of the used features, specially those extracted from the Echonest Analyzer, is another point that should be taken into account. Having said that, it is very likely that the information obtained by means of the performed experiments will help us to develop more tailored features in the future.

Some decisions directly associated with the experimental methodology can also be a matter of discussion, starting from the idea behind the proposed experiments themselves. Even though we have been insisting since the very beginning of this dissertation that our purpose was not to improve classification accuracies, every single experiment performed involved some degree of automatic classification. However, we strongly believe that this procedure has allowed us to better understand which audio features are more suitable for characterizing the stylistic traits of Electronic Music Artists.

Concerning the choice of the learning algorithm, Support Vector Machines is probably the most widely used in similar studies and in our opinion it provides a good balance between prediction power and model understandability. Alternatively, we could have tested different algorithms in order to determine the highest accuracy that we can obtain with the extracted features. For example, using a boosting meta-algorithm, such as ADA Boost, with a set of weak learners, or even Deep Learning Networks, may have been good candidates for that purpose. Nevertheless, we decided to priorize the design of some non-

conventional experiments more closely related with the determination of the most relevant features for each artist. Anyway, we have noticed that those experiments, despite having shown a high degree of coherence in the features selected, can be improved in order to fix some design flaws. This point will be developed more in detail in Section 5.3.

Most of the issues that have been commented in this discussion have been revealed thanks to the study case that we performed in Chapter 4, a detailed analysis that has proven to be very useful in order to check the validity of our analyses. However, we should be very cautious with the generalization of the methodologies that we have developed and tested, as the number of artists considered was extremely reduced, and the selection methods and feature subset sizes used may not lead to optimal results.

5.2 Summary of Thesis Achievements

Despite the issues reported in the previous section, we strongly believe that the developed research has generated relevant contributions. First of all, we have constructed a music collection suitable for being shared with the community by deeply revising the contents of the one previously used by Melidis in his Master Thesis. The creation of a database including several information provided by different trustworthy sources has allowed us to properly organize its large amount of tracks, as well as to filter out all the content that did not fit the minimum requirements established for the task to which it is addressed. The result is a collection specially fitted for authorship detection and artist style analysis in Electronic Music. To the best of our knowledge, no previous collection has been released for this purpose.

We have shown that current state of the art MIR tools and techniques are able to capture a relevant part of the stylistic traits that characterize the works of the analyzed artists, even though we have just scratched the surface of all the possibilities that this discipline can offer. At this respect, we have been able to improve the quantitave results that had been previously reported in similar studies despite the fact that we have always performed the experiments in worst-case scenarios.

Finally, we have proposed a couple of tests specifically designed to obtain the most relevant features for each artist. In this sense, we distinguished two main categories of features that we think are needed in order to fully characterize the style of an artist. Traditionally, only those descriptors that allow to discriminate between the different categories, in this case, artists, are taken into account. However, we are convinced that a much more accurate stylistic analysis can be achieved when also considering those features which values are coherent among most of the works of an artist, even when they do not allow to distinguish between different authors. Even though we think that the proposed methodology can be improved, we are pretty sure that it has the potential to become a good starting point for developing future studies. The qualitative results that were analyzed in the study case are encouraging, as they show that even with a non-tailored feature set we are able to obtain reasonable conclusions about the most relevant characteristics of each artist.

5.3 Future Work

From the very beginning, the aim of this thesis has been mainly exploratory. In this sense, we consider most of the outcomes that we have reported in this dissertation as starting points to future studies. Obviously, the construction of the collection allows other researchers to replicate and improve the experiments that we have performed, but also to develop their own lines of investigation. For example, achieving reliable computational stylistic analysis of Electronic Music has a great potential to be useful for a number of practical applications, from recommender systems and expert assistants in music production softwares, to automatic djing and composition algorithms.

In the most immediate future, we are aware of some aspects of our own analysis that can be improved or, at least, tested:

- Determining if analyzing fixed-length excerpts of the tracks contained in the music collection leads to similar results may be an interesting point to test. In that case, not only it may lead us to detect different stylistic traits that could not be observed due to the employed scale, but also it would allow us to share the contents of the music collection without being restricted by copyright issues, taking profit of platforms such as the Echonest API.
- The performed analyses have provided very useful information that may help to construct more tailored feature sets, as irrelevancies, redundancies and incoherencies have been noticed when observing the obtained results in detail. Moreover, this information may be really useful for developing high level descriptors specially designed for representing stylistic traits of Electronic Music. In this sense, implementing and testing tailored descriptors would be a really interesting path to take.
- The methodology developed for determining the most discriminative features per artist may be improved if a balanced number of instances per class is fixed. For that purpose, it should be tested if a random selection of tracks among all the possible candidates for being included in the "Other" class is feasible, or if a more balanced composition that represents all the possible artists should be ensured.
- With respect to the experiment addressed to capture what we have called "descriptive" features, even though we are convinced of the usefulness of the concept, in our opinion the methodology can be vastly improved. For example, using album identification as a test for the goodness of the selected subsets has been shown to be less revealing than it should. Other tests, such as track identification by splitting each piece in multiple excerpts may be much more informative.

Those are only four of many possible examples of improvements that may be able increase the relevance of this type of analysis even more. Furthermore, during the development of this Thesis, some other questions have emerged and we think that they are interesting enough to be addressed. For example:

- *Does the "Album Effect" have the same impact in Electronic Music as it does in "recorded" music?*
  It is known that the particular sound that producers and mastering engineers imprint to all the tracks of the same album provides sometimes more cues in classification tasks than the characteristics of the creator/performer of those tracks. For that reason, the separation of train and test albums has become a *de facto* requirement of any Artist Identification task. However, as EM artists usually perform not only the role of creator/composer, but also they act as producers and even mastering engineers of their own albums, it is possible that the differential sound of each album may not be as relevant. In that case, many of the artists that did not fulfill the requirements to be included in the collection could be added, allowing it to be larger and, thus, increasing the research possibilities that it may offer.
- *Can we use the proposed techniques to track the evolution of the style of Electronic Music artists and their influences through time?*
  As we mentioned in the background review, many authors evolve their creation style until they reach artistic maturity. In some cases we can even distinguish clearly different styles in their own works. In this sense, it would be interesting to test if the extracted features are suitable for detecting stylistic evolution and capturing possible influences between different artists.
- *At which degree is the style of Electronic Music artists conditioned by the parallel improvement of the technology to which they have access?*
  Electronic Music artists rely undeniably in the available technology in order to obtain the sounds that they want to create. However, some of them seem to be more constant in their style independently than the technology used than others. We think it would be interesting to identify those authors and, on the other hand, determine at which degree contemporary artists show similar trends in their creations.

We expect that in the future the path that has been described in this dissertation may be continued and those questions, among others, may be aswered, leading at their time to others maybe even more interesting.

References

[1] Ackerman, J. S. A Theory of Style. *Aesthetics and Art Criticism, The Journal of 20*, 3 (Spring 1962), 227–237.

[2] Akkermans, V., Serrà, J., and Herrera, P. Shape-Based Spectral Contrast Descriptor. In *6th Sound and Music Computing Conference (SMCC'09), Proc.* (Porto, Portugal, July 2009), pp. 143–148.

[3] Aucouturier, J.-J., and Pachet, F. Representing Musical Genre: A State of the Art. *New Music Research, Journal of 32*, 1 (Aug. 2003), 83–93.

[4] Backer, E., and Kranenburg, P. On Musical Stylometry. A Pattern Recognition Approach. *Pattern Recognition Letters 26*, 3 (Feb. 2005), 299–309.

[5] Bartsch, M. A., and Wakefield, G. H. Singing Voice Identification Using Spectral Envelope Estimation. *Speech and Audio Processing, IEEE Transactions on 12*, 2 (March 2004), 100–109.

[6] Bergstra, J., Casagrande, N., and Eck, D. Two Algorithms for Timbre- and Rhythm-Based Multi-Resolution Audio Classification. In *2005 Music Information Retrieval Evaluation Exchange (MIREX 205)* (London, UK, 2005).

[7] Chatman, S. "Style:" A Narrow View. *College Composition and Communication 18*, 2 (May 1967), 72–76.

[8] Collins, N. Computational Analysis of Musical Influence: A Musicological Case Study using MIR Tools. In *11th International Society for Music Information Retrieval Conference (ISMIR'10), Proc.* (Utrecht, Netherlands, Aug. 2010), pp. 177–182.

[9] Collins, N. Influence in Early Electronic Dance Music: An Audio Content Analysis Investigation. In *13th International Society for Music Information Retrieval Conference (ISMIR'12), Proc.* (Porto, Portugal, Oct. 2012), pp. 1–6.

[10] Cope, D. *Computers and Musical Style.* Oxford University Press, 1991.

[11] Cruz-Alcázar, P. P., and Vidal-Ruiz, E. Modeling Musical Style using Grammatical Inference Techniques: A Tool for Classifying and Generating Melodies. In *3rd*

International Conference on WEB Delivering of Music (WEDELMUSIC'03), Proc. (Leeds, UK, 2003).

[12] Dannenberg, R. B. Style in Music. In *The Structure of Style. Algorithmic Approaches to Understanding Manner and Meaning*, S. Argamon, K. Burns, and S. Dubnov, Eds. Springer-Verlag, 2010, ch. 3, pp. 45–57.

[13] Dannenberg, R. B., Thom, B., and Watson, D. A Machine Learning Approach to Musical Style Recognition. In *1997 International Computer Music Conference, Proc.* (Thessaloniki, Greece, September 1997).

[14] Dayal, G., and Ferrigno, E. Electronic Dance Music. Grove Music Online. Oxford Music Online. Oxford University Press, accessed March 22, 2014, `http://www.oxfordmusiconline.com/subscriber/article/grove/music/A2224259`.

[15] Dor, O., and Reich, Y. An Evaluation of Musical Score Characteristics for Automatic Classification of Composers. *Computer Music Journal 35*, 3 (Sept. 2011), 86–97.

[16] Downie, J. S. The Music Information Retrieval Evaluation Exchange (2005-2007): A window into music information retrieval research. *Acoustical Science and Technology 29*, 4 (2008).

[17] Dubnov, S., Assayag, G., Lartillot, O., and Bejerano, G. Using Machine-Learning Methods for Musical Style Modeling. *Computer 36*, 10 (Oct. 2003), 73–80.

[18] Elkins, J. Style. Grove Art Online. Oxford Art Online. Oxford University Press, accessed March 19, 2014, `http://www.oxfordartonline.com/subscriber/article/grove/art/T082129`.

[19] Emmerson, S., and Smalley, D. Electro-Acoustic Music. Grove Music Online. Oxford Music Online. Oxford University Press, accessed March 22, 2014, `http://www.oxfordmusiconline.com/subscriber/article/grove/music/08695`.

[20] Fu, Z., Lu, G., Ting, K. M., and Zhang, D. A Survey of Audio-Based Music Classification and Annotation. *Multimedia, IEEE Transactions on 13*, 2 (2011), 303–319.

[21] Fucks, W. Mathematical Analysis of Formal Structure of Music. *Information Theory, IRE Transactions on 8*, 5 (1962), 225–228.

[22] Fujihara, H., Goto, M., Kitahara, T., and Okuno, H. G. A Modeling of Singing Voice Robust to Accompaniment Sounds and Its Application to Singer Identification and Vocal-Timbre-Similarity-Based Music Information Retrieval. *Audio, Speech and Language Processing, IEEE Transactions on 18*, 3 (March 2010), 638–648.

[23] Goodman, N. The Status of Style. *Critical Inquiry 1*, 4 (June 1975), 799–811.

[24] Haro, M., Serrà, J., Herrera, P., and Corral, A. Zipf's Law in Short-Time Timbral Codings of Speech, Music, and Environmental Sound Signals. *PLoS ONE 7* (March 2012), e33993.

[25] Herrera, P., Klapuri, A., and Davy, M. Automatic Classification of Pitched Musical Instrument Sounds. In *Signal Processing Methods for Music Transcription*, A. Klapuri and M. Davy, Eds. Springer, 2006, ch. 6, pp. 163–200.

[26] Holmes, D. I. Analysis of Literary Style. A Review. *Royal Statistical Society, Journal of. Series A (General) 148*, 4 (1985), 328–341.

