Universitat Pompeu Fabra Barcelona

MTG Music Technology Group

Evaluation of Set Class Similarity Measures for Tonal Analysis

Nicholas Harley

SMC Masters Thesis

Supervisor: Agustín Martorell

August 30, 2014

Abstract

This work explores the limitations of exiting approaches to computational modelling and description of tonality. Set class theory is presented as alternative or complement to existing approaches. Set class similarity is presented as a tool for the representation of set class information for structural analysis. A survey of set class similarity measures from the literature is conducted as well as a rendering of traditional musicological terminology in the language of set class theory. Six of the most suitable measures are chosen for further evaluation. An analysis methodology is outlined which emphasises systematicity and perceptual relevance. This methodology consists of a number of computational techniques and is used to analyse specific musical examples. The analytical potential of the similarity measures is evidenced through the reconstruction of basic music intuition and analysis using the proposed methodology.

Contents

List of Tables

Set Class Theory Glossary

#X Set cardinality.
#nC Size of nC.
#nCV Vector cardinality.

DV Difference Vector.

I(X) Inversion.
I-Type Inversional SC-Type.
IC Interval Class.
ICV Interval-class Vector.

nC Cardinality Class.
nC%V n-class subset percentage vector.
nCV n-class subset vector.
nSATV n-class subset saturation vector.

PC Pitch Class.
PC-Set Pitch Class Set.
Prime Form PC-set representing all members of an SC.

SC Set Class.

Tn(X) Transposition.
Tn-Type Transpositional SC-type.
TnI-Type Transpositional/Inversional SC-Type.
Trivial Form SCs 1-1, 11-1 and 12-1.

1 Introduction

The objectives of the present work are principally concerned with descriptive modelling of tonality using set class theory. In particular, the analytical potential of set class similarity measures will be assessed and evaluated through practical demonstration within a specific methodological framework. The tentative goal, therefore, is to present this approach in such a way as to make it relevant to the wider community of music researchers, such as those in MIR.

1.1 Problem

1. Existing computational description of tonality is often limited in its scope.
2. The predominant use of template matching for chord and key estimations limits the knowledge of music-capable systems to repertoire of the major-minor paradigm. This narrow view of tonality is insufficient for even some western music.
3. The predominant use of non-musical similarity measures in MIR such as Euclidean distance and correlation seems counter-intuitive in many cases.
4. The above may be contributing to the "semantic gap".

1.2 Objectives

1. To adopt a systematic approach to description of tonality.
2. To justify set class analysis as a systematic descriptive tool, capable of relating useful musical information as well as having a degree of perceptual relevance.
3. To position set class analysis among existing methods of tonal description such as chroma vector time-series.
4. To create a comprehensive and practical survey of set class similarity measures.
5. To examine the measures in terms of systematicity and perceptual relevance.
6. To develop techniques for the representation of set class information so as to expose meaningful musicological information.
7. To utilise set class similarity in these representation.
8. To evaluate the utility of these similarity measures through exploration of SC information to extract a priori knowledge about specific pieces. The analytical potential of the model will be evidenced through specific analysis examples.

1.3 Outline

Part I comprises a concise literature review containing background information and some basic theory. Chapter 2 addresses the challenges and problems involved in descriptive modelling of tonality as a means of justification for the proposed approach. Chapter 3 introduces the basic concepts of set class theory including set class similarity measures. Chapter 4 gives a basic description of multidimensional scaling techniques and how they can play a role in set class analysis. Chapter 5 contains a review of the relevant tonal models that exist so as to give context work that follows.

Part II contains a description of the contribution of this work. Chapter 6 contains the outcomes of a comprehensive survey of set class similarity measures that was carried out. Chapter 7 presents a number of techniques for obtaining and representing a set class description of a musical piece and describes the relationship between parameters involved in the analysis process. Chapter 8 contains a demonstration of each computation technique using real musical examples.

Part III presents a summary of the findings and discussion of the work with Chapters 9 and 10 containing conclusions and future work respectively.

Part IV contains appendices. Appendix A contains a explanation of each similarity measure from the literature. Appendix B contains additional information regarding the chosen measures. Appendix C contains a set class reference guide for pairing common musical objects with their corresponding set classes.

Part I
Background

2 Tonality

2.1 Defining Tonality

2.1.1 General Definition

Tonality is a notoriously complex musical phenomenon and numerous definitions have been proposed from a variety of viewpoints. Perhaps the most general definition is that provided by Hyer (2013): "... refers to the systematic arrangements of pitch phenomena and relations between them." Explanations of tonality have been provided through many different disciplines (acoustics, music theory, linguistics, cognitive psychology) and a detailed discussion of these areas is certainly beyond the scope of this work. However, it is generally agreed that tonality is an abstract cultural and cognitive construct that can have many different physical representations.

2.1.2 Babbitt's Domains

Babbitt (1965) proposed three domains to categorise different types of representation of music: acoustic (physical), auditory (perceived), graphemic (notated). Western music theory provides a lexicon for describing abstract tonal objects with terms such as note, chord and key. These objects have a hierarchical relationship and the meaning of these labels is highly dependent on musical context and the scale of observation. Musicological descriptions, which constitute the majority of reasoning about tonaility, reside mainly in the Babbitt's graphemic domain, although arguably they reflect some aspects of the other two. Each domain, whilst connected to every other, provides only a projection of the musical whole and examination of tonality from just one will most likely result in an incomplete picture. However, these three domains provide a convenient framework for the discussion that follows.

2.2 Modelling Tonality

The challenge of mathematically modelling aspects of tonality has been approached in numerous ways and from different domains. In the graphemic domain, musicologists and composers have proposed theoretical models, attempting to rethink tonal theory from a mathematical perspective. These models employ different branches of mathematics such as geometry (Tymoczko, 2012) or group theory (Ring, 2011) to describe harmonic structure. From the auditory domain, cognitive psychologists have built models of tonal induction based on perceptual ratings of tonal stimuli (Krumhansl, 1990).

2.2.1 Tonality as Context

Many models approach the concept of tonality as a context, within which the relations and hierarchies of tonal phenomena can be understood. A sense of tonality can be induced when musical stimuli resemble some a priori contextual category. For western music of the major-minor period, key signatures comprise a collection of categories that give context to the tonal components of music. Martorell (2013) identifies three important aspects of tonality as context: dimensionality (the relatedness or "closeness" of categories), ambiguity (reference to two or more categories simultaneously) and timing (the dynamics of tonal context). He highlights the importance of a models capability to describe these aspects.

2.2.2 Tonality in MIR

The MIR community is primarily concerned with the extraction of tonal descriptors from audio signals such as chord and key estimates. Most systems use chroma features as a preliminary step, obtained by mapping STFT or CQ transform energies to chroma bins. Template matching is used to compare the chroma vectors to a tonal model (contextual category) using some distance measure. A commonly used tonal model for key estimation are the KK-profiles (Krumhansl, 1990) (5.3) (e.g. in Gómez 2006). Distance measures such as inner product (e.g. in Gómez 2006) and fuzzy distance (e.g. in Purwins et al. 2000) are used to compare vectors. Statistical methods, such as HMMs, have been used for chord and key tracking (Chai, 2005). Of addition interest in the field is the concept of musical similarity (for music recommendation, structure analysis, cover detection etc.). Foote (2000) computed self-similarity matrices for visualisation of structure by correlating the MFCC feature vector time-series. Gómez (2006) proposed the application of this method to tonal feature vectors.

2.2.3 Similarity

The importance of defining the similarity or closeness between musical phenomena, be it theoretical, physical or perceptual, is central to almost every model of tonality and often leads to a geometric configuration of tonal objects. The concepts of similarity and distance is discussed further in Chapter 5 where a review of spatial models of tonality is given.

2.3 The Semantic Gap

2.3.1 Acoustic Domain

Wiggins (2009) discusses, what is referred to in MIR as, the "Semantic Gap": the inability of systems to achieve success rates beyond a conspicuous boundary. He examines the fundamental methodological groundings of MIR in terms of Babbitts three domains, discussing the limits of each representation and regarding the discarnate nature of music. He concludes that the audio signal (acoustic domain) simply cannot contain all of the information that systems seek to retrieve. He points towards the the auditory domain as the chief residence of music information and urges for in not to be overlooked in MIR and wider music research.

2.3.2 Graphemic Domain

Furthermore, Wiggins criticises the purely graphemic approach and the tendency of music research to presuppose musicological axioms. Wiggins (2012) argues that music (tonal) theory is, rather than a theory in the scientific sense, a highly developed folk psychology (internal human theory for explaining common behaviour). Thus, the rules of music theory are not like scientific laws but rather abstract descriptions of a specific musical behaviour. This idea challenges the validity of formalising such rules in mathematics and prompts the question, "What is actually being modelled?" He concludes that to apply mathematical models to musical output alone (scales or chords) without consideration of the musical mind is a scientific failure.

2.3.3 Problems

The two assertions of Wiggins sit contrary to a number of the aspects of the tonal models discussed in 2.2. Firstly, the major-minor paradigm, upon which many systems are based, whilst certainly possessing perceptual significance, is still a musicological concept and therefore a misleading basis for both mathematical and cognitive approaches. A second problem is that of the numerical methods used by some MIR systems, in particular, distance measures. As will be discussed in Chapter 5, similarity (and by extension distance) is a central part of the auditory domain. MIR systems often uses distance measures from mathematics such as Mahalanobis (Tzanetakis and Cook, 1999) or Cosine (Foote, 2000) with little consideration of their perceptual or musical significance.

2.4 Systematicity

2.4.1 The Musical Surface

Having cautioned against a purely musicological approach, Wiggins (2009, pp. 481) proposes a compromise: to adopt a bottom-up approach to music theory, exploring the concepts through systematic mid-level representations. He states that "methods starting at, for example, the musical surface of notes is a useful way of proceeding" The concept of musical surface is illustrated by Huovinen and Tenkanen (2007, pp. 159) with a metaphor: "...to approach a musical landscape not by drawing a map, which necessarily confines itself to a limited set of structurally important features, but by presenting a bird's-eye view of the musical surface – an aerial photograph, as it were, which details the position of every pitched component."

2.4.2 Systematic Description

Martorell (2013) also advocates this mid-level approach, observing that surface description influences analytical observation and that, for an unbiased view, the researcher must be provided with the adequate raw materials with which to make more in-depth observation. Such a systematic, descriptive model would be fundamentally independent of high level concepts such as chords and key but, at the same time, capable of capturing them. Martorell (2013) also discusses the importance of systematicity in terms of dimensionality, ambiguity and timing. He finds that models based on the major-minor paradigm are incapable of adequately describing tonal ambiguity even in some Western music (Martorell, 2013, chap. 3).

With a systematic description of the musical surface, theories and models from different domains can be gathered and evaluated together in the same analytical arena, thus helping to bridge the gap between traditional musicology, cognitive psychology and MIR.

3 Set Class Theory

One such method available for systematic description of the musical surface is set class theory. Set class theory is a system for analysing the pitch content of music. It uses class equivalence relations to reduce the amount of data required to describe any collection of pitches. This chapter will outline the basic principles.

3.1 Pitch Class Set

Set class theory uses octave equivalence. In Western equal temperament (TET), a pitch class (PC) is an integer representing the residue class modulo 12 of a pitch (**?**)Babbitt1965) and indicates the position of a note within the octave. A PC-set is a collection of PCs ignoring any repetitions and the order in which they occur. PC-sets are notated as follows {0,1,2,3,4} with PCs ordered from lowest to highest as a convention (Example 1). The cardinality of a set, denoted #S, is the number of PCs it contains (Example 2). There are 4096 (2^{12}) unique PC-sets with which any segment of music can be represented.

Table 1: Notes and corresponding pitch-classes

Note	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
PC	0	1	2	3	4	5	6	7	8	9	10	11

Example 1:	PC-set	Pitch-set	S = {A4,C5,E5,A5} (A minor)
		PC-set	S = {9,0,4,9} = {0,4,9}
Example 2:	Cardinality		#S = 3

3.2 Set Classification

Defining equivalence classes of PC-sets further reduces the total number of tonal objects. A set-class (SC) is a group of PC-sets related by a transformation or group of transformations. The two types of transformation commonly used are transposition and inversion. A transposition, Tn(S), transposes the set, S, by the interval, n, (by adding n to all PCs, Example 3). An inversion, I(S), inverts the set S, replacing all PCs with their inverse (12-PC, Example 4). From these two transformations it is possible to define three types of SC: Tn, TnI and I.

Example 3:	Transposition	S = {0,4,9}, T3(S) = {3,7,0} = {0,3,7}
Example 4:	Inversion	S = {0,4,9}, I(S) = {11,7,2} = {2,7,11}

Transpositional (Tn):	All PC-sets that can be transformed to each by transposition belong to the same class. There are 351 distinct Tn types.
Inversional (I):	All PC-sets that can be transformed to each other by inversion belong to the same SC. There are 200 distinct I types.
Transpositional/ Inversional (TnI):	All PC-sets that can be transformed to each other by transposition, inversion or both belong to the same SC. There are 223 distinct TnI types.

The Prime Form of a PC-set is a convention for denoting the SC it belongs to. The convention was introduced by Allan Forte (Forte, 1973) for TnI types and has since been adopted by the majority of theorists. In addition, he devised a system for ordering TnI-type SCs and assigning to each one a cardinality-ordinal number. For example, the Forte number 3-11 refers to the 11th SC of cardinality 3. This convention has been modified for use with Tn types by adding A and B to the names of inversionally related SCs.

One additional concept is that of cardinality-class (nC), which refers to all the SCs of cardinality n. Cardinality-class 2 is commonly referred to as interval-class (IC) and there are 6 distinct interval-classes.

Table 2: Forte's Prime form and numbering convention

PC-set	{0,4,9}
Prime Form (TnI)	{0,3,7}
Prime Form (Tn)	{0,4,7}
Forte Name (TnI)	3-11
Forte Name (Tn)	3-11A

Table 3: Numbers of objects

Object type	No. Objects
Pitch	88
Pitch set	3e26
PC	12
PC-set	4096
Tn-Type SC	348
I-Type SC	197
TnI-Type SC	220

3.3 Vector Analysis

3.3.1 Membership and Inclusion

Two concepts that are crucial in set class theory are membership and inclusion. Membership of a set is denoted p $\in$ S and means that PC p is a member of set S (Example 5). Inclusion in a set is

Table 4: Cardinality Class

		$\#nC$	
n	Tn	I	TnI
1C	1	1	1
2C	6	6	6
3C	19	12	12
4C	43	28	29
5C	66	35	38
6C	80	35	50
7C	66	35	38
8C	43	28	29
9C	19	12	12
10C	6	6	6
11C	1	1	1
12C	1	1	1

denoted Q ⊂ S and means that all members of set Q are also members of set S (Example 6). Q is said to be a subset of S.

Example 5:	Membership	$4 \in \{0,4,9\}$
Example 6:	Inclusion	$\{0,4,9\} \subset \{0,1,4,5,9\}$

3.3.2 Embedding Number

Lewin (1979) applied these concepts to SCs to develop his Embedding Number, EMB(X,Y). Given two SCs, X and Y, EMB(X,Y) is the number of instances of SC, X, which are included in (are subsets of) SC, Y (Example 7). X is ring-shifted 11 times and each unique resulting set which is included in Y adds one to the embedding number.

Example 7:	Embedding Number	X = {0,4} and Y = {0,4,8}
		so… EMB(X,Y) = 3

3.3.3 Subset Vectors

An n-class subset vector of X, nCV(X), is an array of values of EMB(A,X) where A is each of the SCs in the cardinality-class, nC (Example 8). The Interval-Class Vector (ICV) is a special instance of the nCV with n equal to 2. Vector cardinality, denoted #nCV(X), is the sum of all the terms in the vector (Example 9). The length of a subset vector is given by the number of SCs in the cardinality class, #nC.

Subset vectors form the basis of the majority of analysis performed by set class theorists. In addition, many theorists have proposed modifications to the basic nCV to suit their specific purposes and some of these modifications will be discussed in context where necessary.

Example 8:	Subset Vector	S = {0,4,9}
		2CV(S) = ICV(S) = [0 0 1 1 1 0]
Example 9:	Vector Cardinality	#ICV(S) = 0+0+1+1+1+0 = 3

3.4 Set Class Similarity

3.4.1 Similarity Relations

The assessment of similarity between two SCs has been discussed in the literature for decades and a large number theoretical models have been proposed. Different models approach the problem from different conceptual standpoints and theorists have different opinions about the contributing factors. All these models are described under the blanket term “similarity relations”. Despite the perennial fascination with the concept, little or no consensus exits as to what constitutes a good similarity relation.

Castrén (1994) provides a comprehensive and in-depth review of a large number of similarity relations and categorises them according to some fundamental principles. Firstly, he distinguishes between methods that produce binary outcomes and those that produce a range of values. The former category, termed “plain relations”, include Forte’s R-relations (Forte, 1973) and indicate whether the two SCs are related in a specific way, which in turn may give some indication of whether they are similar. The latter category, termed “similarity measures”, indicate a degree of similarity, returning a value from a known range. This property appears to be more inline with the perceptual notion of similarity and therefore the focus of this work shall be exclusively on similarity measures.

3.4.2 Similarity Measures

The vast number and diversity of approaches to similarity measures renders concise summation challenging if not impossible. The problem can only be approached by narrowing the focus to a specific type. This work focuses on measures that use the Tn and TnI-type SCs (3.2), and furthermore we will only consider those methods based on vector analysis (3.3). These measures usually involve the comparison of the SCs’ nCVs. Of this (still sizeable) subset, Castrén (1994) identifies two main categories.

Single nC:	Single nC measures compare the nCVs of the two SCs for one particular value of n. Many of the relations in this category compare ICVs (2CVs).
Total Measures:	Total Measures consider the subsets of all cardinalities contained within in two SCs. All the relevant nCVs are compared to produce a final value.

Table 4 shows the majority of the Tn and TnI-Type, vector based similarity measures from the literature organised by theorist. Vector Type indicates whether the measure compares ICVs or nCVs (nC%V, nSATV and CSATV are all variations on the basic nCV). Card (Cardinality) indicates whether the measure is capable of comparing SCs of different cardinalities while the Measure Type indicates which of Castren’s categories it belongs to. nC indicates it is a Single nC measure and TOTAL indicates it is a Total Measure. All these measure are described more thoroughly in A.

3.4.3 Castren’s Criteria

In addition to his categorisation, Castren specifies a number of criteria which a good similarity relation should meet. Later, these criteria will be used in assessing the suitability of the various similarity measures.

Castren says that a similarity measure should:

C1: Allow comparisons between SCs of different cardinalities
C2: Provide a distinct value for every pair of SCs
C3: Provide a comprehensible scale of values such that...
- **C3.1:** All values are commensurable
- **C3.2:** The end points are not just some extreme values but can be meaningfully associated with maximal and minimal similarity.
- **C3.3:** The values are integers or other easily manageable numbers
- **C3.4:** The degree of discrimination is not too coarse and not unrealistically fine

C4: Produce a uniform value for all comparable cases
C5: Observe mutually embeddable subset-classes of all meaningful cardinalities
C6: Observe also the mutual embeddable subset-classes not in common between the SCs being compared.

Table 5: Comparison table of similarity measures

THEORIST	SIMILARITY MEASURE	VECTOR TYPE	CARD	MEASURE TYPE
	K	ICV	SAME	nC
	SIM	ICV	SAME	nC
MORRIS	ASIM	ICV	ANY	nC
LORD	sf	ICV	SAME	nC
TEITELBAUM	s.i.	ICV	SAME	nC
	IcVD1	ICV	ANY	nC
	IcVD2	ICV	ANY	nC
ROGERS	COS	ICV	ANY	nC
	AMEMB2	ICV	ANY	nC
	IcVSIM	ICV	ANY	nC
	ISIM2	ICV	ANY	nC
ISAACSON	ANGLE	ICV	ANY	nC
	AK	ICV	ANY	nC
	MEMBn	nCV	ANY	nC
	TMEMB	nCV	ANY	TOTAL
RAHN	ATMEMB	nCV	ANY	TOTAL
	REL2	ICV	ANY	nC
LEWIN	REL	nCV	ANY	TOTAL
	%RELn	nC%V	ANY	nC
	T%REL	nC%V	ANY	TOTAL
CASTREN	RECREL	nC%V	ANY	TOTAL
	SATSIM	nSATV	ANY	nC
	TSATSIM	nSATV	ANY	TOTAL
BUCHLER	AvgSATSIM	nSATV	ANY	TOTAL

3.5 Perceptual Relevance

The many equivalence relations used in PC-set theory give rise to a highly abstract description of musical objects. Thus, an important question to be asked is whether these theoretical assumptions and models of similarity reflect perceptual equivalence. This chapter contains a summary and discussion of some relevant studies.

3.5.1 Octave Equivalence

Pitch is a percept that derives from a particular harmonic structure and is roughly proportional to the logarithm of the fundamental frequency. This allows pitch to be perceptually modelled as a straight line. Music psychologists have observed a strong perceptual similarity between pitches with fundamental frequencies in the ratio of 2:1. This property of octave similarity leads the straight line model of pitch to be bent into a helix. Division of the octave into a number of categories is thought to offer a more efficient cognitive representation in memory and thus confers evolutionary advantage. The resulting pitch equivalence classes are implicitly learned through repeated exposure. TET has 12 pitch equivalence classes which, in PC-set theory, are modelled as a circular projection of the pitch helix. Thus the two most fundamental components of PC-set theory, i.e. octave equivalence and pitch-class labelling, would appear to have a solid basis in perception.

Gibson (1988) investigated the perceived similarity of pairs of chords with varying numbers of octave related pitches. He found that in general chords with identical PC contents were perceived as more similar than chords with near identical PC contents, regardless of the octave of the pitch components. However, in further studies he his findings suggest that there are other factors that play a significant role (Gibson, 1993).

3.5.2 Set Class Equivalence

Some researchers have attempted to examine whether there is perceived equivalence between different manifestations of a PC-set. Krumhansl et al. (1987) presented subjects with sequences of tones derived by transforming two different PC-sets. They noted that subjects were able to distinguish between the different sets both in neutral and musical contexts.

Millar (1984) investigated the perceptual similarity of different PC-sets derived from the same set class under TnI classification. Subjects were presented with three-note melodies and asked to judge which was equivalent to a reference melody. Some melodies preserved the SC identity whilst others did not. She found transpositions to be perceived more similar than inversions and in addition she discovered that the order of the notes and melodic contour was a strong factor in perceived similarity.

Some authors have questioned the perceptual relevance of using TnI and I equivalence as a basis for set classification. Deutsch (1982) seems unconvinced by evidence for the perceptual similarity of inverted intervals. This can be illustrated by the example of major and minor triads which, while perceptually distinct, are equivalent under TnI and I equivalence.

3.5.3 Perceived vs Theoretical Similarity

A number of studies have been done to ascertain the connection between perceptual similarity ratings and the theoretical values obtained from some set class similarity measures. A large number of relevant studies are summarised by Kuusi (2001) and the most significant ones are mentioned here.

Bruner (1984) used multidimensional scaling on subjects' similarity ratings between trichords and tetrachords and on the similarity values obtained from SIM. She compared the 2-dimensional solutions and found there to be little correlation.

Gibson (1986) investigated non-traditional chords. He compared subjects' ratings with similarity assessments calculated from Forte's R-relations and Lord's similarity function. He also concluded there was little correspondence between the two.

Stammers (1994) compared subjects' ratings of 4 note melodies with the theoretical values obtained from SIM. She found the ratings of subjects with more musical training to be more correlated with the SIM values.

Lane (1997) compared subjects' ratings of pitch sequences with corresponding values of seven ICV-based similarity measures: ASIM, MEMB2, REL2, s.i., IcVSIM and AMEMB2 and concluded there to be a strong relation.

Kuusi (2001) compared subjects' ratings of pentachords with the values obtained from 9 similarity measures. He found there to be a connection between aurally estimated ratings and the theoretical values and concluded that the abstract properties of set-classes do have some perceptual relevance. He also comments on the way in which this kind of study is conducted, suggesting that the way in which subjects are presented with the stimuli has a significant effect on the outcome.

3.6 Set Class Analysis

PC-set theory as means for descriptive modelling of tonality is not widely known outside of highly theoretical circles and the use of set-class similarity measures seems mainly restricted to the theorists who proposed them (for example, Isaacson 1996). The basic premise is simple: a musical piece is segmented and each segment described by its SC. Similarity measures can be used to assess the similarity between segments or between a segment and some reference SC.

Huovinen and Tenkanen (2007) used a pentachordal tail segmentation policy (each successive note defines a segment that includes the preceding four notes) and compared these segments to sets 7-1 (chromaticism) and 7-35 (diatonicism) using the REL distance (A.8.1). They claim that the visual results of their analysis "reflect pertinent aspects of our listening experience" (Huovinen and Tenkanen, 2007, pp. 204).

Martorell (2013, chap. 5.3) uses a more systematic approach to segmentation using multiple time scales. He proposes the class-scape, a two-dimensional visualisation of a piece of music with time on the x-axis and segmentation time-scale on the y-axis. The presence of a single SC can be indicated by highlighting the segment or alternatively each segment can be shaded according to its REL distance from a comparison SC. He emphasises that the class-scape is an exploratory tool rather than an automated analysis system.

Perhaps the most crucial aspect of using SC descriptions for tonal analysis is the way in which a piece of music is segmented. The issue of segmentation will be discussed further in Chapter 7.3.

4 Multidimensional Scaling

Multidimensional scaling (MDS) is a numerical visualisation technique that, given a matrix of pair-wise distances between objects, provides a geometric configuration of the objects in some abstract space. It provides an efficient means of observing relationships in large, complex data sets and the resulting dimensions often give valuable insight into the data as a whole.

4.1 Non-Metric MDS

Non-Metric MDS was first described by Shepard (1962) and it assumes that the distance matrix values are related to points in an abstract N-dimensional Euclidean space. An important consideration is that of the dimensionality of the solution. For comprehension and visualisation it is important to minimise the number of dimensions however, there is a trade-off between the number of dimensions and the accuracy of the model. For a given dimensionality, we obtain a value of stress. Stress is a "goodness of fit" measure which characterises the distortion that occurs in a given number of dimensions. As the number of dimensions increases the stress decreases and the choice of dimensions should be based on interpretation.

4.2 Cluster Analysis

Cluster analysis (CA) is method for dealing with dimensions that are highly separable. First, the most similar pair of objects are selected and grouped together in a cluster. The process is repeated, creating a binary tree structure. The distance between objects is then related to their separation along the branches of the tree.

4.3 MDS with Similarity Measures

Using MDS on the values produced by similarity measures is one way to approach an understanding of the constructs they are measuring. There are two potentially interesting issues to consider. Firstly, a measure may be inconsistent with itself, meaning that the geometries it produces are not "robust"; changing the set of objects changes the distances between the original set. This kind of problem cannot be observed through inspection of the values alone. The second issue is that two different measures that are both self-consistent may produce very different geometries from the same group of SCs. The question then is, what exactly do the measures measure?

5 Spatial Models of Tonality

5.1 Similarity and Distance

Judgements of similarity form the basis of many cognitive processes including the perception of tonality. Similarity between two objects is often conceived as being inversely related to distance between them in geometric space. For example, some tonal objects (chords, for example) are perceived as close to one another whereas others are further apart. In addition, the number of dimensions of the geometric space is in connection with the number of independent properties that are relevant for similarity judgments. Gärdenfors (2000) suggests that humans are naturally predisposed to create spatial cognitive representations of perceptual stimuli due to the geometric nature of the world we have evolved to inhabit. Therefore spatial modelling of tonality, as well as helping to visualise the complex multidimensional relationships between tonal phenomena, has the potential to reflect cognitive aspects of the way they are perceived.