[27] Holmes, D. I. The Evolution of Stylometry in Humanities Scholarship. *Literary and Linguistic Computing 13*, 3 (1998), 111–117.

[28] Hughes, J. M., Graham, D. J., and Rockmore, D. N. Quantification of Artistic Style through Sparse Coding Analysis in the Drawings of Pieter Brugel the Elder. *National Academy of Sciences, Proc. 107*, 4 (January 2010), 1279–1283.

[29] Johnson Jr., C. R., Hendriks, E., Berezhnoy, I. J., Brevdo, E., Hughes, S. M., Daubechies, I., Li, J., Postma, E., and Wang, J. Z. Image Processing for Artist Identification. *Signal Processing Magazine, IEEE 25*, 4 (July 2008), 37–48.

[30] Kaliakatsos-Papakostas, M. A., Epitropakis, M. G., and Vrahatis, M. G. Musical Composer Identification through Probabilistic and Feedforward Neural Networks. In *Applications of Evolutionary Computation, EvoApplications 2010, Proc., Part II* (Istanbul, Turkey, Apr. 2010), Springer, pp. 411–420.

[31] Kim, Y. E., and Whitman, B. Singer Identification in Popular Music Recordings Using Voice Coding Features. In *3rd International Society for Music Information Retrieval Conference (ISMIR'02). Proc.* (Paris, France, 2002).

[32] Kim, Y. E., Williamson, D. S., and Pilli, S. Towards Quantifying the "Album Effect" in Artist Identification. In *7th International Society of Music Information Retrieval Conference (ISMIR'06), Proc.* (Victoria, BC, Canada, Oct. 2006), pp. 393–394.

[33] Kirss, P. Audio Based Genre Classification of Electronic Music. Master's thesis, University of Jyväslylä, June 2007.

[34] Leichtentritt, H. Aesthetic Ideas as the Basis of Musical Styles. *Aesthetics and Art Criticism, The Journal of 4*, 2 (December 1945), 65–73.

[35] Levitin, D. J., Chordia, P., and Menon, V. Musical Rhythm Spectra from Bach to Joplin Obey a 1/F Power Law. *National Academy of Sciences, Proc. 109*, 10 (Oct. 2012), 3716–3720.

[36] Lim, S.-C., Byun, K., Lee, J.-S., Jang, S.-J., and Kim, M. Y. Music Genre/Mood Classification: MIREX 2012. In *2012 Music Information Retrieval Evaluation Exchange (MIREX 2012)* (Porto, Portugal, 2012).

[37] Lomax, A. Notes on a Systematic Approach to the Study of Folk Song. *International Folk Music Council, Journal of 8* (1956), 48–50.

[38] Lutoslawski, W. On Stylometry. *The Classical Review 11*, 6 (July 1897), 284–286.

[39] Manaris, B., Romero, J., Machado, P., Krehbiel, D., Hirzel, T., Pharr, W., and Davis, R. B. Zipf's Law, Music Classification, and Aesthetics. *Computer Music Journal 29*, 1 (Feb. 2005), 55–69.

[40] Mandel, M. I., and Ellis, D. P. W. Song-Level Features and Support Vector Machines for Music Classification. In *6th International Society of Music Information Retrieval Conference (ISMIR'05), Variation 2, Proc.* (London, UK, Sept. 2005).

[41] McLeod, K. Genres, Subgenres, Sub-Subgenres and More: Musical and Social Differentiation Within Electronic/Dance Music Communities. *Popular Music Studies, Journal of 13*, 1 (2001), 59–75.

[42] Mearns, L., Tidhar, D., and Dixon, S. Characterisation of Composer Style Using High-Level Musical Features. In *3rd International Workshop on Machine Learning and Music (MML'10), Proc.* (New York, NY, USA, Oct. 2010), pp. 37–40.

[43] Melidis, P. Electronic Music Artist Identification. Master's thesis, Universitat Pompeu Fabra, 2012.

[44] Meyer, L. B. *Style and Music: Theory, History and Ideology*, 1 ed. University of Pennsylvania Press, Philadelphia, PE, USA, 1989.

[45] Moore, A. F. Categorical Conventions in Music Discourse: Style and Genre. *Music & Letters 82*, 3 (August 2001), 432–442.

[46] Murry, J. M. *The Problem of Style.* Oxford University Press, 1922.

[47] Nettl, B. Some Linguistic Approaches to Musical Analysis. *International Folk Music Council, Journal of 10* (1958), 37–41.

[48] Pachet, F. The Continuator: Musical Interaction With Style. *New Music Research, Journal of 32*, 3 (2002), 333–341.

[49] Pachet, F., Roy, P., and Ghedini, F. Creativity through Style Manipulation: the Flow Machines Project. In *2013 Marconi Institute for Creativity Conference, Proc.* (Bologna, Italy, September 2013).

[50] Pascall, R. Style. Grove Music Online. Oxford Music Online. Oxford University Press, accessed March 19, 2014, `http://www.oxfordmusiconline.com/subscriber/article/grove/music/27041`.

[51] Pérez-Sancho, C., Iñesta, J. M., and Calera-Rubio, J. A Text Categorization Approach for Music Style Recognition. In *2nd Iberian Conference on Pattern Recognition and Image Analysis, Proc.* (Estoril, Portugal, 2005), vol. 2.

[52] Ramirez, R., Maestre, E., and Serra, X. Automatic Performer Identification in Commercial Monophonic Jazz Performances. *Pattern Recognition Letters 31*, 12 (december 2010), 1514–1523.

[53] Ratner, L. G. *Classic Music: Expression, Form and Style*, 1 ed. Schirmer Books, New York, NY, USA, 1980.

[54] Ren, J.-M., Wu, M.-J., and Chang, K. K. MIREX 2011: Music Genre Classification via Sparse Representation. In *2011 Music Information Retrieval Evaluation Exchange (MIREX 2011)* (Miami, FL, USA, 2011).

[55] Sauders, C., Hardoon, D. R., Shawe-Taylor, J., and Widmer, G. Using String Kernels to Identify Famous Performers from their Playing Style. *Intelligent Data Analysis 12* (2008), 425–440.

[56] Schapiro, M. Style. In *Antropology Today*, A. L. Kroeber, Ed. University of Chicago Press, Chicago, IL, USA, 1953, pp. 287–312.

[57] Sesmero Molina, J. Electronic Dance Music Genre Classification. Master's thesis, Universitat Pompeu Fabra, September 2008.

[58] Shen, J., Shepherd, J., Cui, B., and Tan, K.-L. A Novel Framework for Efficient Automated Singer Identification in Large Music Databases. *Information Systems, ACM Transactions on 27*, 3 (May 2009), 1–31.

[59] Siefkes, M. Style: A New Semiotic View on an Old Problem. *Ars Semiotica 34*, 1-2 (2011), 15–25.

[60] Stamatatos, E., and Widmer, G. Automatic Identification of Music Performers with Learning Ensembles. *Artificial Intelligence 165* (2005), 37–56.

[61] Sturm, B. L. Classification Accuracy is not Enough. *Intelligent Information Systems, Journal of 41*, 3 (2013), 371–406.

[62] Tsai, W.-H., and Lin, H.-P. Popular Singer Identification Based on Cepstrum Transformation. In *2010 IEEE International Conference on Multimedia and Expo (ICME 2010), Proc.* (Singapore, July 2010).

[63] Urbano, J., Schedl, M., and Serra, X. Evaluation in Music Information Retrieval. *Intelligent Information Systems, Journal of 41*, 3 (2013), 345–369.

[64] Verdonk, P. Style. In *Encyclopedia of Language and Linguistics*, K. Brown, Ed., 2 ed. Elsevier, 2006, pp. 196–210.

[65] Whitman, B., Flake, G., and Lawrence, S. Artist Detection in Music with Minnowmatch. In *2001 IEEE Workshop on Neural Networks for Signal Processing, Proc.* (Falmouth, MA, USA, 2001), pp. 559–568.

[66] Widmer, G., Dixon, S., Goebl, W., Pampalk, E., and Tobudic, A. In Search of the Horowitz Factor. *AI Magazine 24*, 3 (2003), 111–130.

[67] Zhang, T. System and Method for Automatic Singer Identification. In *2003 IEEE International Conference on Multimedia and Expo (ICME 2003), Proc.* (Baltimore, MD, USA, July 2003).

[68] Zipf, G. K. *Selected Studies of the Principle of Relative Frequency in Language.* Harvard University Press, Cambridge, MA, USA, 1932.

[69] Zipf, G. K. *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology.* Addison-Wesley Press, Inc., Cambridge, MA, USA, 1949.

Appendices

Appendix A | Artists included in the Music Collection

In this appendix we list the artists and albums that are contained in the Electronic Music Artists (EMA) collection.

µ-ziq: "Bluff Limbo", "Lunatic Harness", "Royal Astronomy", "Duntisbourne Abbots Soulmate Devastation Technique", "Chewed Corners"

Amon Tobin: "Bricolage", "Permutation", "Supermodified", "Out From Out Where", "Chaos Theory"

Aphex Twin: "Selected Ambient Works 85-92", "...I Care Because You Do", "Richard D. James Album", "Come To Daddy", "Drukqs"

Asmus Tietchens: "Spät-Europa", "Notturno", "β-Menge", "Biotop", "Litia"

Autechre: "Amber", "Tri Repetae", "Chiastic Slide", "LP5". "Confield"

Bass Comunion: "Bass Communion","Atmospherics", "Bass Communion II", "Bass Communion III", "Ghosts on Magnetic Tape"

Ben Frost: "Music for Sad Children", "Steel Wound", "Theory of Machines", "By the Throat", "The Invisibles"

Biosphere: "Patashnik", "Insomnia: No Peace for the Wicked", "Cirque", "Substrata 2 (CD 1)", "Biosphere", "Shenzhou"

Boards of Canada: "Twoism (2002 CD)", "Hi-Scores", "Music Has the Right to Childre", "Geogaddi", "The Campfire Headphase", "Trans Canada Highway"

Bonobo: "Animal Magic", "Dial 'M' for Monkey", "Days to Come", "Black Sands", "The North Borders"

Ceephax aka Ceephax Acid Crew: "Exidy Tours", "Volume 1", "Drive Time LP", "Ceeland", "United Acid Emirates"

Clark aka Chris Clark: "Clarence Park", "Body Riddle", "Turning Dragon", "Totems Flare", "Iradelphic"

Conrad Schnitzler: “Blau”, “Con 3”, “Congratulacion”, “Ballet Statique”, “Electrocon”

Cosmic Hoffmann: “Shiva Connection”, “Electric Trick”, “Space Gems”, “Outerspace Gems”, “Hypernova”

Daedelus: “Invention”, “Snowdonia”, “A Gent Agent”, “Daedelus Denies the Day’s Demise”, “Drown Out”

deadma5: “Get Scraped”, “Vexillology”, “At Play, Volume 2”, “4x4=12”, “> album title goes here <”

Dj Food: “Jazz Brakes, Volume 2”, “Jazz Brakes, Volume 4”, “Jazz Brakes, Volume 5”, “A Recipe for Disaster”, “Kaleidoscope”

Dntel: “Early Works for Me If It Works for You”, “Life Is Full of Possibilities”, “Something Always Goes Wrong”, “Aimlessness”, “After Parties 1 and 2”

Dom F. Scab: “Innerseed”, “Binary Secrets”, “Analogical Confessions”, “Facta”, “Twelve Stories”

Fennesz: “Hotel Paral.lel”, “Endless Summer”, “Field Recordings 1995:2002”, “Venice”, “Black Sea”

Flying Lotus: “1983”, “Angeles”, “Cosmogramma”, “IIOIO”, “Pattern+Grid World”

Four Tet: “Dialogue”, “Pause”, “Rounds”, “Everything Ecstatic”, “There Is Love in You”

ISAN: “Beautronics”, “Lucky Cat”, “Clockwork Menagerie”, “Meet Next Life”, “Plans Drawn in Pencil”

Jaga Jazzist: “Jaevla Jazzist Grete Stitz”, “Livingroom Hush”, “The Stix”, “What We Must”, “One-Armed Bandit”