5.2 Spatial Representations

Throughout history theorists have proposed many spatial representations of tonality from different domains. From the graphemic domain, Weber (1851) and Schoenberg (1954) both proposed simple 2-dimensional charts to display the proximity between keys. For representation of chords, Riemann (1877) models major and minor triads as regions in a 2-dimensional space whilst Tymoczko (2011) proposes a variety high dimensional, non-euclidean chord spaces that reflect the theoretical principles of voice leading. From the acoustic domain, Shepard (1982) proposes a five-dimensional model to represent interval relations between pitches. Some theorists have attempted to incorporate relations between several levels of tonal hierarchy into one configuration. The "spiral array" of Chew (2000) is a three-dimensional mathematical model which simultaneously captures the relations between pitches, chords and keys. The "chordal-regional space" of Lerdahl (2001) models the relations between chords within a certain key.

5.3 Cognitive Psychology

The auditory domain has been addressed through cognitive psychology by Krumhansl (1990) who used the probe-tone methodology (Krumhansl and Shepard, 1979) to establish major and minor key profiles (12-dimensional vectors containing the perceptual stability ratings of each of the 12 pitch classes within a major or minor context). These profiles, know as Krumhansl-Kessler profiles (KK-profiles), show the hierarchy of pitches in major and minor keys. Correlating each of the 24 major and minor profiles produced a matrix of pairwise distances which was fed to a dimensional scaling algorithm. The resulting geometrical solution was found to have a double circular property (circle of fifths and relative-parallel relations) which can be modelled as the surface a 3D torus. Many spatial models of tonality have this double circular property whether it is implicit (Weber, 1851; Schoenberg, 1954) or stated explicitly (Lerdahl, 2001).

5.4 Set Class Spaces

Most of these models are limited to description of music in the major-minor paradigm and are not capable of generalising beyond the "western common practice". PC-set theory, once again, provides a possible means to generalise to any kind of pitch-based music. By considering a collection of tonal objects described by SCs, a geometric space can be constructed to model their relations based on some theoretical principle. Some PC-set theorists have proposed explicit geometric spaces to model relations between SCs. The distances in these spaces are expressed by models of similarity based on voice leading (Cohn, 2003; Tymoczko, 2012) or ICVs and the Fourier transform (Quinn, 2006, 2007). However, these models are only designed to represent SCs of one cardinality-class at a time and cannot model the relations between arbitrary collections of pitches.

Alternative spatial models are provided by the implicit geometries of the values produced by the SC similarity measures discussed in 3.4. As mentioned in 4.3, MDS can be used on values produced by similarity measure to create a geometric space. Kuusi (2001) and Samplaski (2005) both applied MDS to the values produced from a variety of similarity measures. Samplaski used TnI-type SCs while Kuusi used Tn-type. They both found reasonably low-dimensional solutions and attempted to interpret each of the dimensions. Kuusi interpreted three dimensions as corresponding to chromaticism, wholetoneness and pentatonicism. Samplaski made similar observations but found some dimensions in the higher-dimensional spaces difficult to interpret. Nevertheless, he concluded that values from similarity measure tend to agree (with some exceptions) and that they measure constructs relating to familiar scales (diatonic, hexatonic, octatonic, etc.).

Part II
Contribution

6 Similarity Measure Survey

So far, brief reference has been made to the extensive existing literature on set class similarity measures (3.4). This chapter summarises the outcomes of an extensive survey of the different models. The large number of measures are discussed in relation to Castren's criteria (3.4.3) in order to gauge their suitability for use in systematic surface description models. The most suitable models will be examined further.

6.1 Criteria

Castren's criteria (see 3.4.3) for similarity measures provide a basis for assessment of similarity measures for our purposes. A detailed descriptions and justification for the criteria can be found in Castrén (1994, chap. 2), however here we will focus on one or two specific aspects. Table 6 shows the list of similarity measures with marks indicating whether each of the criteria is met. In sections 6.2 to 6.4 specific criteria are used to exclude measures from further consideration with justification in terms of systematicity and perceptual relevance.

Table 6: Castren's Criteria

SIMILARITY MEASURE	C1	C2	C3.1	C3.2	C3.3	C3.4	C4	C5	C6
K	X	X			X	X	X		
SIM	X	X			X	X	X		
ASIM	X	X	X	X		X	X		
sf					X	X	X		
s.i.					X	X			
IcVD1	X	X	X	X		X	X		
IcVD2	X	X	X	X		X			
COS	X	X	X	X		X			
AK	X	X	X	X		X	X		
MEMBn	X	X			X	X	X		
TMEMB	X	X			X		X	X	
ATMEMB	X	X	X	X		X	X	X	
AMEMB2	X	X	X						
IcVSIM	X	X				X			
ISIM2	X	X				X			
ANGLE	X	X	X	X		X			
REL	X	X	X	X		X	X	X	
REL2	X	X				X			
%RELn	X	X	X	X	X	X	X		
T%REL	X	X	X	X	X	X	X	X	
RECREL	X	X	X	X	X	X	X	X	X
SATSIM	X	X	X						
TSATSIM	X	X	X	X		X		X	
AvgSATSIM	X	X	X	X		X		X	

6.2 Cardinality

Measures which fail to meet criteria C1, i.e. that cannot compare SCs of different cardinalities, are clearly inadequate for systematic analysis of music, which might require the comparison of any two arbitrary segments regardless of how many PCs they contain. Both s.i. (A.4.1) and sf (A.3.1) were proposed specifically for SCs of the same cardinality and so will be excluded from further discussion. Some other measures which were intended to compare SCs of different cardinalities nonetheless have problems. Measures such as SIM (A.2.2) and K (A.2.1) give unintuitive values when the cardinalities of the SCs being compared differ greatly and, in addition, the range of values produced depends on the cardinality of the sets (failure to meet criteria C3.1). Measures of this type will also be excluded.

6.3 Set Class Type

An important consideration when using similarity measures is the type of SC being compared. Many of the measures are designed for comparison of TnI-type SCs, however, owing to issues riased in 3.5 regarding the perceptual relevance of invertionally related sets, here, measures will be selected for use with Tn-type SCs. This means that the measure should be able to discriminate between inversionally related sets. All the single-nC measures which exclusively consider interval content (ICVs) in the comparison procedure can therefore be discounted, as inversionally related sets have identical ICVs.

6.4 Measure Type

Although many theorists have supposed that interval-class subsets are of paramount importance in similarity judgments, no thorough investigation has been carried out as to the exact perceptual significance of subset cardinality. Single-nC measures presuppose that subsets of one particular cardinality contribute to similarity above all others. In the interest of systematicity, we will not make this assumption, instead assuming that subsets of all cardinalities are equally relevant and should be considered. Similarity measures that exhaustively consider all subset cardinalities meet criteria C5 and are "total" measures (see 3.4.2). The six total measures from 3.4.2 shall therefore become the focus of this work. Details on the specific formulations (including three versions of REL) are given in Appendix B.

6.5 Total Measure Comparison

For a preliminary idea of the utility of the total measures it is useful to visualise the values produced for comparisons involving some of the common tonal objects described in Appendix C. This information can be visualised as 2D grids with each square corresponding to the comparison between two tonal objects and shaded according the distance between them i.e. the value of MEASURE-prime (see B.2). Figures 1 and 2 show two such grids for ATMEMB-prime and AvgSATSIM-prime respectively. As can be seen, the values produced by the two measures are quite different. Thus, measure selection will be an important part of the analysis and these grids will form a useful reference guide when selecting parameters.

Plotting the absolute difference between the values in these grids gives a local indication of comparisons for which the measures most disagree. Figure 3 shows such a plot, in which the lighter areas indicate a higher degree of discrepancy between the measures' values. A more quantitative comparison of the measures can be obtained by correlation of the vectors containing all 61425 values. Figure 4 shows a grid with each square corresponding to a comparison between measures and coloured according to the correlation value.

The shear quantity of values for all the measures means that a thorough and meaningful comparison of the would be difficult and time consuming. Still, from a superficial inspection of these grids it is possible to draw some basic conclusions:

1. The values produced by the measures are sufficiently different as to produce different outcomes in the proposed analysis.

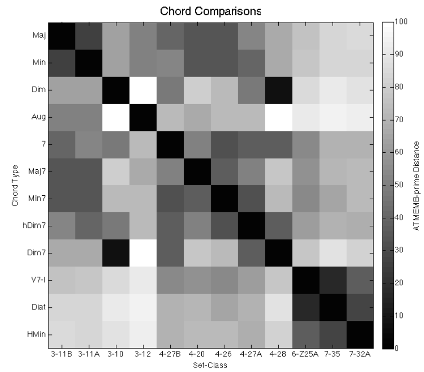

Figure 1: Chord Comparisons: ATMEMB-prime distances between common tonal objects

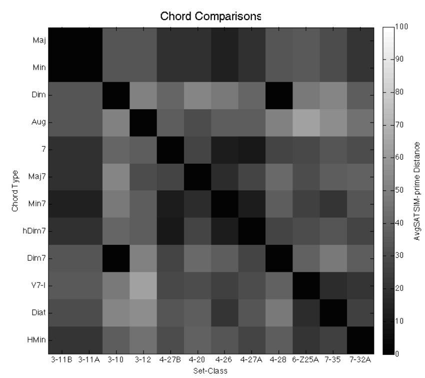

Figure 2: Chord Comparisons: AvgSATSIM-prime distance between common tonal objects

2. Different measure possess different discriminatory power and give different contrast for different collections of set classes.
3. Whilst all the measures can in principle discriminate between inversionally related sets only ATMEMB and REL discriminate between major and minor triads.
4. AvgSATSIM and TSATSIM are very similar and the values they produce are overall lower indicating greater proximity between common chord set classes.
5. AvgSATSIM and TSATSIM discriminate poorly between smaller set (cardinality 3/4) but between between larger sets (cardinality 7+).
6. The values produced by TPREL are overall higher indicating greater separation between common chord set classes.
7. ATMEMB gives a high degree of discrimination between set classes of very different cardinalities.
8. RECREL and REL discriminate the best between smaller set classes (cardinality 3/4).

7 Analysis Methodology

This chapter describes a set of computational techniques that can be used in conjunction for analysis of a musical piece. 7.1 gives a concise overview of the analysis process and introduces the key variables whilst 7.2 to 7.4 give a more detailed explanation of the factors involved in the selection of the parameters. In the interests of clarity, demonstration of the techniques with examples will be postponed until chapter 8.

In this work, analysis is done from digital scores in MIDI format. The advantage of this is that it avoids the potential inaccuracies involved in extracting chroma from an audio signal. Symbolic data such as MIDI allows direct access to the pitch material upon which the analysis techniques are applied.

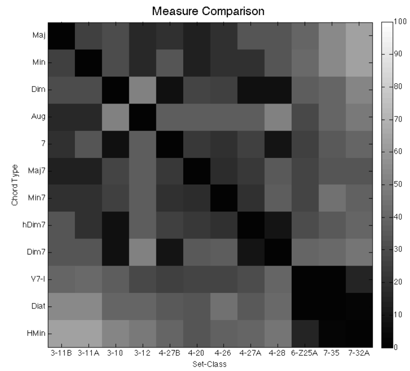

Figure 3: Measure Comparison: Absolute difference between ATMEMB-prime and AvgSATSIM-prime values

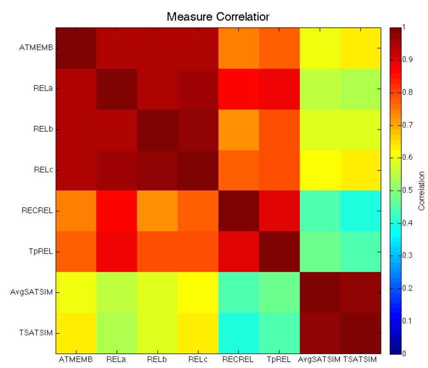

Figure 4: Measure Correlation: Correlation between the all the values produced by all measures.

7.1 Overview

The central component in analysis using similarity measures is the distance time series. This provides a simple means of capturing how the pitch content of a piece evolves in time with respect to a specific set class or sonority. It involves segmenting the piece using a fixed sliding window and calculating the distance between each segment and a reference set class. Such representations of tonal progression in time lend themselves very well to analysis of harmonic and musical structure. Specific features of the curve can indicate structurally important events while repetitions in the time series can indicate passages with similar tonal progressions. The first examples of distance time series are in Figure 10 (8.2).

There are three interdependent parameters which must be selected according to the specific intentions of the analyst: Segmentation (window and hop size), the reference set class and the similarity/distance measure. The segmentation determines the captured set class content, which should be targeted according to its relationship to the reference set class. This relationship is determined by the measure used, which must possess an adequate degree of discrimination so as to produce characteristic changes in the time series.

7.2 Reference Set Class Selection

Selection of the reference SC will vary depending on the intentions of the analyst; different selections will reveal different musical features and types of harmonic structure. In traditional musicology, the components of harmonic structure are described by scales, chords and chord progressions. As a preliminary step towards the reconstruction of this conventional analysis, it is necessary to make some connections between common musical objects and set class theory. Appendix C contains tables that list common chords, cadential progressions and scales with their corresponding set classes. Before selecting a reference set it is necessary to identify which of these basic groups are most relevant to the analysis aims i.e. to establish the `sets of interest`. Below, three potential reference sets are proposed with and given musicological motivation.