Jean Michel Jarre: “Deserted Palace”, “Oxygène”, “Les Chants Magnétiques”, “Zoolook”, “Geometry of Love”

Jeff Mills: “Waveform Transmission, Volume 1”, “Waveform Transmission, Volume 3”, “2087”, “Fantastic Voyage”, “The Power”

Legowelt: “Reports From the Backseat Pimp”, “Cat Disco”, “Dark Days 2”, “Manuel Noriega”, “Vatos Locos”

Lemongrass: “Drumatic Universe”, “Voyage au Centre de la Terre”, “Windows”, “Skydiver”, “Fleur Solaire”, “Ikebana”, “Filmothèque”, “Pour l’Amour”

Luke Vibert: “Big Soup”, “YosepH”, “Lover’s Acid”, “Chicago, Detroit, Redruth”, “Ridmik”

Matmos: “A Chance to Cut Is a Chance to Cure”, “Civil War”, “The Rose Has Teeth in the Mouth of a Beast”, “Supreme Balloon”, “Treasure State”

Michael Stearns: “Planetary Unfolding”, “Sacred Site”, “Encounter: A Journey in the Key of Space”, “Whispers”, “Lyra”, “M’Ocean”

Monolake: “Interstate”, “Cinemascope”, “Gravity”, “Momentum”, “Silence”

Mouse on Mars: “Iaora Tahiti”, “Autoditacker”, “Glam”, “Niun Niggung”, “Idiology”

Muslimgauze: “Silknoose”, “Blue Mosque”, “Fakir Sind”, “Lo-Fi India Abuse”, “Baghdad”

Neuronium: “Supranatural”, “Mystykatea”, “Caldea Music”, “Synapsia”, “Nihilophobia”

Oneohtrix Point Never: “Betrayed in the Octagon (2009 Vinyl)”, “Rifts (CD 2)”, “Without People”, “Returnal”, “Replica”

Oöphoi: “The Spirals of Time”, “Time Fragments, Volume 2: The Archives 1998/1999”, ‘Static Soundscapes: Three Lights at the End of the World ‘”, “The Rustling of Leaves (2005 CD)”, “Dreams”

Orbital: “Orbital”, “Orbital 2”, “Snivilisation”, “The Altogether”, “Blue Album”

Pan Sonic: “Kulma”, “Aaltopiiri”, “Kesto (234.48:4) (CD 1)”, “Katodivaihe”, “Gravitoni”

Plaid: “Mbuki Mvuki”, “Not for Threes”, “Rest Proof Clockwork”, “Double Figure”, “Spokes”

Plastikman: “Sheet One”, “Musik”, “Recycled Plastik”, “Consumed”, “Closer”

Prefuse 73: “Vocal Studies + Uprock Narratives”, “One Word Extinguisher”, “Surrounded by Silence”, “Preparations”, “Interregnums”

Ratatat: “Ratatat”, “9 Beats”, “Classics”, “LP3”, “LP4”

Rudy Adrian: “SubAntarctica: Atmospheric Works, Volume 1”, “The Healing Lake”, “Iridescence: Sequencer Sketches, Volume 2”, “Par Avion: Sequencer Sketches, Volume 4”, “Desert Realms”, “Starfields: Sequencer Sketches, Volume 3”

SND: “Newtables”, “Stdio”, “Systems Medley / Planets”, “Tender Love”, “Atavism”

Squarepusher aka Chaos A.D.: “Feed Me Weird Things”, “Normal Daddy”, “Buzz Caner”, “Go Plastic”, “Hello Everything”, “Ufabulum”

Steve Roach: “Dreamtime Return”, “World’s Edge”, “Mystic Chords & Sacred Spaces, Part 1”, “Mystic Chords & Sacred Spaces, Part 2”, “Spiral Meditations”

Sun Electric: “O’locco”, “Kitchen”, “Aaah!”, “Present”, “Via Nostra”

Susumu Yokota: “1998”, “Magic Thread”, “Sakura”, “Will”, “Baroque”

Tetsu Inoue: “Ambiant Otaku”, “Slow and Low”, “Receiver”, “Psycho-Acoustic”, “Waterloo Terminal”

The Album Leaf: “An Orchestrated Rise to Fall”, “One Day I’ll Be On Time”, “In a Safe Place”, “Into the Blue Again”, “The Enchanted Hill”, “Forward/Return”

The American Dollar: “The American Dollar”, “The Technicolour Sleep”, “A Memory Stream”, “Atlas”, “Awake in the City”

The Caretaker: “A Stairway to the Stars”, “We’ll All Go Riding on a Rainbow”, “Persistent Repetition of Phrases”, “An Empty Bliss Beyond This World”, “Patience (After Sebald)”

The Future Sound of London: “Accelerator”, “Lifeforms”, “ISDN”, “The Isness”, “Teachings From the Electronic Brain: The Best of FSOL”, “Dead Cities”

The Orb: “The Orb’s Adventures Beyond the Ultraworld”, “U.F.Orb”, “Orbus Terrarum”, “Orblivion”

Throbbing Gristle: “The Second Annual Report (1991 CD)”, “D.o.A. The Third and Final Report (1991 CD)”, “20 Jazz Funk Greats”, “The First Annual Report of Throbbing Gristle”, “The Third Mind Movements”

Tim Hecker: “Radio Amor”, “Mirages”, “Harmony in Ultraviolet”, “An Imaginary Country”, “Ravedeath, 1972”

Troum: “Ryna”, “Tjukurrpa, Part One: Harmonies”, “Tjukurrpa, Part Three: Rhythms and Pulsations”, “Seeing-Ear Gods”, “AIWS”

Tycho: “The Science of Patterns”, “Sunrise Projector”, “Past Is Prologue”, “Dive”, “Awake”

Vangelis: “Albedo 0.39”, “Spiral”, “Opera Sauvage”, “See You Later”, “Mask”, “Blade Runner (Limited Edition)”

Venetian Snares: “Fuck Canada // Fuck America”, “The Chocolate Wheelchair Album”, “Meathole”, “Winnipeg Is a Frozen Shithole”, Cavalcade of Glee and Dadaist Happy Hardcore Pom Poms”

Vidna Obmana: “Passage in Beauty”, “Revealed by Composed Nature”, “Spiritual Bonding”, “The River of Appearance”, “Crossing the Trail”

Wagon Christ: “Phat Lab. Nightmare”, “Throbbing Pouch”. “Tally Ho!”, “Musipal”, “Sorry I Make You Lush”

William Orbit: “Strange Cargo”, “Strange Cargo III”, “Strange Cargo Hinterland”, “Hello Waveforms”, “My Oracle Lives Uptown”

Appendix B | Dataset Contents

In this appendix we have included some tables which contain the information about the excerpts that are included in each of the six small datasets that we have considered during this dissertation. Each of those tables indicate:

- the excerpt identificator number
- the name of the artist of the album from where the excerpt was extracted -who, in some cases, may correspond to an alias of the real artist-
- the year in which the album was released
- the title of the album
- the number of the track in the album that corresponds to the extracted excerpt
- the name of the artist as it is considered for classification purposes -that, as already mentioned, may be different from the album artist-
- the title of the track, if any
- the Echonest Identificator that corresponds to the considered track, for reproducibility purposes

Atmospheric Ambient

We have included in the "Atmospheric Ambient" Dataset works by Michael Stearns, Tetsu Inoue and Vidna Obmana. More precisely, and as it can be seen in Table B.1, the albums that have been considered are:

Michael Stearns
: Lyra (1983), Planetary Unfolding (1984), Floating Whispers (1987), Encounter: A Journey in the Key of Space (1988), Sacred Site (1993)

Tetsu Inoue
: Ambiant Otaku (1994), Slow and Low (1995), World Receiver (1996), Waterloo Terminal (1998), Psycho-Acoustic (1998)

Vidna Obmana
: Passage in Beauty (1991), Revealed by Composed Nature (1993), The Spiritual Bonding (1994), The River of Appearance (1996), Crossing the Trail (1998)

IDM Ambient

In the "IDM Ambient" Dataset there are included works by ISAN, Monolake and Tycho. More precisely, and as it can be seen in Table B.2, the albums that have been considered are:

ISAN

Beautronics (1998), Lucky Cat (2001), Clockwork Menagerie (2002), Meet Next Life (2004), Plans Drawn in Pencil (2006)

Monolake

Interstate 1999), Gravity (2001), Cinemascope (2001), Momentum (2003), Silence (2009)

Tycho

The Science of Patterns (2002), Sunrise Projector (2004), Past Is Prologue (2006), Dive (2011), Awake (2014)

Post-Rock Ambient

The "Post-Rock Ambient" Dataset consists on works by Boards of Canada, The American Dollar and The Album Leaf. More precisely, and as it can be seen in Table B.3, the albums that have been considered are:

Boards of Canada

Twoism (1995), Hi-Scores (1996), Music Has the Right to Children (1998), Geogaddi (2002), The Campfire Headphase (2005)

The American Dollar

The American Dollar (2006), The Technology Sleep (2007), A Memory Stream (2008), Atlas (2010), Awake in the City (2012)

The Album Leaf

An Orchestrated Rise to Fall (1999), One Day I'll Be On Time (2001), In a Safe Place (2004), Into the Blue Again (2006), The Enchanted Hill (2007)

IDM

We have included in the "IDM" Dataset works by Aphex Twin, Autechre and Squarepusher. More precisely, and as it is shown in Table B.4, the albums that have been considered are:

Aphex Twin

Selected Ambient Works 85-92 (1992), ...I Care Because You Do (1995), Richard D. James Album (1996), Drukqs (2001)

Autechre

Amber (1994), Tri Repetae (1995), Chiastic Slide (1997), LP5 (1998), Confield (2001)

Squarepusher aka Chaos A.D.

Feed Me Weird Things (1996), Buzz Caner (1998), Go Plastic (2001), Hello Everything (2006), Ufabulum (2012)

Nu Jazz

In the "Nu Jazz" Dataset there are included works by Bonobo, Four Tet and Lemongrass. More precisely, and as it is shown in Table B.5, the albums that have been considered are:

Bonobo

Animal Magic (2000), Dial 'M' for Monkey (2003), Days to Come (2006), Black Sands (2010), The North Borders (2013)

Four Tet

Dialogue (1999), Pause (2001), Rounds (2003), Everything Ecstatic (2005), There Is Love in You (2010)

Lemongrass

Drumatic Universe (1998), Voyage au Centre de la Terre (2001), Windows (2001), Fleur Solaire (2004), Filmothèque (2007)

Techno

The "Techno" Dataset consists on works by jeff Mills, Legowelt and Plastikman. More precisely, and as it is shown in Table B.6, the albums that have been considered are:

Jeff Mills

Waveform Transmission, Volume 1 (1992), Waveform Transmission, Volume 3 (1994), The Power (2011), Fantastic Voyage (2011), 2087 (2011)

Legowelt

Reports From Backseat Pimp (1998), Astro Cat Disco (2006), The Rise and Fall of Manuel Noriega (2008), Dark Days 2 (2008), Vatos Locos (2009)

Plastikman

Sheet One (1993), Recycled Plastik (1994), Musik (1994), Consumed (1998), Closer (2003)

Table B.1: Content of the Atmospheric Ambient Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
1	Vidna Obmana	1991	Passage in Beauty	3	Vidna Obmana	Awaken in Floating Colours - Composition 3	TREDGGJ14641A56ED2
2	Vidna Obmana	1991	Passage in Beauty	8	Vidna Obmana	Passage in Beauty	TRLASPI14642FECFA1
3	Vidna Obmana	1991	Passage in Beauty	7	Vidna Obmana	Awaken in Floating Colours - Composition 4	TRNIOTO14644E226A4
4	Vidna Obmana	1991	Passage in Beauty	4	Vidna Obmana	Mood in Pearls	TRBGVEJ14644E3F34D
5	Vidna Obmana	1991	Passage in Beauty	5	Vidna Obmana	Bright Fall	TRNHMIA14644E55E95
6	Vidna Obmana	1994	The Spiritual Bonding	3	Vidna Obmana	Challenging Boundaries	TRLIZIN14644E705EC
7	Vidna Obmana	1994	The Spiritual Bonding	4	Vidna Obmana	The Spiritual Bonding	TROFCYX14644EB40BE
8	Vidna Obmana	1994	The Spiritual Bonding	1	Vidna Obmana	The Feather Cycle	TRWUBGY14644EF1B48
9	Vidna Obmana	1994	The Spiritual Bonding	7	Vidna Obmana	From the Stepping Stone	TRMTVNG14644F0691A
10	Vidna Obmana	1994	The Spiritual Bonding	2	Vidna Obmana	Spatial Prophecy (Song of the Tribal)	TRVLOPU14644F45502
11	Vidna Obmana	1996	The River of Appearance	1	Vidna Obmana	The Angelic Appearance	TRYBYXC14644F5CD73
12	Vidna Obmana	1996	The River of Appearance	2	Vidna Obmana	Ephemeral Vision	TRQHMIZ14644F7D49E
13	Vidna Obmana	1996	The River of Appearance	3	Vidna Obmana	A Scenic Fall	TRYQLYE14644FA5460
14	Vidna Obmana	1996	The River of Appearance	6	Vidna Obmana	Weaving Cluster	TRNVESB14644FBC0D7
15	Vidna Obmana	1996	The River of Appearance	4	Vidna Obmana	Night-Blooming	TRNGOBN14644FDCA95
16	Vidna Obmana	1998	Crossing the Trail	5	Vidna Obmana	The Esoteric Source	TRZMESE14644FF88E1
17	Vidna Obmana	1998	Crossing the Trail	3	Vidna Obmana	Forest Arrow	TRCMEAV14645014C1D
18	Vidna Obmana	1998	Crossing the Trail	1	Vidna Obmana	Encountering Terrain	TRUKCWA14645040130
19	Vidna Obmana	1998	Crossing the Trail	6	Vidna Obmana	The Giant Traveller	TRXRHYL1464506340B
20	Vidna Obmana	1998	Crossing the Trail	4	Vidna Obmana	Mission Ground	TRZRJBN14645089EDD
21	Vidna Obmana	1993	Revealed by Composed Nature	9	Vidna Obmana	The Gilding Marrow	TREJLXN146450BD91D
22	Vidna Obmana	1993	Revealed by Composed Nature	5	Vidna Obmana	Still Wandering	TRGZXRS146450E1160
23	Vidna Obmana	1993	Revealed by Composed Nature	3	Vidna Obmana	Unfold Gradient	TRGURTL146450F2D8D
24	Vidna Obmana	1993	Revealed by Composed Nature	2	Vidna Obmana	Out From the Garden Reminded	TRCOMZN1464510B74F
25	Vidna Obmana	1993	Revealed by Composed Nature	7	Vidna Obmana	Until the Glowering Space 1	TRTPTPA1464511FAEB
26	Tetsu Inoue	1996	World Receiver	5	Tetsu Inoue	Invitable Colour	TRSAUGJ1464512832D
27	Tetsu Inoue	1996	World Receiver	1	Tetsu Inoue	Inter Link	TRGMFOZ14645159D75
28	Tetsu Inoue	1996	World Receiver	6	Tetsu Inoue	Mood Swing	TRIBSZC14645185B29
29	Tetsu Inoue	1996	World Receiver	2	Tetsu Inoue	Health Loop	TRYCKKF146451AECAC
30	Tetsu Inoue	1996	World Receiver	4	Tetsu Inoue	Background Story	TRUBAYL146451D73E3
31	Tetsu Inoue	1998	Waterloo Terminal	3	Tetsu Inoue	Digital Fiction	TRWWNCC146451EE0A4
32	Tetsu Inoue	1998	Waterloo Terminal	6	Tetsu Inoue	Hi-Fi Static	TRCXFAX1464520B165
33	Tetsu Inoue	1998	Waterloo Terminal	5	Tetsu Inoue	Synthetic Doom	TRZFLPE14645224052
34	Tetsu Inoue	1998	Waterloo Terminal	4	Tetsu Inoue	Arc Texture	TRJJDGL14645239F7A
35	Tetsu Inoue	1998	Waterloo Terminal	2	Tetsu Inoue	DSP Terminal	TRKFAPN14645256552
36	Tetsu Inoue	1998	Psycho-Acoustic	5	Tetsu Inoue	Modu Lotion	TRVQDUG1464526ECF0
37	Tetsu Inoue	1998	Psycho-Acoustic	8	Tetsu Inoue	Plug	TRVDJLT1464528125A
38	Tetsu Inoue	1998	Psycho-Acoustic	2	Tetsu Inoue	Tonic Bit	TRHRWIH1464529CB54
39	Tetsu Inoue	1998	Psycho-Acoustic	4	Tetsu Inoue	Psycho Plastic	TRFZTYR146452B4400
40	Tetsu Inoue	1998	Psycho-Acoustic	3	Tetsu Inoue	Dot Hack	TREMVAP146452D2CD2
41	Tetsu Inoue	1995	Slow and Low	1	Tetsu Inoue	Man Made Heaven	TRESIPL146452FD0FF
42	Tetsu Inoue	1995	Slow and Low	5	Tetsu Inoue	Polychrome Chant	TRSXHCQ146453188F5
43	Tetsu Inoue	1995	Slow and Low	2	Tetsu Inoue	Static Soul	TRVBRJZ1464534A97A
44	Tetsu Inoue	1995	Slow and Low	3	Tetsu Inoue	Automatic Motion	TRDSAIM14645371E1B
45	Tetsu Inoue	1995	Slow and Low	6	Tetsu Inoue	Speculative Vision	TRXAUGA1464539A51E
46	Tetsu Inoue	1994	Ambiant Otaku	1	Tetsu Inoue	Karmic Light	TRYBWSA146453D814E
47	Tetsu Inoue	1994	Ambiant Otaku	3	Tetsu Inoue	Ambiant Otaku	TRTRSWM14645401D50
48	Tetsu Inoue	1994	Ambiant Otaku	5	Tetsu Inoue	Magnetic Field	TRQEJWR1464543A441
49	Tetsu Inoue	1994	Ambiant Otaku	2	Tetsu Inoue	Low of Vibration	TRAWYKF146454632EA
50	Tetsu Inoue	1994	Ambiant Otaku	4	Tetsu Inoue	Holy Dance	TRKCXKO1464549BAF0
51	Michael Stearns	1984	Planetary Unfolding	6	Michael Stearns	Something's Moving	TRLUKNE146454B5A9E
52	Michael Stearns	1984	Planetary Unfolding	5	Michael Stearns	As the Earth Kissed the Moon	TROEFJO146454E4D94
53	Michael Stearns	1984	Planetary Unfolding	3	Michael Stearns	Wherever Two or More Are Gathered...	TRGYSOE1464550F85B
54	Michael Stearns	1984	Planetary Unfolding	4	Michael Stearns	Life in the Gravity Well	TRUSQTH14645531063
55	Michael Stearns	1984	Planetary Unfolding	2	Michael Stearns	Toto, I've a Feeling We're Not in Kansas Anymore	TRPYHKV14645551143
56	Michael Stearns	1988	Encounter: A Journey in the Key of Space	10	Michael Stearns	Star Dreams (Peace Eternal)	TRGICXO1464556A4B5
57	Michael Stearns	1988	Encounter: A Journey in the Key of Space	7	Michael Stearns	Distant Thunder (Solitary Witness)	TREDDWX1464558F671
58	Michael Stearns	1988	Encounter: A Journey in the Key of Space	6	Michael Stearns	Within (Choir of the Ascending Spirit)	TRAPDGF146455A9AFF
59	Michael Stearns	1988	Encounter: A Journey in the Key of Space	5	Michael Stearns	Dimensional Shift (Across the Threshold)	TRJZKMW146455CC6BD
60	Michael Stearns	1988	Encounter: A Journey in the Key of Space	3	Michael Stearns	The Beacon (Those Who Have Gone Before)	TRHVBFL146455F7FC2
61	Michael Stearns	1983	Lyra	8	Michael Stearns	Return	TRZDCZO14645625D89
62	Michael Stearns	1983	Lyra	1	Michael Stearns	Arrival	TRERKIZ14645644C5C
63	Michael Stearns	1983	Lyra	3	Michael Stearns	Intervals and Echoes	TRXERJH1464565B225
64	Michael Stearns	1983	Lyra	7	Michael Stearns	The Dragon's Dream World	TREKRAF146456786BE
65	Michael Stearns	1983	Lyra	6	Michael Stearns	Invocation	TREPNPW14645688388
66	Michael Stearns	1993	Sacred Site	5	Michael Stearns	Land Light	TRRFTXJ146456A4EA2
67	Michael Stearns	1993	Sacred Site	6	Michael Stearns	Sacred Site Soundtrack	TRBFQOZ146456CE152
68	Michael Stearns	1993	Sacred Site	2	Michael Stearns	Baraka Theme	TRMXBSP146456F3B67
69	Michael Stearns	1993	Sacred Site	7	Michael Stearns	Twin Flame	TREFLSG14645711AC3
70	Michael Stearns	1993	Sacred Site	4	Michael Stearns	Tropical Rain Forest	TRMVLVY14645734D30
71	Michael Stearns	1987	Floating Whispers	8	Michael Stearns	Floating Whispers	TRZEUPW1464573EAF0
72	Michael Stearns	1987	Floating Whispers	5	Michael Stearns	A Moment Before	TRUUQBH1464574AF5F
73	Michael Stearns	1987	Floating Whispers	1	Michael Stearns	Spanish Twilight	TRPUUZV146457602C0
74	Michael Stearns	1987	Floating Whispers	4	Michael Stearns	At the Bath	TRTAMKQ1464577C133
75	Michael Stearns	1987	Floating Whispers	6	Michael Stearns	The Reflecting Heart	TREFOGC146457963EB