1. Major/Minor Triad (3-11A/3-11B): The triad is widely considered to be the basic building block of western harmony. A distance time series with reference to the basic major or minor triad will give an indication as to the complexity of the chords and harmonic progressions.
2. Perfect Authentic Cadence (6-Z25A): In much of western music cadences punctuate harmonic progressions by suggesting varying degrees of resolution. A distance time series with reference to the perfect authentic cadence might contain characteristic features at the boundaries between distinct passages.
3. Diatonic Scale (7-35): The pitch content of much western music is confined to the diatonic scale. A distance time series with reference to the diatonic set would indicate the degree to which the music is diatonic or signal the use of other scales and modulations.

7.3 Segmentation

Using a fixed sliding window to segment the music requires considered selection of the window length and hop size so as to best target the `sets of interest`. The selection of these parameters is a crucial stage in the analysis process, which is highly sensitive to the scale of observation. The window size determines the cardinality of the sets which are captured, with larger windows typically containing larger cardinality sets. The relationship between hop size and captured set class contents is rather subtle: smaller hop sizes are required for observing the note-wise change in set class that occurs in passages of sequential (horizontal) notes, whereas larger hop sizes can be used when the harmony is built from concurrent (vertical) notes.

When working in the MIDI domain, an alternative segmentation policy can be adopted to supplement the analysis process. This method, from Martorell (2013, chap. 5.3), is a fully systematic segmentation policy which exhaustively windows every possible combination of adjacent notes. Martorell (2013, chap. 5.3.5) also specifies two compact representations of this data as a means of observing the the global set class content of a piece:

Class Matrix is a 2d plot with time on the x axis and set class on the y axis. The set class of each window is calculated and plotted as a horizontal line.

Class Vector shows the relative active duration of each class in the class-matrix expressed as a percentage of the total duration of the piece.

The first examples of class matrix and class vector are shown in Figure 6 (8.1.1). This information gives a global indications as to the types of sonorities contained within a piece and can aid reference set selection. From this complete information it is also possible to view statistical information about the time scale in which specific set classes or cardinality classes occur. This information can inform the selection of window and hop size so as to best target the sets of interest. The first example of this is in Figure 7 (8.1.1). Once the sliding window segmentation has been performed, the captured set class contents can be viewed by superimposing them on top of the class matrix and class vector. This representation gives an indication of the proportion of the overall class contents that have been retrieved and thus the efficacy of the sliding window parameters. The first example of this is in Figure 11 (8.3).

7.4 Measure Selection

As mentioned previously, different measures may be appropriate for different analysis contexts. Grids such as those shown in 6.5 form a simple way to visualise the values produced by a measure and can give an indication of the relationship between the set classes in a in a specific time scale. Comparison of these grids can reveal the strengths and weaknesses of the different measures. Often it is useful to directly compare the time series produced by two measures by plotting both. This technique is used throughout 8. In many cases this is the simplest way to select the best measure.

A further method available for understanding the relationship between set classes is through multidimensional scaling. Spaces formed from the set classes contained in a specific segmentation time scale can be used for comparing different measures and can aid the selection of comparison set. Examples of this are in 8.6.

8 Analysis Examples

In this chapter the analytical potential of the similarity measures is evaluated through specific analysis examples. The subjects of the analysis are described in 8.1 while, 8.2 to 8.6 contain examples of the computation techniques described in 7.

In examples where distance time series are displayed, different measures are plotted in different colours. Figure 5 shows a colour key for these plots. In each example a selection of measures are presented together for comparison. Each selection is intended to demonstrate the variation in measure selection.

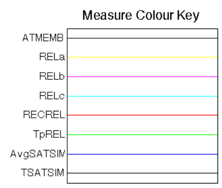

Figure 5: Measure colour key

8.1 Musical Examples

Two pieces were chosen as subjects for analysis: *a*) BWV-846: C Major Prelude from Book I of The Well-Tempered Clavier by JS Bach *b*) Dvorak-Op101-1: Humoresque No. 1, Vivace by Antonín Dvořák. Each piece is short and for solo piano/keyboard, which limits the number of voices complicit in the harmony. In addition, each piece exemplifies some example of common musical practise.

8.1.1 BWV-846

This piece was chosen for the relative simplicity of its tonal contents. The harmonic progression is expressed through a series of arpeggiated chords which are mainly confined to familiar triads and seventh chords and arranged in common cadential progression. The structure of the piece is as follows:
Bars 1-4: A full cadence in C major
Bars 5-11 Modulation to G major
Bars 12-19 Modulation back to C major
Bars 20-35 Complex extended cadence in C major
Figure 6 shows the pianoroll (top), class matrix (middle) and class vector (bottom) for the prelude. Peaks in the class matrix correspond to 3-11B (major triads), 4-27B (dominant seventh chords), 6-Z25A (perfect authentic cadences) and 7-35 (diatonic scales). Figure 7 shows the window length statistics. From these plots it can be see that three and four mainly occur in windows of around 2 beats.

8.1.2 Dvorak-Op101-1

This piece was chosen for its regular structure. It is comprised of several distinct sections of contrasting tonal material. The piece starts with the main theme which appears unambiguous in its mode and tonal centre. This theme is repeated at intervals throughout the piece. A number of other sections can also clearly be identified. The sections appear to depart from the tonality of the main theme to varying degrees. Some sections appear similar to each other save for a transposition. The identifiable sections of the piece are as follows:
A Main theme in Eb (natural) minor
B Harmonic minor scale (D)
C Dolce, major mode
D Stacatto, major mode
C* Related to C
D* Related to D

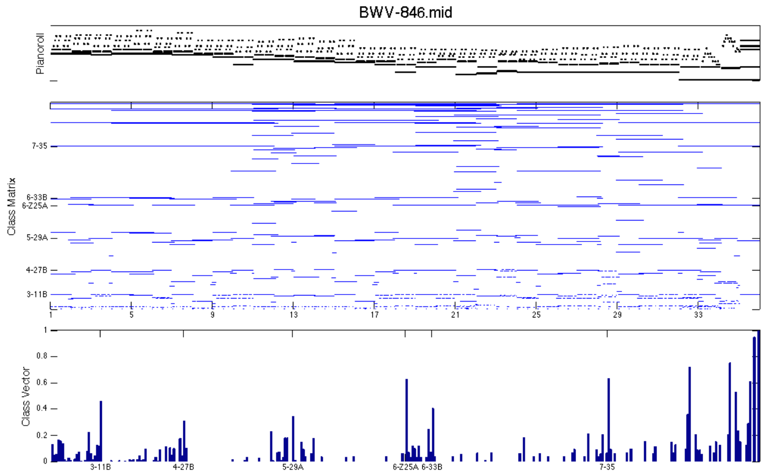

Figure 6: BWV-846: Pianoroll (top), class matrix (middle) and class vector (bottom)

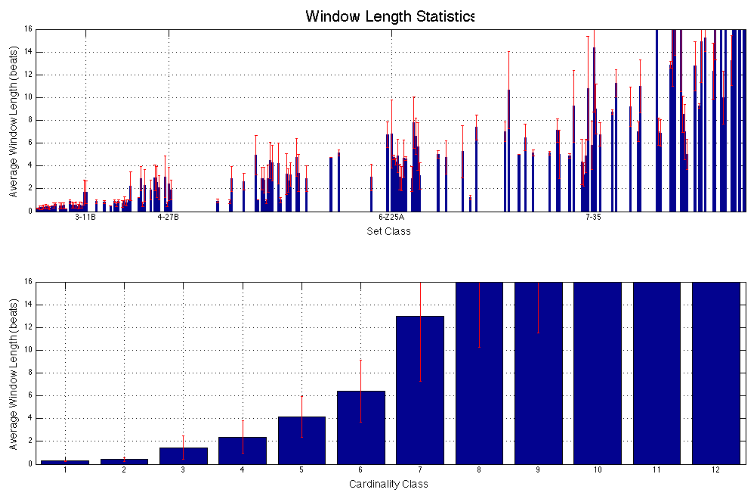

Figure 7: BWV-846: Window length statistics

E Meno Mosso, related to A
F Finale
Figure 8 shows the pianoroll(top), class matrix (middle) and class vector (bottom) for the piece. Peaks in the class matrix correspond to 4-26 (Min7), 5-27B (Min9), 6-32 (Min11), 6-Z25B (minor cadence) and 7-35. It should be noted that extended chords and cadences often have the same set class (e.g. Min11 and a $iv - i$ progression). Figure 9 shows the window length statistics for the piece.

8.2 Distance Time Series

- This example shows how the distance time series can be plotted to represent the tonal characteristics of a piece. To facilitate an initial, rudimentary analysis example, an reduction of the Bach prelude is used (BWV-846-Chords), in which each bar was replaced with a single semibreve chord containing every note from that bar. This reduction, in effect, replaces the piece with its underlying chord progression, removing the rhythmic element of the arpeggiation and providing a clearer expression of the tonal contents.
- BWV-846-Chords was segmented using 3 separate sliding windows.
- Figure 10 shows the pianoroll (top) and three distance time series, plotted as lines, with

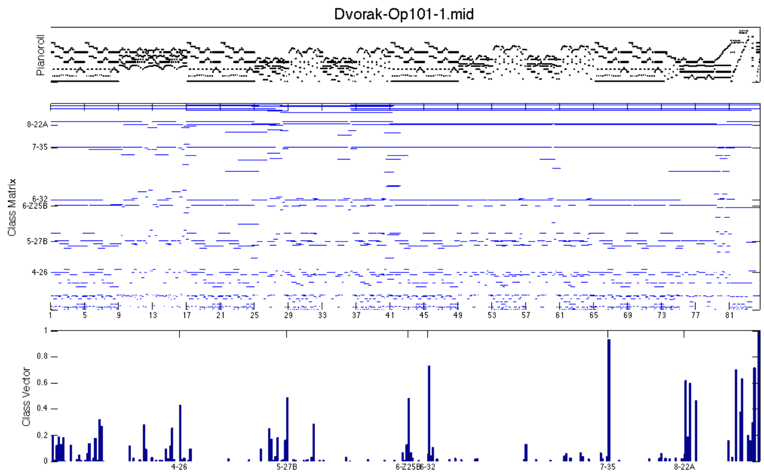

Figure 8: Dvorak-Op101-1: Pianoroll (top), class matrix (middle) and class vector (bottom)

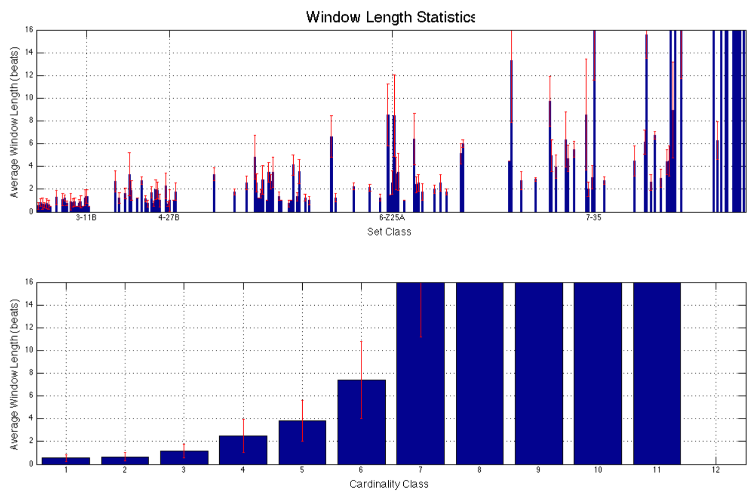

Figure 9: Dvorak-Op101-1: Window length statistics

reference sets 3-11B, 6-Z25A and 7-35.

- 3-11B: Single chords were targeted using a window length of 1 bar and a hop size of 1 bar. Thus, each point corresponds to a single bar/chord in the progression. Points where the curve is at zero correspond to bars containing major triads, while peaks in the curve correspond to more complex of less familiar chords.
- 6-Z25A: Cadential progression were targeted using a window length of 2 bars and a hop size of 1 bar. Thus, each point corresponds adjacent pairs of bars/chords. The occurrence of cadences is marked by zeros in the time series.
- 7-35: Diatonicism was targeted using a window length of 4 bars and a hop size of 1 bar. Common in diatonic music of this type are chord progressions that move by descending fifths, three of which in succession comprise a diatonic set (eg. ii-V-I). Areas of steady flatness at zero denote diatonic passages where as the higher points in the curve indicate more chromatic passages or less familiar scales.

8.3 Time Series Differential

- This example shows how the cadential punctuation of a musical piece can be detected by calculating the approximate differential of the distance time series.

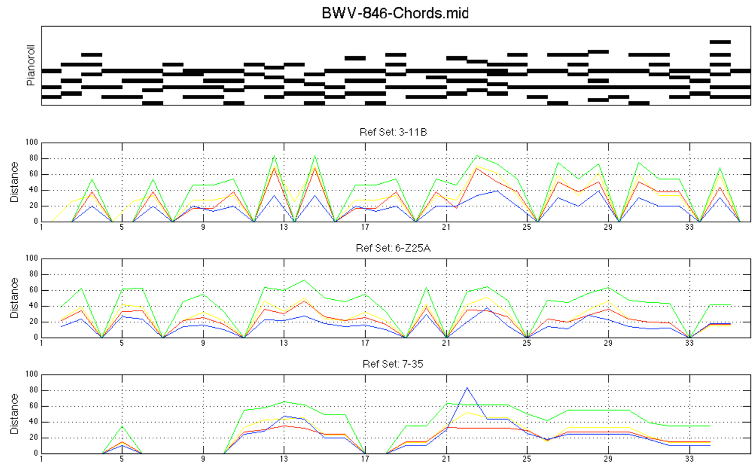

Figure 10: BWV-846-Chords: Distance time series targeting chords (top), cadences (middle) and diatonicism (bottom).

- BWV-846 was segmented so as to target 3 and 4 note chords using a window length of 2 beats. Window size selections was informed by Figure 7. A hop size of 1 beat was chosen so as to capture the cadential overlap of these chords.
- Figure 11 shows the class matrix and class vector for BWV-846 with the contents of the sliding superimposed on top in red. The class vector shows that a high proportion of major triads and seventh chords were captured at this time scale while the class matrix shows the position of these captured set classes in time.