Table B.2: Content of the IDM Ambient Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
76	Tycho	2002	The Science of Patterns	5	Tycho	In the End	TRWOJBC14648F15DBF
77	Tycho	2002	The Science of Patterns	3	Tycho	Red Bridge	TRFQFAM14648F35F15
78	Tycho	2002	The Science of Patterns	1	Tycho	Dream as Memory	TRVOPPN14648F492B7
79	Tycho	2002	The Science of Patterns	4	Tycho	Systems	TRWWAHK14648F6E880
80	Tycho	2002	The Science of Patterns	2	Tycho	Human Condition	TRSVCZB14648F894AC
81	Tycho	2011	Dive	10	Tycho	Elegy	TRFKATJ13B73068025
82	Tycho	2011	Dive	5	Tycho	Coastal Brake	TRDFHLW13B9C8BC41F
83	Tycho	2011	Dive	3	Tycho	Daydream	TRAXCSA139F12DABD9
84	Tycho	2011	Dive	1	Tycho	A Walk	TRIKYCF137600CB327
85	Tycho	2011	Dive	9	Tycho	Epigram	TRLZAPW13B74EC3F16
86	Tycho	2014	Awake	3	Tycho	L	TRCMWNN145D49120F4
87	Tycho	2014	Awake	6	Tycho	Apogee	TRAKACU14648FCA070
88	Tycho	2014	Awake	7	Tycho	Spectre	TRBVIZZ14648FE053A
89	Tycho	2014	Awake	8	Tycho	Plains	TRCCLCO14648FF55F7
90	Tycho	2014	Awake	4	Tycho	Dye	TRFODMS14649015DF2
91	Tycho	2004	Sunrise Projector	3	Tycho	PBS/KAE	TRIZGDY1464902EF9A
92	Tycho	2004	Sunrise Projector	11	Tycho	Cloud Generator	TRXLJJA146490449F8
93	Tycho	2004	Sunrise Projector	6	Tycho	Lapse	TRXBNAT1464905F3A1
94	Tycho	2004	Sunrise Projector	9	Tycho	You Should Be More Like Your Brother	TRGEBGZ14649069721
95	Tycho	2004	Sunrise Projector	8	Tycho	Past Is Prologue	TRACDIL146490832C5
96	Tycho	2006	Past Is Prologue	7	Tycho	A Circular Reeducation	TRQXAVH1464909D2C3
97	Tycho	2006	Past Is Prologue	6	Tycho	Brother	TRZIEZS146490AA18C
98	Tycho	2006	Past Is Prologue	5	Tycho	Send and Receive	TRLMGGF146490BE610
99	Tycho	2006	Past Is Prologue	1	Tycho	From Home	TRGFDNS146490E7482
100	Tycho	2006	Past Is Prologue	10	Tycho	The Disconnect	TRWBUQQ13CB939A6A3
326	ISAN	2001	Lucky Cat	10	ISAN	Caddis	TRJPSIH146C09940C4
327	ISAN	2001	Lucky Cat	12	ISAN	You Can Use Bamboo as a Ruler	TROTEAY146C0A5A9EE
328	ISAN	2001	Lucky Cat	3	ISAN	Fueled	TRCVVUM146C0B25CCF
329	ISAN	2001	Lucky Cat	9	ISAN	Cathart	TRGZRLW146C58B3AC0
330	ISAN	2001	Lucky Cat	11	ISAN	Scraph	TRALRFM146C58EA6D7
331	ISAN	1998	Beautronics	16	ISAN	Anklet	TRVDKEU146C592282D
332	ISAN	1998	Beautronics	5	ISAN	Ampule	TRMXLNV146C595EA1C
333	ISAN	1998	Beautronics	7	ISAN	Bolselin	TRXVKKX146C5998558
334	ISAN	1998	Beautronics	14	ISAN	Tint7-Bloody Mary	TRGULJK146C59C4CD7
335	ISAN	1998	Beautronics	11	ISAN	Skeek	TRQHUIK146C59E2213
336	ISAN	2006	Plans Drawn in Pencil	13	ISAN	Ruined Feathers	TRALFMR146C5A1ECC8
337	ISAN	2006	Plans Drawn in Pencil	9	ISAN	Corundum	TRSFOPT146C5A43BF9
338	ISAN	2006	Plans Drawn in Pencil	11	ISAN	Seven Mile Marker	TRDFHDL146C5A74DD1
339	ISAN	2006	Plans Drawn in Pencil	12	ISAN	Working in Dust	TREXVAD146C5AA9633
340	ISAN	2006	Plans Drawn in Pencil	8	ISAN	Five to Four, Ten to Eleven	TRBANOR146C5ADABA3
341	ISAN	2002	Clockwork Menagerie	9	ISAN	Phoeb	TRBJYVW147C72E62A1
342	ISAN	2002	Clockwork Menagerie	5	ISAN	Cubillo	TRNEAOX146C5B5E2C3
343	ISAN	2002	Clockwork Menagerie	1	ISAN	Autolung	TROSXTQ146C5B8A8B3
344	ISAN	2002	Clockwork Menagerie	8	ISAN	Eusa's Head	TRFGTGW146C5BD5376
345	ISAN	2002	Clockwork Menagerie	7	ISAN	Eeriel	TRZFKDT147C72957B9
346	ISAN	2004	Meet Next Life	7	ISAN	Iron Eyes	TRLEKVO146C5C4892E
347	ISAN	2004	Meet Next Life	6	ISAN	Gunnera	TRGBUJY146C5C8647A
348	ISAN	2004	Meet Next Life	2	ISAN	First Date - Jumble Sale	TRMJWWT146C5CC9E32
349	ISAN	2004	Meet Next Life	11	ISAN	Slow Bulb Slippage	TRBSUBD146C5D111C7
350	ISAN	2004	Meet Next Life	1	ISAN	Birds Over Barges	TRCBIDA146C5D56971
351	Monolake	2003	Momentum	5	Monolake	Tetris	TRUJSBS146C5DBA478
352	Monolake	2003	Momentum	3	Monolake	Atomium	TRNFLLE146C5E42514
353	Monolake	2003	Momentum	7	Monolake	Reminiscence	TRBGJVJ146C5EDCC1F
354	Monolake	2003	Momentum	8	Monolake	Stratosphere (edit)	TREYYIR147C99F5D1B
355	Monolake	2003	Momentum	1	Monolake	Cern	TRSKTUI146C6006BC4
356	Monolake	1999	Interstate	11	Monolake	Terminal	TRYTREB146C6098720
357	Monolake	1999	Interstate	9	Monolake	[untitled]	TRZDQQU146C610DEE2
358	Monolake	1999	Interstate	4	Monolake	[untitled]	TRFXRUD146C6167D4C
359	Monolake	1999	Interstate	7	Monolake	Perpetuum	TROKWAE146C61CA10D
360	Monolake	1999	Interstate	3	Monolake	Gecko	TRIDKOS146C625A831
361	Monolake	2009	Silence	5	Monolake	Avalanche	TRKRQJL146C6296D4F
362	Monolake	2009	Silence	1	Monolake	Watching Clouds	TRGYDYJ146C62F4E9B
363	Monolake	2009	Silence	4	Monolake	Far Red	TRRUACV146C6330210
364	Monolake	2009	Silence	7	Monolake	Internal Clock	TRIDXJZ146C63B104F
365	Monolake	2009	Silence	2	Monolake	Infinite Snow	TRONCEH146C644A513
366	Monolake	2001	Gravity	8	Monolake	Nucleus	TRGFQNA146C64AFFF2
367	Monolake	2001	Gravity	3	Monolake	Frost	TRPNYII146C656D8A9
368	Monolake	2001	Gravity	1	Monolake	Mobile	TRFXJAU146C65F87FD
369	Monolake	2001	Gravity	2	Monolake	Ice	TRUKMWS146C66A8580
370	Monolake	2001	Gravity	5	Monolake	Zero Gravity	TRMXTQN146C6758FB4
371	Monolake	2001	Cinemascope	10	Monolake	Indigo	TRXHUUU146C67BAA9B
372	Monolake	2001	Cinemascope	5	Monolake	Ionized	TRVTQSE146C684617F
373	Monolake	2001	Cinemascope	6	Monolake	Remoteable	TRUHUWS146C69B64DA
374	Monolake	2001	Cinemascope	8	Monolake	Cut	TRSNHGT146C69FD96A
375	Monolake	2001	Cinemascope	3	Monolake	Cubicle	TRWIEHZ146C6A5EC43

Table B.3: Content of the Post-Rock Ambient Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
101	The Album Leaf	2006	Into the Blue Again	1	The Album Leaf	The Light	TROUJID1464910527D
102	The Album Leaf	2006	Into the Blue Again	10	The Album Leaf	Broken Arrow	TRXOZWI146C3D85D70
103	The Album Leaf	2006	Into the Blue Again	3	The Album Leaf	Shine	TRGSZGY1464913B70F
104	The Album Leaf	2006	Into the Blue Again	9	The Album Leaf	Wishful Thinking	TRHSZLZ14649153D9D
105	The Album Leaf	2006	Into the Blue Again	7	The Album Leaf	Into the Sea	TRMPZDF14649169F0B
106	The Album Leaf	2007	The Enchanted Hill	2	The Album Leaf	Fear of Flying	TREUMWL1464918D4DF
107	The Album Leaf	2007	The Enchanted Hill	5	The Album Leaf	Kevlar	TRKJQOK146491A2AD0
108	The Album Leaf	2007	The Enchanted Hill	6	The Album Leaf	San Simeon	TRYHIXC146491BC4FB
109	The Album Leaf	2007	The Enchanted Hill	3	The Album Leaf	Drawing Mountains	TRVMNLV146491CDDDC
110	The Album Leaf	2007	The Enchanted Hill	4	The Album Leaf	Enchanted Hill	TRWHMPF146491E5361
111	The Album Leaf	1999	An Orchestrated Rise to Fall	8	The Album Leaf	A Short Story	TRXPWUR146492393A2
112	The Album Leaf	1999	An Orchestrated Rise to Fall	5	The Album Leaf	We Once Were (One)	TRCZMMU1464925093C
113	The Album Leaf	1999	An Orchestrated Rise to Fall	4	The Album Leaf	September Song	TRCQRGV146C3C8202C
114	The Album Leaf	1999	An Orchestrated Rise to Fall	1	The Album Leaf	Wander	TRXNZET147CA847600
115	The Album Leaf	1999	An Orchestrated Rise to Fall	9	The Album Leaf	We Once Were (Two)	TRONTAP1464927906E
116	The Album Leaf	2001	One Day I'll Be On Time	5	The Album Leaf	The Audio Pool	TRGXQNM1464928DCEC
117	The Album Leaf	2001	One Day I'll Be On Time	12	The Album Leaf	Glimmer	TRFBRGV146492A1A50
118	The Album Leaf	2001	One Day I'll Be On Time	11	The Album Leaf	Vermillion	TRNIDZD146492BACBE
119	The Album Leaf	2001	One Day I'll Be On Time	9	The Album Leaf	Asleep	TRPANDX146492CFFE0
120	The Album Leaf	2001	One Day I'll Be On Time	7	The Album Leaf	In Between Lines	TRWUXWX146492E32D4
121	The Album Leaf	2004	In a Safe Place	4	The Album Leaf	TwentyTwoFourteen	TRWKSLO146492FD174
122	The Album Leaf	2004	In a Safe Place	2	The Album Leaf	Thule	TRQGFWS14767C0AEFC
123	The Album Leaf	2004	In a Safe Place	5	The Album Leaf	The Outer Banks	TRGHSKS14767C6573F
124	The Album Leaf	2004	In a Safe Place	1	The Album Leaf	Window	TRPSWPI146C3CA7855
125	The Album Leaf	2004	In a Safe Place	7	The Album Leaf	Another Day (Revised)	TRVSUQQ146C3CD72E8
126	The American Dollar	2006	The American Dollar	10	The American Dollar	Everyone Gets Shot	TRVJOSD14649391125
127	The American Dollar	2006	The American Dollar	3	The American Dollar	Cambian	TREEWGK146493A7B86
128	The American Dollar	2006	The American Dollar	2	The American Dollar	Glow	TRXATVE146493CE492
129	The American Dollar	2006	The American Dollar	8	The American Dollar	Thompson	TRDJUQU146493DAF13
130	The American Dollar	2006	The American Dollar	6	The American Dollar	Separate but Equal	TROARDR146493EDF3D
131	The American Dollar	2010	Atlas	11	The American Dollar	Frontier Melt	TRPYPAB14649408E65
132	The American Dollar	2010	Atlas	10	The American Dollar	Second Sight	TRKBFMN1464941F9E3
133	The American Dollar	2010	Atlas	4	The American Dollar	Shadows	TRAZTDA146494389BE
134	The American Dollar	2010	Atlas	3	The American Dollar	Fade in Out	TRZRBMB14649456F22
135	The American Dollar	2010	Atlas	2	The American Dollar	Age of Wonder	TRLQRVL1464947354F
136	The American Dollar	2008	A Memory Stream	8	The American Dollar	Our Hearts Are Read	TRIXUYK1464948CF47
137	The American Dollar	2008	A Memory Stream	3	The American Dollar	Call	TRINSLK146494ADC25
138	The American Dollar	2008	A Memory Stream	7	The American Dollar	Transcendence	TRHTOBW146494C44ED
139	The American Dollar	2008	A Memory Stream	2	The American Dollar	The Slow Wait, Part 2	TRFQWVI146494DC705
140	The American Dollar	2008	A Memory Stream	6	The American Dollar	Lights Dim	TRAURCV146494F5B6E
141	The American Dollar	2012	Awake in the City	3	The American Dollar	Ether Channels	TRZHEFS1464950939B
142	The American Dollar	2012	Awake in the City	12	The American Dollar	Oracle	TRSILQZ1464951C18A
143	The American Dollar	2012	Awake in the City	4	The American Dollar	First Day	TRQRMNH1464952EE27
144	The American Dollar	2012	Awake in the City	5	The American Dollar	Steeltown, Part 1	TRLQJKV14649539FAE
145	The American Dollar	2012	Awake in the City	1	The American Dollar	Faces in the Haze	TRXAJME1464954F5C2
146	The American Dollar	2007	The Technicolour Sleep	3	The American Dollar	Signaling Through the Flames	TREHKKF1464956FB3F
147	The American Dollar	2007	The Technicolour Sleep	12	The American Dollar	Palestine	TRQSUCN1464957FC30
148	The American Dollar	2007	The Technicolour Sleep	7	The American Dollar	Supernova Landslide	TRCRDIR14649599AC4
149	The American Dollar	2007	The Technicolour Sleep	11	The American Dollar	Raided by Waves	TRHYOHF146495B5B13
150	The American Dollar	2007	The Technicolour Sleep	9	The American Dollar	Summer of War	TRHWPKD146495CBB94
226	Boards of Canada	1998	Music Has the Right to Children	10	Boards of Canada	Roygbiv	TRPVDFW1308B66B4D2
227	Boards of Canada	1998	Music Has the Right to Children	13	Boards of Canada	Olson	TRWGMXV146BFED18BE
228	Boards of Canada	1998	Music Has the Right to Children	16	Boards of Canada	Open the Light	TRDVVTK146BFEEE193
229	Boards of Canada	1998	Music Has the Right to Children	6	Boards of Canada	Sixtyten	TRYWXMV147C6D7E733
230	Boards of Canada	1998	Music Has the Right to Children	2	Boards of Canada	An Eagle in Your Mind	TRFTZAX147C6D14034
231	Boards of Canada	2002	Geogaddi	13	Boards of Canada	Opening the Mouth	TROJYVT13C6DAC17B8
232	Boards of Canada	2002	Geogaddi	19	Boards of Canada	Dawn Chorus	TRVEHXR13965A6FD8F
233	Boards of Canada	2002	Geogaddi	12	Boards of Canada	The Beach at Redpoint	TRHNIPW13C6DAC697C
234	Boards of Canada	2002	Geogaddi	16	Boards of Canada	The Devil Is in the Details	TRFXAXQ13C6DACDED5
235	Boards of Canada	2002	Geogaddi	18	Boards of Canada	Over the Horizon Radar	TRFXNSG13C6DACCDA0
236	Boards of Canada	1995	Twoism (2002 CD)	3	Boards of Canada	Iced Cooly	TRNWVCU146C00EDB2C
237	Boards of Canada	1995	Twoism (2002 CD)	1	Boards of Canada	Sixtyniner	TRBKYEF13D26D7790D
238	Boards of Canada	1995	Twoism (2002 CD)	7	Boards of Canada	Melissa Juice	TROIHSG13D26DA9ECF
239	Boards of Canada	1995	Twoism (2002 CD)	2	Boards of Canada	Oirectine	TRFOTVP13D26D8B7F9
240	Boards of Canada	1995	Twoism (2002 CD)	4	Boards of Canada	Basefree	TRTQKHQ13D26DA3882
241	Boards of Canada	2005	The Campfire Headphase	12	Boards of Canada	Constants Are Changing	TRUSFDO146C01F22CC
242	Boards of Canada	2005	The Campfire Headphase	2	Boards of Canada	Chromakey Dreamcoat	TRCYZOK13C6DB05731
243	Boards of Canada	2005	The Campfire Headphase	15	Boards of Canada	Farewell Fire	TRGXYOA13C6DB56353
244	Boards of Canada	2005	The Campfire Headphase	7	Boards of Canada	'84 Pontiac Dream	TRPVPUQ13C6DB666A5
245	Boards of Canada	2005	The Campfire Headphase	11	Boards of Canada	Hey Saturday Sun	TRIZPOD13C6DB7D6BE
246	Boards of Canada	1996	Hi-Scores	4	Boards of Canada	June 9th	TRAEGRH13D2A6FA5FC
247	Boards of Canada	1996	Hi-Scores	1	Boards of Canada	Hi Scores	TREQDTC13D26E0D5DB
248	Boards of Canada	1996	Hi-Scores	3	Boards of Canada	Nlogax	TRJXVHT13D26DE8AA8
249	Boards of Canada	1996	Hi-Scores	6	Boards of Canada	Everything You Do Is a Balloon	TRHYFTK13B048B1BD5
250	Boards of Canada	1996	Hi-Scores	5	Boards of Canada	Seeya Later	TRCAUHL13D26E4B9E9