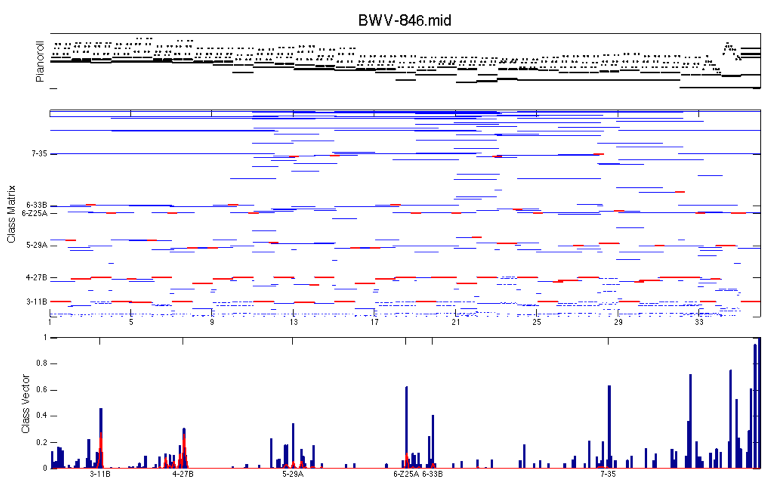

Figure 11: BWV-846: Pinaoroll (top), class matrix (middle) and class vector (bottom) with sliding window contents (red)

- Figure 12 shows the pianoroll (top), distance time series (middle) and approximate second order differential of the time series. The time series was computed using 6-Z25A as a reference set.
- The highest peaks correspond to the centre of windows containing perfect authentic cadences ($V^7 - I$) and denote the boundaries of distinct musical units. The first two peaks have been highlighted with red arrows and labelled “A” and “B”.
- The peak at “A” in bars 3-4 is at the conclusion of a full cadence in C which establishes the tonality of the prelude.
- The peak at “B” in bars 6-7 is where, following a modulation, the new key of G major is

confirmed with a $V^7 - I$ cadence.

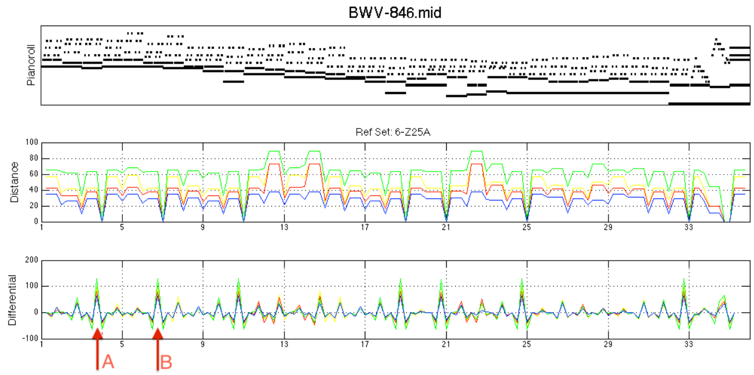

Figure 12: BWV-846: Cadential punctuation Example 1

- Other types of cadence can be observed by using different comparison sets, for example, 7-32B ($vii^{o7} - I$ cadence).
- A more general approach can involve exploiting the distance between subclasses within the cadential classes, such as that between the dominant and tonic chords.
- Figure 13 shows the approximate first order differential of the time series computed with reference to 3-11B (tonic chord).
- The negative peaks correspond to sudden drops in the time series resulting from the distance between dominant and tonic chords. The plot not only marks $V^7 - I$ cadences but also other types. An example has been highlighted with a red arrow and labelled "C".
- The peak marked by "C" in bars 14-15 is a $vii^{o7} - I$ cadence as part of the modulation back to C major.

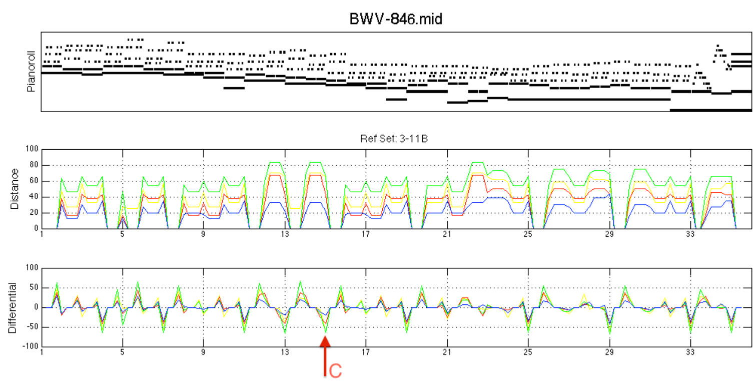

Figure 13: BWV-846: Cadential punctuation Example 2

8.4 Time Series Autocorrelation

- These examples demonstrates how autocorrelation of the distance time series can be used to detect repetitions and some structural aspects of the tonal progression.
- BWV-846 was segmented using a window length of 2 beats and a hop size of 1 beat so as to target 3 and 4 note chords.
- Figure 14 shows the autocorrelation of the time series which was calculated using a reference set of 3-11B. Peaks occur at regular intervals indicating a certain degree of structural repetition in the tonal progression. Put another way, the time varying distance between the music and 3-11B is periodic, repeating at 2 bar intervals.

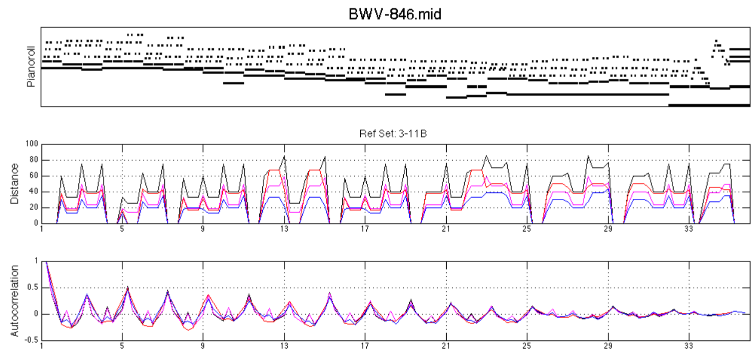

Figure 14: BWV-846: Detection of structural boundaries with autocorrelation

- Dvorak-Op101-1 was segmented so as to target 3 and 4 note chords using a window length of 2 beats and hop size of 1 beat. Window length selection was informed by figure 9.
- Figure 15 shows the pianoroll (top) with red dotted lines marking the boundaries between the structural elements of the piece described in 8.1.2. The distance time series (middle) was computed using a reference set of 3-11A. The autocorrelation (bottom) shows a similar type of periodicity as figure 14 and the pattern of major pattern of peaks correspond to the boundaries between section at regular 4 bar intervals.

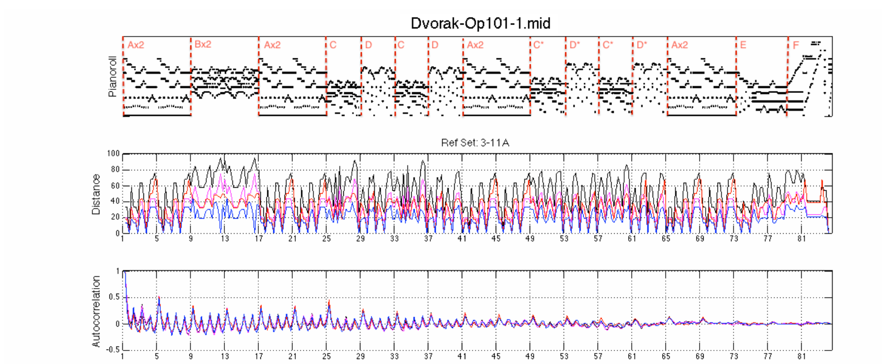

Figure 15: DvorakDvorak-Op101-1: Detection of structural boundaries with autocorrelation

8.5 Self-Similarity Matrix

- The self-similarity matrix is a widespread technique for detecting repetitions in a time series. In this context, the time series is the set class time series and the metric used is a set class similarity measure.
- This examples demonstrates how a self-similarity matrix computed in this way can be used to obtain structural information about a musical piece. Here, there is no need for a reference set as each window is systematically compared with every other.
- Dvorak-Op101-1 was segmented using a window length of 2 beats and hop size of 1 beat.
- Figure 16 shows the self-similarity matrix computed using the ATMEMB-prime distance.
- The area highlighted in red corresponds to section A, the main theme, and it can be clearly seen to repeat at various points within the piece. The area highlighted in blue shows how the sections C and D are related to sections C* and D*. The broken black diagonal down the middle of this section indicates the points at which the set class material of C* and D* are not identical to C and D.
- BWV-846 was segmented using a window length of 4 beats and hop size of 2 beats.
- Figure 17 shows the self similarity matrix computed using the ATMEMB-prime distance.

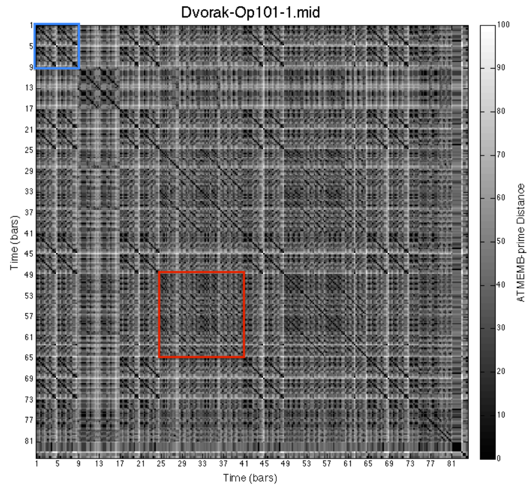

Figure 16: Dvorak-Op101-1: Self-similarity matrix (ATMEMB-prime)

- The area highlighted in red contains a black diagonal indicating the repetition of a 5 bar sequence: bars 7-11 in G major repeated in bars 15-19 in C major. Of additional interest is the region to the top left of the area where the diagonal is continued by way of several dark grey squares. This shows that there is a small and constant distance between these sections, indicating a degree of musical similarity that goes beyond mere transposition.

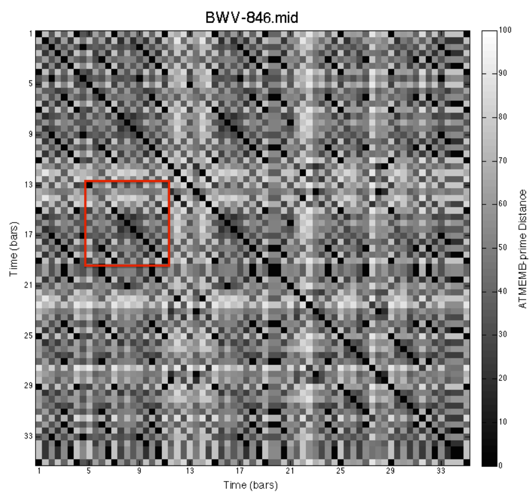

Figure 17: BWV-846: Self-similarity matrix (ATMEMB-prime)

8.6 Set Class Space

- This example demonstrates how multidimensional scaling can be used to visualise a geometric configuration of the set classes captured by a specific segmentation policy.
- Grids such as those shown in 6.5
- BWV-486-Chords was segmented with a window length of 1 bar and a hop size of 1 bar to obtain the basic chord progression.
- Figure 18 shows a grid containing the distance values between the captured set classes and a reference set of 3-11B. Each row contains the values from a different measure and gives a basic, 1-dimensional projection of the implicit set class space. WHAT DOES THIS SHOW?

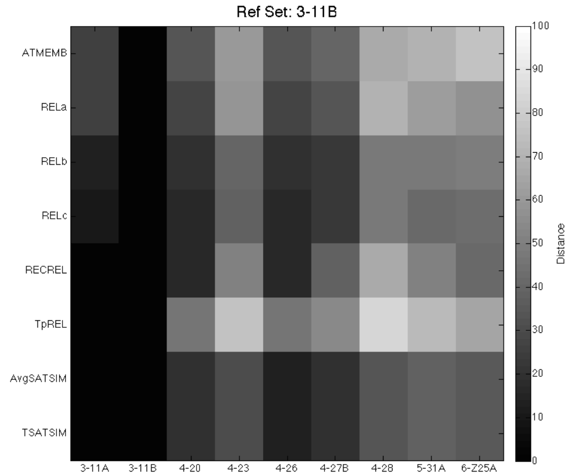

Figure 18: BWV-846-Chords: 1-dimensional set class space

- Figure 19 shows the 2-dimensional configuration resulting from non-metric multidimensional scaling of the RELa-prime distances between the captured set classes. The size of each point is proportional to the relative active duration and they are coloured according to a ternary cluster analysis. Although the global stress of the configuration is high (0.1967) it gives an indication as to the relationship between harmonic components. The green cluster contains the chords Maj, Min, 7, Maj7 and Min7 and these constitute the greater part of the music. The blue cluster contains the perfect cadence and 7sus4 chord which both contain a higher number of 4th intervals and occur less frequently. The red cluster contains chords Dim7 and Dim7b9 both of which contain high numbers of minor 3rds and also occur infrequently.

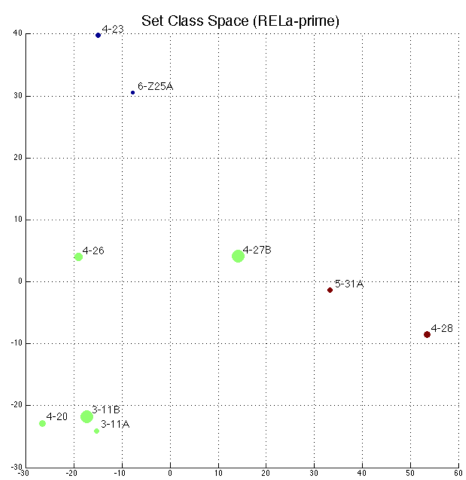

Figure 19: BWV-846-Chords: 2-dimensional chord space (RELa-prime)

- BWV-846-Chords was segmented using a window length of 4 bars and a hop size of 1 bar.
- Figure 20 shows a similar set class space based on larger sets (window length of 4 bars and hop size of 1 bar). The global stress is 0.0760.
- The clusters here can be interpreted by the cardinality of the sets and amount of chromaticism they contain. A space such as this might contain some familiar set classes and others less familiar. Visualisation of the space is helpful in understanding the material being analysed and could lead to selection of a less conventional reference based on its geometric location. A key concept to consider when viewing these spaces is that of orthogonality. By identifying dimensions that correspond to linearly independent properties, the set class space can be

better exploited in analysis and a better musical and/or mathematical comparison of the measures can be performed.

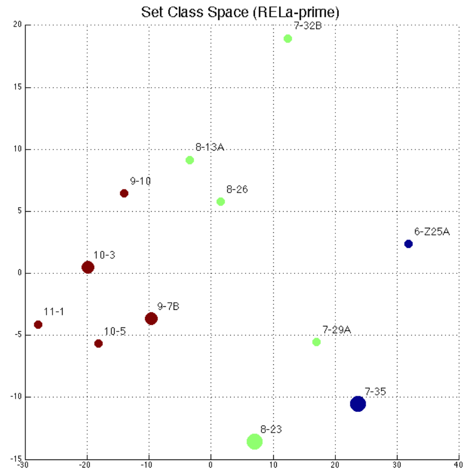

Figure 20: BWV-846-Chords: 2-dimensional scale space (RELa-prime)

Part III

Discussion

9 Conclusions

1. This work has successfully demonstrated the analytical potential of set class similarity measures.
2. Systematic set class descriptions of music have been discussed in terms of their perceptual and analytical relevance and have been shown to possess a lot potential as a starting point in many music research areas.
3. A comprehensive rendering of musicological concepts and terminology in the language of set class theory has facilitated the reconstruction of some basic traditional analysis.
4. In isolation, individual sets that correspond to chords of interest are of less interest than the hierarchical relationship between these sets' subsets and supersets and their evolution in time.
5. A comprehensive survey of set class similarity measures from the literature resulted in the selection of six models as suitable for systematic descriptive modelling.
6. Basic conclusions were drawn about the different capabilities and discriminatory power of the measures, however, it is presumed that no one measure is the best: "There is, after all, no single tool makes all other tools obsolete. It is up to each theorist and analyst to decide which are appropriate in any given circumstance" (Buchler, 1997).
7. Two segmentation policies have been presented to work in conjunction for extracting a set class description of a musical piece: Systematic and sliding window.
8. When working in the symbolic domain, the systematic segmentation can be used to supplement the analysis process.
9. The class matrix and class vector are concise and informative representations of systematic segmentation data and convey global characteristics about a musical piece.
10. The efficacy of a specific sliding window can be assessed by plotting the captured contents on top of the class matrix and class vector. This technique can be used to tune the window parameters so as to target the `sets of interest`.
11. Five techniques have been presented for representing set class information: Distance time series, differential, autocorrelation, self-similarity matrix and MDS. These techniques exploit set class similarity measures to retrieve structural information.
12. The distance time series has been shown to reflect intuitions about the tonal progression of a piece.
13. Approaches to parameter selection can be divided into three categories depending on the sets of interest: chords, cadences, scales.
14. By targeting common chords and using a reference set of the basic triad, the distance time series can be interpreted as reflecting harmonic complexity or perhaps musical tension.
15. By targeting two-chord progressions and using a reference set of a cadence, the distance time series can characteristic features points of cadential punctuation.
16. By targeting scales and using a reference set of the diatonic scale, the distance time series can indicate whether the music is harmonically stable, very chromatic or modulating.
17. The first and second order differentials of the time series can reveal the cadential punctuation of a piece or locate instances of any particular set class.
18. Repetitions in the distance time series can be quantified through autocorrelation. Peaks in the autocorrelation can point to important structural boundaries in a piece.
19. The self-similarity matrix can reveal the structural makeup of a piece including repeated sections and related sections.
20. Both autocorrelation and the self-similarity matrix are capable of capturing not only exact/transposed repetitions, but passages in which some quality of distance or ratio is preserved. These relationships are determined by the measure used and might have a sophisticated musicological or perceptual basis that is not easily observed from listening or from the score.
21. Visualisation of set class space through multi-dimensional scaling can give insight into the

relationship between tonal objects.

22. Through interpretation of the dimensions, set class spaces can be used to understand and compare different measures.
23. Set class spaces constructed using set classes from a specific timescale can be used to inform the reference set selection.
24. The large number of interdependent parameters prompted the development of an Analysis Tool for MATLAB in which these techniques can be used in conjunction, enabling the analyst to explore the numerous combinations and approaches. The demonstrations here are just the beginning of what could potentially be explored.
25. A systematic set class description combined with these representation techniques could be employed in MIR systems for the automatic detection of structure and musical similarity.

10 Future Work

1. A greater understanding of the set class contents of a piece could be achieved through a more exhaustive exploration of systematic set class descriptions of simple examples.
2. Understanding the hierarchical relationship between set classes and the role of the sets of interest would allow for a more discerning selection of parameters.
3. One particularly fascinating line of inquiry is whether the class matrix is unique for a given piece. Can pitch classes in the piece be changed without changing the class matrix? To what extent can a piece be reconstructed from its class matrix?
4. A more concrete and quantitative analysis of the discriminatory power of the similarity measures will better inform the selection of appropriate comparison sets and measure.
5. A more thorough mathematical evaluation of the measures could be starting point for comparison as well as yield information relevant to parameter selection.
6. Comparison of set class similarity measures with numerical distance measures such as Euclidean and Mahalanobis distance would be a necessary component in the justification of their use.
7. Examination of the implicit spaces created by different measures could provide an intuitive method of measure comparison. Interpretation of the dimensions could reveal the musical quantities that they are measuring.
8. Computationally combining different distance plots (multiplying, convolving, correlating etc.) could reinforce or weaken a particular analytical hypothesis.
9. The addition of peak selecting and structural marking in the analysis tool would inform tests as to the suitability of the proposed techniques for automatic structual segmentation.
10. The use of multiple measures, comparison sets and sliding windows could allow for a more finely tuned targeting of structural information.
11. The use of multiple distances simultaneously would allow the combining of approaches specified here. It would also be a step towards describing tonal progressions in multidimensional space.
12. The use of set class spaces could provide deeper understanding of tonal progressions by analogy to trajectories in space.

References

M. Babbitt. The Use of Computers in Musicological Research. *Perspectives of New Music*, 3(2): 74–83, 1965.

C. L. Bruner. The Perception of Contemporary Pitch Structures. *Music Perception: An Interdisciplinary Journal*, 2(1):25–39, 1984.

MH Buchler. Relative Saturation of Subsets and Interval Cycles as a Means for Determining Set-Class Similarity. (ic):1–204, 1997.

M. Castrén. *RECREL. A similarity measure for set-classes*. Sibelius Academy, Helsinki, 1994.

W. Chai. *Automated Analysis of Musical Structure*. PhD thesis, MIT, 2005.

E. Chew. *Towards a mathematical model of tonality*. PhD thesis, M.I.T., 2000.

R. Cohn. A Tetrahedral Graph of Tetrachordal Voice-Leading Space. *Music Theory Online*, 9(4), 2003. URL `http://www.mtosmt.org/issues/mto.03.9.4/mto.03.9.4.cohn_frames.html`.

D. Deutsch. The Procession of Pitch Combinations. In *Deutsch D. (ed.): The Psychology of Music*, chapter 9, pages 271–316. Academic Press, New York, 1982.

J. Foote. Automatic audio segmentation using a measure of audio novelty. In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE*, volume 00, pages 452–455, 2000. ISBN 0780365364.

A. Forte. *The Structure of Atonal Music*, volume 61. Yale University Press, London, 1973.

P. Gärdenfors. *Conceptual Spaces: The Geometry of Thought*, volume 4. MIT Press, Cambridge, MA, 2000.

D. Gibson. The Aural Perception of Nontraditional Chords in Selected Theoretical Relationships: A Computer-Generated Experiment. *Journal of Research in Music Education*, 34(1):5–23, 1986.

D. Gibson. The Aural Perception of Similarity in Nontraditional Chords Related by Octave Equivalence. *Journal of Research in Music Education*, 36(1):5–17, 1988.

D. Gibson. The Effects of Pitch and Pitch-class Content on their Aural Perception of Dissimilarity in Complementary Hexachords. *Psychomusicology*, 12(1):58–72, 1993.

E. Gómez. *Tonal description of music audio signals*. PhD thesis, UPF, 2006.

E. Huovinen and A. Tenkanen. Bird's-Eye Views of the Musical Surface: Methods for Systematic Pitch-Class Set Analysis. *Music Analysis*, 26(1-2):159–214, March 2007. ISSN 02625245.

B Hyer. Tonality. *Grove Music Online*, 2013.

E. J. Isaacson. Similarity of Interval-Class Content between Pitch-Class Sets : The IcVSIM Relation. *Journal of Music Theory*, 34(1):1–28, 1990.

E. J. Isaacson. Issues in the Study of Similarity in Atonal Music. *Music Theory Online*, 2(7), 1996. URL `http://www.mtosmt.org/issues/mto.96.2.7/mto.96.2.7.isaacson.html`.

C. L. Krumhansl. *Cognitive Foundations of Musical Pitch*. Oxford University Press, New York, 1990.

C. L. Krumhansl and R. N. Shepard. Quantification of the Hierarchy of Tonal Functions Within a Diatonic Context. *Journal of Experimental Psychology*, 5(4):579–594, 1979.

C L Krumhansl, G J Sandell, and D C Sergeant. The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music. *Music Perception: An Interdisciplinary Journal*, 5(1):31–77, 1987.

T. Kuusi. *Set-class and chord: examining connection between theoretical resemblance and perceived closeness*. PhD thesis, 2001.

R. C. Lane. *A Multidimensional Scaling Study of Seven Theoretical Indexes of Intervallic Similarity and Musicians' Perceptions Among Twenty-One Pitch-Class Sets*. Phd, University of North Carolina, 1997.

F. Lerdahl. *Tonal Pitch Space*. Oxford University Press, New York, 2001.

D. Lewin. A response to a response: on Pcset relatedness. *Perspectives of new music*, 18(1):498–502, 1979.

C. H. Lord. Intervallic Similarity Relations In Atonal Set Analysis. *Journal of Music Theory*, 25 (1):91–111, 1981.

A. Martorell. *Modelling Tonal Context Dynamics by Temporal Multi-Scale Analysis*. PhD thesis, UPF, 2013.

J. K. Millar. *The Aural Perception of Pitch-Class Set Relations: A Computer-Assisted Investigation*. Phd, University of North Texas, 1984.

R. Morris. A Similarity Index for Pitch-Class Sets. *Perspectives of New Music*, 18(1):445–460, 1979.

H. Purwins, B. Blankertz, and K. Obermayer. A new method for tracking modulations in tonal music in audio data format. In *International Joint Conference on Neural Network*, volume 6, pages 270–275, 2000. ISBN 0-7695-0619-4.

I. Quinn. General Equal-Tempered Harmony (Introduction and Part I). *Perspectives of New Music*, 44(2):114–158, 2006.

I. Quinn. General Equal-Tempered Harmony: Parts 2 and 3. *Perspectives of New Music*, 45(1): 4–63, 2007.

J. Rahn. Relating sets. *Perspectives of New Music*, 18(1):483–498, 1979. URL `http://www.jstor.org/stable/10.2307/832998`.

H. Riemann. *Musikalische Syntaxis: Grundriss einer harmonischen Satzbildungslehre*. Leipzig, 1877.

S. Ring. *Tonality and Transformation*. Oxford University Press, 2011.

D. W. Rogers. *Interval-Class Content and Similarity Functions: A Critique of a Critique*. Eastman School of Music, A seminar handout, 1992.

A. Samplaski. Mapping the Geometries of Pitch-Class Set Similarity Measures via Multidimensional Scaling. *Music Theory Online*, 11(2), 2005. URL `http://www.mtosmt.org/issues/mto.05.11.2/mto.05.11.2.samplaski.html`.

A. Schoenberg. *Structural Functions of Harmony*. Norton, New York, 1969 edition, 1954.

D. Scott and E. J. Isaacson. The interval angle: A similarity measure for pitch-class sets. *Perspectives of New Music*, 36(2):107–142, 1998.

R. N. Shepard. The Analysis of Proximities: Multidimensional Scaling With an Unknown Distance Function (I & II). *Psychometrika*, 27(2,3):125–140, 219–246, 1962.

R. N. Shepard. Geometrical approximations to the structure of musical pitch. *Psychological Review*, 89(4):305–333, 1982. ISSN 1939-1471(Electronic);0033-295X(Print).

D. Stammers. *Set Theory in the Perception of Atonal Pitch Relations*. Phd, Cambridge University, 1994.

R. Teitelbaum. Intervallic Relations in Atonal Music. *Journal of Music Theory*, 9(1):72–127, 1965.

D. Tymoczko. *A Geometry of Music: Harmony and Counterpoint in the Extended Common Practice*. Oxford University Press, New York, 2011.

D. Tymoczko. The Generalized Tonnetz. *Journal of Music Theory*, 56(1):1–5, 2012.

G. Tzanetakis and P. Cook. A framework for audio analysis based on classification and temporal segmentation. In *EUROMICRO Conference, 1999. ...*, volume 2, pages 61–67. Ieee, 1999. ISBN 0-7695-0321-7.