Table B.4: Content of the IDM Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
251	Aphex Twin	1996	Richard D. James Album	6	Aphex Twin	To Cure a Weakling Child	TRBDAAS146C0473556
252	Aphex Twin	1996	Richard D. James Album	1	Aphex Twin	4	TRSXPQC146C04AABB5
253	Aphex Twin	1996	Richard D. James Album	5	Aphex Twin	Carn Marth	TRLSLRP146C04D535E
254	Aphex Twin	1996	Richard D. James Album	8	Aphex Twin	Yellow Calx	TRBAGDV146C04FD3F5
255	Aphex Twin	1996	Richard D. James Album	12	Aphex Twin	Inkey$	TRBFMNB146C0525751
256	Aphex Twin	1992	Selected Ambient Works 85-92	4	Aphex Twin	Ageispolis	TRZVWOV146C05508E6
257	Aphex Twin	1992	Selected Ambient Works 85-92	13	Aphex Twin	Actium	TRFBKWE146C0591331
258	Aphex Twin	1992	Selected Ambient Works 85-92	6	Aphex Twin	Green Calx	TRQXRIN146C05F90D5
259	Aphex Twin	1992	Selected Ambient Works 85-92	7	Aphex Twin	Heliosphan	TRGGPHG146C065E103
260	Aphex Twin	1992	Selected Ambient Works 85-92	8	Aphex Twin	We Are the Music Makers	TRCDQPM146C06A1F98
261	Aphex Twin	2001	Drukqs	6	Aphex Twin	Gwely Mernans	TRZNEAG146C0706B03
262	Aphex Twin	2001	Drukqs	5	Aphex Twin	Strotha Tynhe	TRYVDGH146C073AD56
263	Aphex Twin	2001	Drukqs	15	Aphex Twin	Kesson Dalef	TRFGAYE146C0752370
264	Aphex Twin	2001	Drukqs	20	Aphex Twin	Meltphace 6	TRPRGRL146C077429D
265	Aphex Twin	2001	Drukqs	8	Aphex Twin	Cock/Ver10	TRTOIBW146C07D968B
266	Aphex Twin	1997	Come To Daddy	8	Aphex Twin	IZ-US	TRNOQGG146C0828095
267	Aphex Twin	1997	Come To Daddy	5	Aphex Twin	To Cure a Weakling Child (Contour Regard)	TRCWCNC146C48D3812
268	Aphex Twin	1997	Come To Daddy	2	Aphex Twin	Flim	TROVMEU146C08879E0
269	Aphex Twin	1997	Come To Daddy	1	Aphex Twin	Come to Daddy (Pappy Mix)	TRPHCTT146C08AFE62
270	Aphex Twin	1997	Come To Daddy	4	Aphex Twin	Bucephalus Bouncing Ball	TRIGTRS146C08FD1AC
271	Aphex Twin	1995	...I Care Because You Do	3	Aphex Twin	Wax the Nip	TRESMIF146C0955C08
272	Aphex Twin	1995	...I Care Because You Do	5	Aphex Twin	Ventolin (Video Version)	TRBQONZ13C5E2982F5
273	Aphex Twin	1995	...I Care Because You Do	1	Aphex Twin	Acrid Avid Jam Shred	TRAUADW146C09D72CF
274	Aphex Twin	1995	...I Care Because You Do	11	Aphex Twin	Cow Cud Is a Twin	TROKVMR146C0ABB59A
275	Aphex Twin	1995	...I Care Because You Do	4	Aphex Twin	Icct Hedral (Edit)	TRTLJYL146C0B6C4CA
276	Chaos A.D.	1998	Buzz Caner	2	Squarepusher	Mess Head	TRGPIRA146C0BC410D
277	Chaos A.D.	1998	Buzz Caner	8	Squarepusher	Psultan, Part 1	TRSFOQZ146C0C3092F
278	Chaos A.D.	1998	Buzz Caner	5	Squarepusher	Dreaded Pestilence	TRMGRMR147C6B15017
279	Chaos A.D.	1998	Buzz Caner	3	Squarepusher	Bioslate	TRVHGFR146C0CC6474
280	Chaos A.D.	1998	Buzz Caner	7	Squarepusher	Friend Track	TRFIXMU146C0D29D31
281	Squarepusher	2006	Hello Everything	6	Squarepusher	Circlewave 2	TRIFTTT146C0D6B918
282	Squarepusher	2006	Hello Everything	1	Squarepusher	Hello Meow	TRZEIJQ146C0D9EBF1
283	Squarepusher	2006	Hello Everything	9	Squarepusher	Welcome to Europe	TRTCPLZ146C0DE269F
284	Squarepusher	2006	Hello Everything	10	Squarepusher	Plotinus	TRMLZJB146C0E2F818
285	Squarepusher	2006	Hello Everything	5	Squarepusher	Vacuum Garden	TREUHVS146C0EABAEF
286	Squarepusher	2001	Go Plastic	8	Squarepusher	Tommib	TRAKJUS146C268F819
287	Squarepusher	2001	Go Plastic	3	Squarepusher	Go! Spastic	TRIALSY146C284971D
288	Squarepusher	2001	Go Plastic	9	Squarepusher	My Fucking Sound	TROWGOM146C28C44F7
289	Squarepusher	2001	Go Plastic	2	Squarepusher	Boneville Occident	TRMLHRN146C4977E62
290	Squarepusher	2001	Go Plastic	4	Squarepusher	Metteng Excuske v1.2	TRLMNCN146C297245C
291	Squarepusher	1996	Feed Me Weird Things	1	Squarepusher	Squarepusher Theme	TRJJXNI146C298611E
292	Squarepusher	1996	Feed Me Weird Things	10	Squarepusher	U.F.O.'s Over Leytonstone	TRQRAQN146C29B3E05
293	Squarepusher	1996	Feed Me Weird Things	8	Squarepusher	Goodnight Jade	TRVWKKU146C29F1A43
294	Squarepusher	1996	Feed Me Weird Things	6	Squarepusher	Windscale 2	TRNYTVF146C2A17049
295	Squarepusher	1996	Feed Me Weird Things	3	Squarepusher	The Swifty	TRAOGDB146C2A5DA2F
296	Squarepusher	2012	Ufabulum	2	Squarepusher	Unreal Square	TROFPVR13B2EEF0510
297	Squarepusher	2012	Ufabulum	1	Squarepusher	4001	TRCJOIE13B00B44C74
298	Squarepusher	2012	Ufabulum	5	Squarepusher	Red in Blue	TRBYEYU13B41214E7A
299	Squarepusher	2012	Ufabulum	9	Squarepusher	303 Scopem Hard	TRHGVFC146C2BB6E82
300	Squarepusher	2012	Ufabulum	6	Squarepusher	The Metallurgist	TRFEHTH146C2C0889E
301	Autechre	2001	Confield	8	Autechre	Uviol	TREPGUM145A60B2885
302	Autechre	2001	Confield	6	Autechre	Bine	TRBFWKI146C2CE7964
303	Autechre	2001	Confield	4	Autechre	Sim Gishel	TRAOHQB145A60944DE
304	Autechre	2001	Confield	1	Autechre	VI Scose Poise	TRFJKYL146C2DE456E
305	Autechre	2001	Confield	3	Autechre	Pen Expers	TRHYBPH146C2E5664D
306	Autechre	1994	Amber	6	Autechre	Piezo	TRMCTSR146C2F0C5BD
307	Autechre	1994	Amber	1	Autechre	Foil	TRLEQOJ146C2F8AE72
308	Autechre	1994	Amber	3	Autechre	Silverside	TRYRYME145A5EBCE72
309	Autechre	1994	Amber	10	Autechre	Nil	TRVYUES145A5F0F064
310	Autechre	1994	Amber	2	Autechre	Montreal	TRULENB146C30E135D
311	Autechre	1995	Tri Repetae	10	Autechre	Rsdio	TRAJGGT146C317B03F
312	Autechre	1995	Tri Repetae	3	Autechre	Leterel	TRZSPND146C3237068
313	Autechre	1995	Tri Repetae	9	Autechre	Overand	TRPMYKI146C32C8014
314	Autechre	1995	Tri Repetae	5	Autechre	Stud	TRJIHKU146C333E56A
315	Autechre	1995	Tri Repetae	1	Autechre	Dael	TRQAQKK146C33E4ECB
316	Autechre	1997	Chiastic Slide	5	Autechre	Hub	TRTFJYN145A5FFF619
317	Autechre	1997	Chiastic Slide	8	Autechre	Pule	TRMULPV146C34FEEC6
318	Autechre	1997	Chiastic Slide	4	Autechre	Cichli	TRERGBJ145791988CB
319	Autechre	1997	Chiastic Slide	6	Autechre	Calbruc	TROCIPS146C362914E
320	Autechre	1997	Chiastic Slide	3	Autechre	Tewe	TRQLDIR145A5FE67FE
321	Autechre	1998	LP5	1	Autechre	Acroyear2	TRKOSZO146C3AF1CD8
322	Autechre	1998	LP5	2	Autechre	777	TRHGTLD146C388AFA1
323	Autechre	1998	LP5	8	Autechre	Corc	TRNBCSJ146C38EEDAD
324	Autechre	1998	LP5	5	Autechre	Vose In	TRQHSFC146C39447BA
325	Autechre	1998	LP5	4	Autechre	Melve	TRSTGQC146C398B4BD