G. Weber. *The Theory of Musical Composition*. Cocks, London, 1851.

G. A. Wiggins. Semantic Gap?? Schemantic Schmap!! Methodological Considerations in the Scientific Study of Music. In *2009 11th IEEE International Symposium on Multimedia*, pages 477–482, 2009. ISBN 978-1-4244-5231-6.

G. A. Wiggins. Music, mind and mathematics: theory, reality and formality. *Journal of Mathematics and Music*, 6(2):111–123, July 2012. ISSN 1745-9737.

Part IV
Appendices

A Set Class Similarity Measures

This chapter contains a concise summary of the set class similarity measures from the literature organised by theorist. Each section specifies the publication in which the measure was proposed and brief description of the theoretical approach adopted by the theorist. A mathematical formula is given where possible using standard notation. A reference for notation can be found in A.1 and commonly used symbols are defined in the glossary. Where a mathematical formula does suffice, the comparison procedure is described in words. In addition, each section contains a table specifying important statistics:

- SC-Type: the type of SC the measure compares (Tn or TnI)
- Cardinality: whether the measure can compare SCs of different cardinalities.
- Vector Type: the type of vector used in the comparison procedure (see 3.3).
- Criteria Met: a list of Castren's criteria which the measure meets.
- I-related: whether the measure discriminates between inversionally related sets.
- Z-related: whether the measure discriminates between Z-related sets.

A.1 On Notation

Many of the formulas for similarity measures in the following sections appear differently to the way they were originally published. The reason for this is an attempt to standardise their symbolic representation through common vector notation in order to illuminate and compare the underlying mathematical concepts. Below are definitions of the of the required symbols.

Difference Vector is the absolute difference between corresponding terms in the nCVs of two SCs, X and Y:

$$DV(nCV(X), nCV(Y)) = |nCV(X) - nCV(Y)|$$

Vector Magnitude is the length of the nCV in euclidean space:

$$\|nCV(X)\| = \sqrt{\sum_{i=1}^{\#nC} (nCV(X)_i)^2}$$

Unit Vector is the normalised nCV (unit length):

$$n\hat{CV}(X) = \frac{nCV(X)}{\|nCV(X)\|}$$

Euclidean Distance is the distance between the points defined by two nCVs in n-dimensional Euclidean space:

$$d(X, Y) = \sqrt{\sum_{i=1}^{n} (X_i - Y_i)^2} = \|DV(X, Y)\|$$

A.2 MORRIS

A.2.1 K

Presented in Morris (1979, pp. 448), the K measure gives the number of intervals-classes (dyad-classes) shared by two SCs, X and Y.

$$K(X, Y) = \sum_{i=1}^{6} MIN(x_i, y_i)$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.3,C3.4,C4
I-related:	No
Z-related:	No

A.2.2 SIM

Presented in Morris (1979, pp. 446), SIM compares the ICVs of two SCs (the value is the cardinality of the DV).

$$SIM\left(X,Y\right) = \#DV\left(ICV\left(X\right), ICV\left(Y\right)\right)$$

SIM is also a function of K:

$$SIM(X,Y) = \#ICV(X) + \#ICV(Y) - 2.K(X,Y)$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.3,C3.4,C4
I-related:	No
Z-related:	No

A.2.3 ASIM

Presented in Morris (1979, pp. 450), ASIM (Absolute SIM) is a scaled version of SIM to address criteria C3.1. Scaling is done as a final step. Whilst the scale of values is now fixed, the resolution is course when cardinalities differ greatly.

$$ASIM\left(X,Y\right) = \frac{SIM\left(X,Y\right)}{\#ICV\left(X\right) + \#ICV\left(Y\right)}$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4,C4
I-related:	No
Z-related:	No

A.3 LORD

A.3.1 sf

Presented in (Lord, 1981, pp. 93), sf (Similarity Function) is similar to SIM but developed independently. sf is a subset of SIM:

$$sf\left(X,Y\right) = \frac{\#DV\left(ICV\left(X\right), ICV\left(Y\right)\right)}{2} = \frac{SIM(X,Y)}{2}$$

SC Type:	TnI
Cardinality:	Same
Vector Type:	ICV
Criteria Met:	C3.3,C3.4,C4
I-related:	No
Z-related:	No

A.4 TEITELBAUM

A.4.1 s.i.

Presented in Teitelbaum (1965, pp. 88), s.i. (Similarity Index) is the Euclidean distance between the Cartesian coordinates defined by the ICVs of two SCs. This is equivalent to the magnitude of the difference vector.

$$s.i.(X,Y) = \|DV(ICV(X), ICV(Y))\|$$

SC Type:	TnI
Cardinality:	Same
Vector Type:	ICV
Criteria Met:	C3.3,C3.4
I-related:	No
Z-related:	No

A.5 ROGERS

A.5.1 IcVD_1

Presented in Rogers (1992), IcVD_1 (Distance Formula 1) is a modification of SIM (A.2.2). The ICV components are scaled before being summed. IcVD_1 is related to Castren's $\%\text{REL}_2$ (A.9.3): $\%\text{REL}_2(\text{X,Y}) = \text{IcVD}_1(\text{X,Y}) \times 50$.

$$IcVD_1(X,Y) = \#DV\left(\frac{ICV(X)}{\#ICV(X)}, \frac{ICV(Y)}{\#ICV(Y)}\right)$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4,C4
I-related:	No
Z-related:	No

A.5.2 IcVD_2

Presented in Rogers (1992), IcVD_2 (Distance Formula 2) is similar to s.i. (A.4.1), but instead returns the Euclidean distance between the ends of the normalised ICVs.

$$IcVD_2(X,Y) = \left\|DV(IC\hat{V}(X), IC\hat{V}(Y))\right\|$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4
I-related:	No
Z-related:	No

A.5.3 Cos(θ)

Presented in Rogers (1992), Cosθ, gives the cosine of the angle between the ICVs in six-dimensional Euclidean space. As the angle decreases the similarity approaches 1.

$$Cos\theta(X,Y) = \frac{ICV(X) \cdot ICV(Y)}{\|ICV(X)\| \times \|ICV(Y)\|}$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4
I-related:	No
Z-related:	No

A.6 RAHN

A.6.1 AK

Presented in Rahn (1979, pp. 489), AK is an absolute or adjusted version of Morris' K (A.2.1), addressing the C3.1 criteria. AK is related to Morris' ASIM: AK(X,Y)=1-ASIM(X,Y).

$$AK\left(X,Y\right)=\frac{2K\left(X,Y\right)}{\#ICV\left(X\right)+\#ICV\left(Y\right)}$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4,C4
I-related:	No
Z-related:	No

A.6.2 MEMB$_n$

Presented in Rahn (1979, pp. 492), MEMB$_n$ (Mutual Embedding Number) compares the nCVs of two SCs for one nC at a time. It measures the mutual embedding of subsets such that only non-zero components of the nCVs contribute. By setting n = 2 (MEMB$_2$) it compares ICVs.

$$MEMB_n\left(X,Y\right)=\sum_{i=1}^{\#nC}nCV(X)_i+nCV(Y)_i$$

such that nCV(X)$_i$>0 and nCV(Y)$_i$>0.

SC Type:	TnI or Tn
Cardinality:	Any
Vector Type:	nCV
Criteria Met:	C1,C2,C3.3,C3.4,C4
I-related:	Yes*
Z-related:	Yes*

A.6.3 TMEMB

Presented in Rahn (1979, pp. 492), TMEMB (Total Mutual Embedding Number) counts the mutually embedded subsets of every cardinality. TMEMB is a total measure.

$$TMEMB\left(X,Y\right)=\sum_{n=2}^{12}MEMB_n\left(X,Y\right)$$

SC Type:	TnI or Tn
Cardinality:	Any
Vector Type:	nCV
Criteria Met:	C1,C2,C3.3,C4,C5
I-related:	Yes
Z-related:	Yes

A.6.4 ATMEMB

Presented in Rahn (1979, pp. 494), ATMEMB (Adjusted Total Mutual Embedding Number) is a scaled version of TMEMB to address criteria C3.1 (like SIM and ASIM; A and AK). ATMEMB is a total measure.

$$ATMEMB\,(X,Y) = \frac{TMEMB\,(X,Y)}{2^{\#X} + 2^{\#Y} - (\#X + \#Y + 2)}$$

SC Type:	TnI or Tn
Cardinality:	Any
Vector Type:	nCV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4,C4,C5
I-related:	Yes
Z-related:	Yes

A.7 ISAACSON

A.7.1 AMEMB2

Proposed by Isaacson (1990, pp. 8), AMEMB_2 (Adjusted MEMB_2) is a scaled version MEMB_2 (A.6.2), measuring the mutual embedding of ICs.

$$AMEMB_2 = \frac{2 \times MEMB_2(X,Y)}{(\#X\,(\#X-1) + \#Y\,(\#Y-1))}$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1

A.7.2 IcVSIM

Presented in Isaacson (1990, pp. 18), IcVSIM (Interval-Class Vector Similarity Relation) is the standard deviation of the entries in the ICVs of two SCs. IcVSIM is a scaled version of s.i. (A.4.1). IdV_i is the ith term in the vector defined by ICV(X)-ICV(Y) and $\overline{DV}$ is the average (mean) of its entries.

$$IcVSIM(X,Y) = \sqrt{\frac{\sum(IdV_i - \overline{IdV})^2}{6}}$$

SC Type	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.4
I-related:	No
Z-related:	No

A.7.3 ISIM2

Presented in Isaacson (1996), ISIM2 is a scaled version of IcVSIM (A.7.2). The squre root is taken of each term in the ICVs. Isaacson argues that each additional instance of an IC contributes less to similitude. However, Samplaski (2005) found ISIM2 to be inconsistent with itself when applying MDS to the values produced.

SC Type	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.4
I-related:	No
Z-related:	No

A.7.4 ANGLE (Isaacson & Scott)

Scott and Isaacson (1998) proposes a geometric method which is identical to that of Cos/theta (A.5.3) but instead gives the size of the angle in degrees.

$$ANGLE(X,Y) = \arccos Cos\theta(X,Y)$$

SC Type	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4
I-related:	No
Z-related:	No

A.8 LEWIN

A.8.1 REL

Presented in Lewin (1979), REL compares the nCVs of two SCs for all the nCs. Like MEMB_n (A.6.2), REL only considers non-zero entries however, this is achieved by multiplication (taking the geometric mean) of corresponding nCV terms.

$$REL(X,Y) = \frac{\sum_{i=1}^{p} \sqrt{SUB(X)_i \times SUB(Y)_i}}{\sqrt{\#SUB(X) \times \#SUB(Y)}}$$

where SUB(X) consists of concatenated nCVs and has a length p.

SC Type:	TnI or Tn
Cardinality:	Any
Vector Type:	nCV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4,C4,C5
I-related:	Yes
Z-related:	Yes

A.8.2 REL_2

Rahn (1979) suggested a number of manifestations of the basic REL concept including REL_2 which measures only intervallic similarity.

$$REL_2(X,Y)\frac{2 \times \sum \sqrt{(x_i y_i)}}{\sqrt{(\#X(\#X-1)\#Y(\#Y-1))}}$$

SC Type:	TnI
Cardinality:	Any
Vector Type:	ICV
Criteria Met:	C1,C2,C3.1,C3.2,C3.4
I-related:	No
Z-related:	No

A.9 CASTREN

A.9.1 Castrén's Difference Vector

Castrén specifies a different type of DV, which we shall call cDV to distinguish it from the regular DV. It consists of two rows, $cDV_x(X,Y) = X - Y$ and $cDV_y(X,Y) = Y - X$. Any negative values in either of the rows are set to zero. In addition Castren defines the weighted difference vector (wcDV) of two vectors X and Y as:

$$wcDV = \frac{cDV(X,Y)}{\#cDV(X,Y)} \times 100$$

A.9.2 nC%V

Presented in Castrén (1994) for use in %REL$_n$, nC%V(X) (n-class subset percentage vector) gives the percentage subset-class contents of an SC, X. The 2C%V is the Interval percentage vector.

$$nC\%V(X)_i = \frac{nCV(X)_i}{\#nCV(X)} \times 100$$

A.9.3 %REL$_n$

Presented in Castrén (1994), %REL$_n$ (Percentage Relation) is a modification of sf (A.3.1) using the nC%Vs (A.9.2) instead of ICVs. %REL$_n$ can be used as a stand-alone measure, however it is primarily intended as an intermediate step in T%REL and RECREL (A.9.4 and A.9.5).

$$\%REL_n(X, Y) = \frac{\#DV(nC\%V(X), nC\%V(Y))}{2}$$

SC Type	TnI or Tn
Cardinality:	Any
Measure Type:	Single nC
Vector Type:	nC%V
Criteria Met:	C1,C2,C3.1,C3.2,C3.3,C3.4,C4
I-related:	Sometimes
Z-related:	Sometimes

A.9.4 T%REL

Presented in Castrén (1994), T%REL (Total Percentage Relation) is the mean average of the values of %REL$_n$ for all values of n from 2 to m where, if $\#X \neq \#Y$, $m = MIN(\#X, \#Y)$ else $m = \#X - 1$.

$$T\%REL(X, Y) = \frac{\sum_{n=2}^{m} \%REL_n\,(X, Y)}{m - 1}$$

SC Type:	TnI or Tn
Cardinality:	Any
Measure Type:	Total
Vector Type:	nC%V
Criteria Met:	C1,C2,C3.1,C3.2,C3.3,C3.4,C4,C5
I-related:	Yes
Z-related:	Yes

A.9.5 RECREL

Presented in Castrén (1994), RECREL (Recursive Relation) recursively compares the subsets and subsets of subsets of two SCs using %REL$_n$ (A.9.3). The comparison procedure is quite complicated and potentially involves evaluating %REL$_n$ thousands of times. The full algorithm is explained in Castrén (1994).