Table B.5: Content of the Nu Jazz Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
376	Bonobo	2003	Dial 'M' for Monkey	5	Bonobo	Wayward Bob	TRGSYDU146C3EB1BE0
377	Bonobo	2003	Dial 'M' for Monkey	6	Bonobo	Pick Up	TRDCZNI146C3EF57E6
378	Bonobo	2003	Dial 'M' for Monkey	7	Bonobo	Something for Windy	TRPNKPD146C3F299DC
379	Bonobo	2003	Dial 'M' for Monkey	9	Bonobo	Light Pattern	TRVREAO146C3F4FC16
380	Bonobo	2003	Dial 'M' for Monkey	2	Bonobo	Flutter	TRXVDBL146C3F9A218
381	Bonobo	2013	The North Borders	6	Bonobo	Jets	TRCAFEL13DA35045F1
382	Bonobo	2013	The North Borders	8	Bonobo	Don't Wait	TRFFLEL13DA328A560
383	Bonobo	2013	The North Borders	3	Bonobo	Cirrus	TRBPBXX13DA345FFCB
384	Bonobo	2013	The North Borders	11	Bonobo	Ten Tigers	TRMKXES13DA332AC21
385	Bonobo	2013	The North Borders	10	Bonobo	Antenna	TRJDCCK13DA32E47A2
386	Bonobo	2000	Animal Magic	4	Bonobo	Kota	TRTJOUL147C6F30799
387	Bonobo	2000	Animal Magic	5	Bonobo	Terrapin	TRALMRC146C4195C2A
388	Bonobo	2000	Animal Magic	7	Bonobo	Shadowtricks	TRAIZXO146C41D33EF
389	Bonobo	2000	Animal Magic	3	Bonobo	Dinosaurs	TRTUNQK147C6EE8DDE
390	Bonobo	2000	Animal Magic	10	Bonobo	Silver	TRXCYAO147C6F92B9C
391	Bonobo	2010	Black Sands	2	Bonobo	Kiara	TRQGVIC13B0136AB55
392	Bonobo	2010	Black Sands	11	Bonobo	Animals	TRBAPGK147C706A294
393	Bonobo	2010	Black Sands	6	Bonobo	We Could Forever	TRJJSPS146C43339B4
394	Bonobo	2010	Black Sands	1	Bonobo	Prelude	TRUYFOR146C4365407
395	Bonobo	2010	Black Sands	5	Bonobo	El Toro	TRGVNDB144126BBCBE
396	Bonobo	2006	Days to Come	4	Bonobo	The Fever	TRHZADZ146C43C4DFE
397	Bonobo	2006	Days to Come	8	Bonobo	On Your Marks	TRRFYXO146C4406C03
398	Bonobo	2006	Days to Come	7	Bonobo	Transmission 94, Parts 1 & 2	TRWXQMY146C445C48F
399	Bonobo	2006	Days to Come	11	Bonobo	Recurring	TRLYYDW146C44CF582
400	Bonobo	2006	Days to Come	5	Bonobo	Ketto	TRICVEE146C451B13C
401	Lemongrass	2001	Voyage au Centre de la Terre	2	Lemongrass	Nightingales	TRFFUTS146C456235D
402	Lemongrass	2001	Voyage au Centre de la Terre	11	Lemongrass	Je Sais	TRVYICP147CAEF2C2B
403	Lemongrass	2001	Voyage au Centre de la Terre	7	Lemongrass	La Mer	TRWAQJN146C45F46C3
404	Lemongrass	2001	Voyage au Centre de la Terre	1	Lemongrass	Falling Star	TRKAPYS146C4632378
405	Lemongrass	2001	Voyage au Centre de la Terre	4	Lemongrass	Marchant	TRSZXXT147CAEA994D
406	Lemongrass	1998	Drumatic Universe	11	Lemongrass	Little Alien	TRAFKIV146C46CE3F2
407	Lemongrass	1998	Drumatic Universe	5	Lemongrass	Why?	TROVCSH147CAC3B560
408	Lemongrass	1998	Drumatic Universe	6	Lemongrass	Tell It to My Heart	TRVYHDX147CACA204D
409	Lemongrass	1998	Drumatic Universe	9	Lemongrass	Spira	TRAVFED146C4A98BC0
410	Lemongrass	1998	Drumatic Universe	12	Lemongrass	Wings	TRDIPKK146C4AE9321
411	Lemongrass	2007	Filmotheque	4	Lemongrass	The Camera	TRRHGGI137356E292A
412	Lemongrass	2007	Filmotheque	13	Lemongrass	Les Affaires	TRPMKEU13734CFAA80
413	Lemongrass	2007	Filmotheque	6	Lemongrass	Aloha	TRLKTWD1373570A3BF
414	Lemongrass	2007	Filmotheque	5	Lemongrass	Fritz the Cat	TRJFZOK147CB118786
415	Lemongrass	2007	Filmotheque	8	Lemongrass	Lonely Beach	TRRMVOL137356147FF
416	Lemongrass	2001	Windows	3	Lemongrass	Sunrise on Fujiyama	TRARRNX146C4C79FB0
417	Lemongrass	2001	Windows	16	Lemongrass	Desert Sand	TRMWSSS147CAFCE796
418	Lemongrass	2001	Windows	4	Lemongrass	Winnetou Melody	TRUHMFP146C4D32F56
419	Lemongrass	2001	Windows	6	Lemongrass	Librabelle	TRRGRRI147CAF45725
420	Lemongrass	2001	Windows	13	Lemongrass	Braindance	TRHZTSZ147CAF8BDBA
421	Lemongrass	2003	Skydiver	5	Lemongrass	Pacifique	TRRZDDI147CB02F6A9
422	Lemongrass	2003	Skydiver	2	Lemongrass	A Fabula	TRFZEDB146C4E23257
423	Lemongrass	2003	Skydiver	10	Lemongrass	Bicycles	TRLWDNA146C4E68D6C
424	Lemongrass	2003	Skydiver	12	Lemongrass	Restful Motion	TRKHDHI146C4EB45FF
425	Lemongrass	2003	Skydiver	11	Lemongrass	Mira	TRZTHRM146C4EF10A1
426	Four Tet	2010	There Is Love in You	9	Four Tet	She Just Likes to Fight	TRQSJOM147C71219CC
427	Four Tet	2010	There Is Love in You	6	Four Tet	This Unfolds	TRYKHDP146C4FB8B70
428	Four Tet	2010	There Is Love in You	8	Four Tet	Plastic People	TRELOTI146C4F3AC7E
429	Four Tet	2010	There Is Love in You	7	Four Tet	Reversing	TRIHQTM146C50ABFAE
430	Four Tet	2010	There Is Love in You	3	Four Tet	Circling	TRNWQZN147C71E5A08
431	Four Tet	2003	Rounds	5	Four Tet	Spirit Fingers	TRZRGHQ146C51993DE
432	Four Tet	2003	Rounds	6	Four Tet	Unspoken	TRVMHQX146C51DBFA6
433	Four Tet	2003	Rounds	4	Four Tet	My Angel Rocks Back and Forth	TRYXGUD146C5265CD2
434	Four Tet	2003	Rounds	3	Four Tet	First Thing	TRCLIEL146C52A1C71
435	Four Tet	2003	Rounds	8	Four Tet	As Serious as Your Life	TRVFBVW146C52BF181
436	Four Tet	2005	Everything Ecstatic	7	Four Tet	High Fives	TRPFZGD13B3B0383F3
437	Four Tet	2005	Everything Ecstatic	4	Four Tet	Sun Drums and Soil	TRGLXFS146C5360BD2
438	Four Tet	2005	Everything Ecstatic	2	Four Tet	Smile Around the Face	TRKHVIK146C53CFD35
439	Four Tet	2005	Everything Ecstatic	8	Four Tet	Turtle Turtle Up	TRUAXJX146C541145C
440	Four Tet	2005	Everything Ecstatic	10	Four Tet	You Were There With Me	TRPAAMN13B3AE6D3FC
441	Four Tet	1999	Dialogue	6	Four Tet	Liquefaction	TRUUDGU146C548D533
442	Four Tet	1999	Dialogue	4	Four Tet	3.3 Degrees From the Pole	TRTJORK146C54D1A54
443	Four Tet	1999	Dialogue	1	Four Tet	The Space of Two Weeks	TRMIFGY146C552E5E4
444	Four Tet	1999	Dialogue	8	Four Tet	Calamine	TRLNZNB146C558686D
445	Four Tet	1999	Dialogue	5	Four Tet	Misnomer	TRSFVCP146C55E133F
446	Four Tet	2001	Pause	1	Four Tet	Glue of the World	TRTJNXO146C5612037
447	Four Tet	2001	Pause	3	Four Tet	Harmony One	TRMWUSR146C564B716
448	Four Tet	2001	Pause	2	Four Tet	Twenty Three	TRDOBCF146C5668781
449	Four Tet	2001	Pause	5	Four Tet	Leila Came Round and We Watched a Video	TRDUSXY146C5697348
450	Four Tet	2001	Pause	10	Four Tet	You Could Ruin My Day	TROCWVM146C56BBA45

Table B.6: Content of the Techno Dataset

ID	Album Artist	Year	Album Title	Track	Ground Truth	Title	Echonest ID
151	Jeff Mills	2011	The Power	6	Jeff Mills	Hallucitaions	TRYPUIS146BBBA18C6
152	Jeff Mills	2011	The Power	3	Jeff Mills	The Intruder Emerges	TRBPPYT146BB0F0ED8
153	Jeff Mills	2011	The Power	12	Jeff Mills	Transformation Complete	TROACWC146B731455E
154	Jeff Mills	2011	The Power	1	Jeff Mills	The Power (Theme)	TRBEOXB146BB19A5A3
155	Jeff Mills	2011	The Power	4	Jeff Mills	A Race to Save Thoughts	TRDMSBS146B727D578
156	Jeff Mills	1992	Waveform Transmission, Volume 1	7	Jeff Mills	DNA	TRDZGHG146B78A00A5
157	Jeff Mills	1992	Waveform Transmission, Volume 1	6	Jeff Mills	Late Night	TRIQJUW146B7804F6B
158	Jeff Mills	1992	Waveform Transmission, Volume 1	8	Jeff Mills	Man-Like	TRSSYGW146B7850E82
159	Jeff Mills	1992	Waveform Transmission, Volume 1	5	Jeff Mills	The Hacker	TRAWCDI146B77B573A
160	Jeff Mills	1992	Waveform Transmission, Volume 1	2	Jeff Mills	Jerical	TRTVQCM146BBBBE42B
161	Jeff Mills	1994	Waveform Transmission, Volume 3	8	Jeff Mills	Basic Human Design	TRKGCQM146BB1FAB3F
162	Jeff Mills	1994	Waveform Transmission, Volume 3	1	Jeff Mills	The Extremist	TRDGBUL146B7713270
163	Jeff Mills	1994	Waveform Transmission, Volume 3	2	Jeff Mills	Solid Sleep	TRPMYBN146B76D0DCE
164	Jeff Mills	1994	Waveform Transmission, Volume 3	4	Jeff Mills	Workers	TRPJMUD146B7651AF5
165	Jeff Mills	1994	Waveform Transmission, Volume 3	5	Jeff Mills	Wrath of the Punisher	TRANZVR146B7617E1A
166	Jeff Mills	2011	Fantastic Voyage	11	Jeff Mills	Wait Until He Exhales	TRAGHJQ146BB2C4CBD
167	Jeff Mills	2011	Fantastic Voyage	16	Jeff Mills	The Loss of Power	TRAMGPQ146BB378170
168	Jeff Mills	2011	Fantastic Voyage	8	Jeff Mills	Passing Through the Heart	TRGEZZR146BB43F332
169	Jeff Mills	2011	Fantastic Voyage	2	Jeff Mills	Into the Body (Inner World)	TRWFFXX146B73C8BFE
170	Jeff Mills	2011	Fantastic Voyage	13	Jeff Mills	Blown Away	TRPJGWE146B7493BCF
171	Jeff Mills	2011	2087	3	Jeff Mills	Programmed to Kill	TRQMBEU146BBBDEB7E
172	Jeff Mills	2011	2087	12	Jeff Mills	Mason's Relationship	TRGZRZB146BBBF32B9
173	Jeff Mills	2011	2087	15	Jeff Mills	Free Thinkers (The Reality)	TRQCBVT146BBC02F13
174	Jeff Mills	2011	2087	9	Jeff Mills	Zeller	TRWCWKA146B75C60A0
175	Jeff Mills	2011	2087	8	Jeff Mills	Operation to Free Garth	TRRNPFH146BBC19033
176	Legowelt	2006	Astro Cat Disco	7	Legowelt	Drivin' for Our Love	TRFGPMX147C92E2025
177	Legowelt	2006	Astro Cat Disco	1	Legowelt	BerlinOstbahnhof	TRAEDOW146BBC3D63D
178	Legowelt	2006	Astro Cat Disco	11	Legowelt	Make Your Move	TRXVMUL147C933EF6E
179	Legowelt	2006	Astro Cat Disco	6	Legowelt	Disco Bitch	TRAFJJA146BBC63BA8
180	Legowelt	2006	Astro Cat Disco	13	Legowelt	Strada 83	TRSCSDL146BBC6EAFE
181	Legowelt	2009	Vatos Locos	10	Legowelt	Escape	TREXGRL146BBC7A78F
182	Legowelt	2009	Vatos Locos	5	Legowelt	Aquajam	TRCSUZD146BBC8814F
183	Legowelt	2009	Vatos Locos	11	Legowelt	Topaz Lagoon	TRKWAOB146BBC990B2
184	Legowelt	2009	Vatos Locos	12	Legowelt	Schooldayz	TRXRBWG146BBCA2AD6
185	Legowelt	2009	Vatos Locos	6	Legowelt	Slowjam Deeptechno	TRHHGFS147C9615C5A
186	Legowelt	2008	The Rise and Fall of Manuel Noriega	5	Legowelt	Avianca	TRYCEVC147C9512B54
187	Legowelt	2008	The Rise and Fall of Manuel Noriega	12	Legowelt	Eve of War	TRTIOGC147C954D058
188	Legowelt	2008	The Rise and Fall of Manuel Noriega	1	Legowelt	Capitan Ortega	TRMSISC146BBCDEAA8
189	Legowelt	2008	The Rise and Fall of Manuel Noriega	11	Legowelt	Axis of the Armadillo	TRDXCFO147C9594BB6
190	Legowelt	2008	The Rise and Fall of Manuel Noriega	3	Legowelt	In the Shadow of the Mossad	TRBDDSZ146BBCFC0DC
191	Legowelt	1998	Reports From the Backseat Pimp	7	Legowelt	Disco-tron	TRONMJP146BBD0707F
192	Legowelt	1998	Reports From the Backseat Pimp	2	Legowelt	Swimming Pool	TRKSXNO147C91C5486
193	Legowelt	1998	Reports From the Backseat Pimp	13	Legowelt	Perron Oost (Instrumental)	TRWNQZS147C9213CCE
194	Legowelt	1998	Reports From the Backseat Pimp	10	Legowelt	Gina Fly to Space	TRXOHCS146BBD322B1
195	Legowelt	1998	Reports From the Backseat Pimp	11	Legowelt	Minimal Report	TRHDPFF146BBD3EA0A
196	Legowelt	2008	Dark Days 2	2	Legowelt	HAM Star Flowers	TRKSTWX146BBD4D0FE
197	Legowelt	2008	Dark Days 2	6	Legowelt	Crystobal Theory	TRBFULD146BBD585D8
198	Legowelt	2008	Dark Days 2	5	Legowelt	Manpulse	TRLTISH147C93B3D3A
199	Legowelt	2008	Dark Days 2	7	Legowelt	Chicago Snow Flakes	TRLJQIO146BBD70427
200	Legowelt	2008	Dark Days 2	3	Legowelt	Future Land	TRUQRHE147C9424E0C
201	Plastikman	1998	Consumed	1	Plastikman	Contain	TRBLPYU146B6A62DA3
202	Plastikman	1998	Consumed	8	Plastikman	Locomotion	TRGOYHN146B69BEBC5
203	Plastikman	1998	Consumed	5	Plastikman	Convulse (sic)	TROHNLF146BBD8B365
204	Plastikman	1998	Consumed	6	Plastikman	Ekko	TRPJPYY146B696A121
205	Plastikman	1998	Consumed	9	Plastikman	In Side	TRUOKYC146BBDB809E
206	Plastikman	1993	Sheet One	1	Plastikman	Drp	TRBXQOM147C9BD0242
207	Plastikman	1993	Sheet One	2	Plastikman	Plasticity	TRTOQVP146B66A7D44
208	Plastikman	1993	Sheet One	6	Plastikman	Glob	TRJSBHF147C9C5D938
209	Plastikman	1993	Sheet One	5	Plastikman	Helikopter	TROSMVI146B684BF24
210	Plastikman	1993	Sheet One	8	Plastikman	Koma	TRNIHCM146BBDD530F
211	Plastikman	2003	Closer	5	Plastikman	Slow Poke (Twilight Zone mix)	TRECGWQ146B6C6630C
212	Plastikman	2003	Closer	6	Plastikman	Headcase	TRFJNPL146BBDFED4B
213	Plastikman	2003	Closer	2	Plastikman	Mind Encode	TRCEVBU146BBE19622
214	Plastikman	2003	Closer	10	Plastikman	I Don't Know	TRRXMLW146BBE4690C
215	Plastikman	2003	Closer	7	Plastikman	Ping Pong	TRFGKWZ147CA4D16A7
216	Plastikman	1994	Recycled Plastik	6	Plastikman	Spastik	TRPKPHA146B6E89A30
217	Plastikman	1994	Recycled Plastik	3	Plastikman	Spaz	TRNJZSQ146B6F56E80
218	Plastikman	1994	Recycled Plastik	2	Plastikman	Elektrostatik	TRTYLTL146B710FC45
219	Plastikman	1994	Recycled Plastik	1	Plastikman	Krakpot	TRVJMOV146B702402A
220	Plastikman	1994	Recycled Plastik	5	Plastikman	Naturalistik	TREERAS146B71D9DD9
221	Plastikman	1994	Musik	2	Plastikman	Plastique	TRGMQWR147CA29C334
222	Plastikman	1994	Musik	8	Plastikman	Goo	TRIATJW146B48B2DD5
223	Plastikman	1994	Musik	3	Plastikman	Kriket	TRAPHZU146B655EC47
224	Plastikman	1994	Musik	1	Plastikman	Konception	TRCQNHI146B477DE9D
225	Plastikman	1994	Musik	7	Plastikman	Plasmatik	TRTVKKR146BBE6CA51

Appendix C | Audio Features

In this appendix we include a comprehensive list of the audio features that were extracted from the different software libraries (Essentia (ESS), MIRToolbox (MTB) and the Echonest Analyzer (EN)), organized according to the facet that they represent, as well as the derived descriptors that were computed from them.

C.1 Standard Audio Features

Dynamics

- ESS: Average Loudness
- MTB: Low Energy
- MTB: Low Energy ASR
- MTB: RMS
- MTB: RMS Median
- EN: Loudness
- EN; End of Fade In
- EN: Start of Fade Out
- EN: Energy

Rhythm

- ESS: BPM
- ESS: First Peak BPM
- ESS: First Peak Spread
- ESS: First Peak Weight
- ESS: Second Peak BPM
- ESS: Second Peak Spread
- ESS: Second Peak Weight
- ESS: Onset Rate
- ESS: Beats Loudness [Mean, Var]
- ESS: Beats Loudness Band Ratio [1-6] [Mean, Var]
- MTB: Tempo

- MTB: Event Density
- EN: Tempo
- EN: Tempo Confidence
- EN: Time Signature
- EN: Time Signature Confidence

Timbre

- ESS: MFCC [1-13] [Mean, Var]
- ESS: Zero Crossing Rate [Mean, Var]
- ESS: Spectral Centroid [Mean, Var]
- ESS: Spectral Complexity [Mean, Var]
- ESS: Spectral Flux [Mean, Var]
- ESS: Spectral Kurtosis [Mean, Var]
- ESS: Spectral Rolloff [Mean, Var]
- ESS: Spectral Skewness [Mean, Var]
- ESS: Spectral Spread [Mean, Var]
- ESS: Spectral RMS [Mean, Var]
- ESS: Spectral Strongpeak [Mean, Var]
- ESS: Spectral Contrast [1-6] [Mean, Var]
- ESS: Spectral Contrast Valleys [1-6] [Mean, Var]
- ESS: Dissonance [Mean, Var]
- MTB: MFCC [1-13]
- MTB: Spectral Kurtosis
- MTB: Spectral Rolloff
- MTB: Spectral Skewness
- MTB: Spectral Spread
- MTB: Spectral Entropy
- MTB: Spectral Flatness
- MTB: Roughness [Mean, Median, Std, Slope, Period Frequency, Period Entropy]
- MTB: Irregularity

Tonality and Pitch

- ESS: Pitch [Mean, Var]
- ESS: Pitch Salience [Mean, Var]
- ESS: Number of Pitch Changes
- ESS: Pitch Change Rate
- ESS: Pitch Detection Confidence [Mean, Median, Var]
- ESS: Tuning Frequency [Mean, Var]

- ESS: HPCP [1-36] [Mean, Var]
- MTB: Key
- MTB: Key Strength [1-12]
- MTB: Mode
- MTB: Chroma [1-12]
- MTB: HCDF [Mean, Median, Std, Slope, Period Frequency, Period Entropy]
- MTB: Inharmonicity
- MTB: Tonal Centroid [1-6]
- EN: Key
- EN: Key Confidence
- EN: Mode
- EN: Mode Confidence

SFX

- ESS: Pitch Centroid
- ESS: Pitch After Max to Before Max Energy Ratio
- ESS: Pitch Max to Total
- ESS: Pitch Min to Total

High Level

- ESS: Danceability
- EN: Acousticness
- EN: Danceability
- EN: Valence
- EN: Liveness

C.2 Derived Audio Features

From Rhythm Positions

- ESS: Inter Onset Intervals [Mean, Median, Var]
- ESS: Inter Beat Intervals [Mean, Median, Var]
- MTB: Inter Onset Intervals [Mean, Median, Var]
- EN: Bar Duration [Mean, Median, Var]
- EN: Bar Confidence [Mean, Median, Var]
- EN: Beat Duration [Mean, Median, Var]
- EN: Beat Confidence [Mean, Median, Var]
- EN: Tatum Duration [Mean, Median, Var]
- EN: Tatum Confidence [Mean, Median, Var]

From the Echonest Structural Analysis

- EN: Number of Sections
- EN: Section Rate
- EN: Section Duration [Mean, Median, Var]
- EN: Section Confidence [Mean, Median, Var]
- EN: Loudness Var
- EN: Tempo Var
- EN: Tempo Confidence Var
- EN: Time Signature Var
- EN: Time Signature Confidence Var
- EN: Key Var
- EN: Key Confidence Var
- EN: Mode Var
- EN: Mode Confidence Var
- EN: Number of Segments
- EN: Segment Rate
- EN: Segment Duration [Mean, Median, Var]
- EN: Segment Confidence [Mean, Median, Var]
- EN: Pitch [1-12] [Mean, Median, Var]
- EN: Timbre [1-12] [Mean, Median, Var]

From the Echonest Rhythm Structure Analysis

- EN: Segments per Section [Mean, Median, Var]
- EN: Bars per Segment [Mean, Median, Var]
- EN: Beats per Bar [Mean, Median, Var]
- EN: Tatums per Beat [Mean, Median, Var]