SC Type:	TnI or Tn
Cardinality:	Any
Measure Type:	Total
Vector Type:	nC%V
Criteria Met:	All
I-related:	Yes
Z-related:	Yes

A.10 BUCHLER

A.10.1 nSATV

Presented in Buchler (1997, chap. 2.3) nSATV(X) (Saturation Vector) is a dual vector consisting of two rows, nSATV_A(X) and nSATV_B(X). It shows extent to which an SC is saturated with subclasses of cardinality n. The steps for computing nSATV(X) are as follows:

1. Compute the nCVs for all SCs of cardinality #X.
2. Find the minimum and maximum values for each vector position. These values form vectors $Max_n(\#X)$ and $Min_n(\#X)$.
3. Compute the following two vectors: $MaxMinus = DV(nCV(X), Max_n(\#X))$ and $MinPlus = DV(nCV(X), Min_n(\#X))$
4. $nSATV_A(X)_i = MIN(MaxMinus_i, MinPlus_i)$ and $nSATV_B(X)_i = MAX(MaxMinus_i, MinPlus_i)$
5. If $MaxMinus_i = MinPlus_i$, $nSATV_A(X)_i = MaxMinus_i$ and $nSATV_B(X)_i = MinPlus_i$

A.10.2 SATSIM_n

Presented in Buchler (1997, chap. 2.4), SATSIM_n (Saturation Similarity index) compares the nSATVs of two SCs and involves the following steps:

1. Calculate nSATV(X) and nSATV(Y)
2. Calculate the vectors $\text{nSATV}_{\text{row}}$(X) and $\text{nSATV}_{\text{row}}$(Y).
3. The function "row" maps the MaxMinus values of one nSATV to the MaxMinus values of the other. If $\text{nSATV}_A(\text{X})_i$ is a MaxMinus value and $\text{nSATV}_A(\text{X})_i$ is also a MaxMinus value, row = A ($\text{nSATV}_{\text{row}}(\text{X})_i = \text{nSATV}_A(\text{X})_i$), otherwise row = B.
4. Finally SATSIM_n(X,Y) is given by the formula:

$$SATSIM_n(X,Y) = \frac{\#DV(nSATV_A(X), nSATV_{row}(Y)) + \#DV(nSATV_A(Y), SATV_{row}(X))}{\#DV(nSATV_A(X), SATV_B(X)) + \#DV(SATV_A(Y), SATV_B(Y))}$$

SC Type:	TnI or Tn
Cardinality:	Any
Measure Type:	nC
Vector Type:	nSATV
Criteria Met:	C1,C2,C3.1
I-related:	Sometime
Z-related:	Sometime

A.10.3 AvgSATSIM

Presented in Buchler (1997, chap. 2.10), AvgSATSIM (Average Saturation Similarity index) is the mean of SATSIM_n values where $m = MIN(\#X, \#Y)$.

$$AvgSATSIM(X,Y) = \frac{\sum_{n=2}^{m-1} SATSIM_n(X,Y)}{m-2}$$

SC Type:	TnI or Tn
Cardinality:	Any
Measure Type:	TOTAL
Vector Type:	nSATV
Criteria Met:	C1,C2,C3.1,C5
I-related:	Yes
Z-related:	Yes

A.10.4 TSATSIM

Presented in Buchler (1997, chap. 2.10), TSATSIM (Total Saturation Vector Similarity index) is an extension of SATSIM_n. TSATSIM is the quotient of the sum of all SATSIM_n numerators and denominators for all values of n from 2 to m-1 where $m = MIN(\#X, \#Y)$.

SC Type:	TnI or Tn
Cardinality:	Any
Measure Type:	TOTAL
Vector Type:	nSATV
Criteria Met:	C1,C2,C3.1,C5
I-related:	Yes
Z-related:	Yes

B Total Measures: Additional Information

This appendix contains information regarding the specific formulation of the total measure, in particular, how they deal with comparisons involving the trivial forms.

B.1 Trivial Forms

Three of the 351 Tn-type SCs are known as trivial forms: 1-1, 11-1 and 12-1. Due to their lack of musical or harmonic significance, these SCs are usually excluded from the work of SC-theorists. However, it is important that they be included in any systematic description and that their similarity to other sets be given a meaningful value.

The total measures make comparisons based on the subset content of a set. SC 1-1, which has no subsets, is rarely accounted for in such measures and in these cases a simple method will be used: Comparisons involving X = 1-1 and Y will be given the value $1\#Y$. Thus, the value will be the ratios of the cardinalities with 1 indicating maximum similarity.

Table 7: Trivial Forms

1-1	{0}
11-1	{0,1,2,3,4,5,6,7,8,9,10}
12-1	{0,1,2,3,4,5,6,7,8,9,10,11}

B.2 Scale of Values

The values of each measure were adjusted to the same scale for comparability by the same method as Kuusi (2001, pp. 48)). This scale is from 0 to 100 with with 0 indicating maximum similarity. The modified values are signalled by adding the symbol "prime" to the name.

Table 8: Adjustment for MEASURE-prime scale

ATMEMB-prime(X,Y)	=	(1-ATMEMB(X,Y))*100
REL-prime(X,Y)	=	(1-REL(X,Y))*100
AvgSATSIM-prime(X,Y)	=	T%REL(X,Y)
TSATSIM-prime(X,Y)	=	RECREL(X,Y)
TpREL-prime(X,Y)	=	AvgSATSIM(X,Y)*100
RECREL-prime(X,Y)	=	TSATSIM(X,Y)*100

B.3 ATMEMB (Rahn)

Details on how to calculate ATMEMB are give in A.6.4. In his analysis of the measure, Castren concludes that "divisor term is flawed, resulting in values suggesting suspiciously high degrees of dissimilarity between SCs of clearly different cardinalities. The general reliability and usefulness of the measure is difficulty to determine" (Castrén, 1994, pp. 89). The trivial forms 11-1 and 12-1 are accommodated explicitly by the formulation of Rahn (1979), however SC 1-1 is not and thus values will be obtained using the method specified in B.1.

B.4 REL (Lewin)

Details on how to calculate REL are given in A.8.1. From the basic equation it is possible to define three different formulations depending on the exact nature of SUB(X). In each formulation the trivial forms 11-1 and 12-1 are accommodated. The three formulations are as follows:

1. SUB(X) consists of the concatenated nCVs from 2 to 12. Here comparisons involving SC 1-1 will be evaluated with the method specified in B.1.
2. SUB(X) consists of the concatenated nCVs from 1 to 12 ($1CV(X) = #X%). This formulation accommodates SC 1-1.
3. Martorell (2013) specifies an alternative formulation where SUB(X) begins with the ICV (2CV) followed by the concatenated nCVs from 1 to 12. This formulation accommodates SC 1-1.

B.5 AvgSATSIM and TSATSIM (Buchler)

Details on how to calculate AvgSATSIM and TSATSIM are given in A.10.3 and A.10.4 respectively. Comparisons involving SC 1-1 are not accommodated and thus the method specified in B.1 will be used to provide values. Comparisons involving SCs 11-1 and 12-1 are accommodated except for the single comparison that involves both. This is because their $\mathrm{MAX}_n(\#\mathrm{X})$ and $\mathrm{MIN}_n(\#\mathrm{X})$ vectors are equal and thus all terms of the nSATVs are 0. The value for this comparison will be set to 0 (indicating maximal similarity). For comparisons involving ICs the value will be given by $\mathrm{SATSIM}_2(\mathrm{X,Y})$ (see A.10.2).

B.6 T%REL and RECREL (Castren)

Details on how to calculate T%REL and RECREL are given in A.9.4 and A.9.5 respectively. Comparisons involving SCs 11-1 and 12-1 are accommodated in both by Castren's formulation. Comparisons involving SC 1-1 will be given values by the method specified in B.1. Castren comments that some T%REL values are too high to be intuitively plausible. Finally, it should be noted that the basic algorithm provided by Castren for calculating RECREL is not feasible for large sets. Comparisons of such sets require tables of pre-computed branch values.

C Set Class Reference

This appendix contains a reference guide for converting common musicological terminology into the language of set class theory. Table 9 contains common chord types and their corresponding set classes. Table 10 contains common scales and their corresponding set classes. Both tables also include the Forte Name and the Tn-type index of the set classes. Table 11 contains common chord pairs and cadential progression. Each position in the table contains the set class composed of the two chords corresponding to the row and the column. Chord symbols are in standard roman numeral notation.

Table 9: Chords to set classes

Chord	Set Class	Name	Idx
Maj	{0,4,7}	3-11B	25
Maj7	{0,1,5,8}	4-20	57
Maj9	{0,1,3,5,8}	5-27A	116
Maj11	{0,1,3,5,6,8}	6-Z25A	176
Maj13	{0,1,3,5,6,8,10}	7-35	276
Maj(6)	{0,3,5,8}	4-26	64
Maj(6/9)	{0,2,4,7,9}	5-35	130
7	{0,3,6,8}	4-27B	66
9	{0,2,4,6,9}	5-34	129
11	{0,2,4,6,7,9}	6-33B	189
13	{0,1,3,5,6,8,10}	7-35	276
7b9	{0,2,3,6,9}	5-31B	125
min	{0,3,7}	3-11A	24
min7	{0,3,5,8}	4-26	64
min9	{0,3,5,7,8}	5-27B	117
min11	{0,2,4,5,7,9}	6-32	187
min13	{0,1,3,5,6,8,10}	7-35	276
min(6)	{0,2,5,8}	4-27A	65
min(b6)	{0,1,5,8}	4-20	57
min(6/9)	{0,2,5,7,8}	5-29B	121
min7b9	{0,3,5,6,8}	5-23B	113
min/Maj7	{0,1,4,8}	4-19A	55
min/Maj9	{0,1,3,4,8}	5-Z17	98
min/Maj11	{0,1,3,4,6,8}	6-Z24B	174
min/Maj13	{0,1,3,4,6,8,10}	7-34	275
Dim	{0,3,6}	3-10	23
hDim7	{0,2,5,8}	4-27A	65
hDim7(9)	{0,2,4,5,8}	5-26A	114
Dim7	{0,3,6,9}	4-28	67
Dim7(b9)	{0,1,3,6,9}	5-31A	124
Aug	{0,4,8}	3-12	26
Aug7	{0,2,4,8}	4-24	62
Aug(maj7)	{0,3,4,8}	4-19B	56
Aug(maj9)	{0,3,4,6,8}	5-26B	115
Aug(maj11)	{0,1,3,5,6,9}	6-Z28	181
Aug(maj13)	{0,1,3,4,6,8,9}	7-32A	272
Sus4	{0,2,7}	3-9	22
7Sus4	{0,2,5,7}	4-23	61
Maj7Sus4	{0,2,6,7}	4-16B	51
Sus2	{0,2,7}	3-9	22
7Sus2	{0,3,5,7}	4-22B	60
Maj7Sus2	{0,4,5,7}	4-14B	47
N6 (bII)	{0,4,7}	3-11B	25
It6 (bVI7/no5)	{0,2,7}	3-8A	20
Fr6 (bVI7/b5)	{0,2,6,8}	4-25	63
Gr6 (bVI7)	{0,3,6,8}	4-27B	66

Table 10: Scales to set classes

Scale	Set Class	Name	Idx
Diatonic	{0,1,3,5,6,8,10}	7-35	276
Melodic Minor	{0,1,3,4,6,8,10}	7-34	275
Harmonic Minor	{0,1,3,4,6,8,9}	7-32A	272
Harmonic Major	{0,1,3,5,6,8,9}	7-32B	273
Neapolitan	{0,1,2,4,6,8,10}	7-33	274
Neapolitan Minor	{0,1,2,4,6,8,9}	7-30A	268
Double Harmonic	{0,1,2,5,6,8,9}	7-22	253
Hungarian	{0,1,3,4,6,7,9}	7-31A	270
Octatonic	{0,1,3,4,6,7,9,10}	8-28	322
Whole-Tone	{0,2,4,6,8,10}	6-35	192
Augmented	{0,1,4,5,8,9}	6-20	168
Pentatonic	{0,2,4,7,9}	5-35	130
Blues	{0,2,3,4,7,9}	6-Z47B	212

Table 11: Chord pairs and cadential progressions

	I	i	ii	ii7	iio	ii07	IV	iv	V	V7	bVI	vi	viio	vii07	viiO7
I	3-11B	4-17	6-33B	6-33B	6-Z24B	6-Z24B	5-27A	5-Z17	5-27A	6-Z25A	5-21B	4-26	6-Z25A	7-35	7-32B
i	4-17	3-11A	6-Z29	6-Z29	6-Z25B	6-Z25B	5-34	5-27B	5-Z17	6-Z24A	4-20	5-32A	6-Z24A	7-34	7-32A
ii	6-33B	6-Z29	3-11B	4-27B	5-16B	6-Z49	5-32B	6-Z49	5-27A	6-Z46A	6-30B	5-34	5-32A	5-32A	6-27A
ii7	6-33B	6-Z29	4-27B	4-27B	6-Z49	6-Z49	5-32B	6-Z49	6-Z25A	7-29A	6-30B	5-34	6-Z50	6-Z50	7-31A
iio	6-Z24B	6-Z25B	5-16B	6-Z49	3-10	4-27A	5-32A	4-27A	5-31B	5-31B	5-25A	6-31B	4-28	5-31A	4-28
ii07	6-Z24B	6-Z25B	6-Z49	6-Z49	4-27A	4-27A	5-32A	4-27A	6-Z29	6-Z29	5-25A	6-31B	5-31A	6-27A	5-31A
IV	5-27A	5-34	5-32B	5-32B	5-32A	5-32A	3-11B	4-17	6-33B	6-33B	5-32B	4-20	5-25A	5-25A	6-27A
iv	5-Z17	5-27B	6-Z49	6-Z49	4-27A	4-27A	4-17	3-11A	6-Z29	6-Z29	4-26	5-21A	5-31A	6-27A	5-31A
V	5-27A	5-Z17	5-27A	6-Z25A	5-31B	6-Z29	6-33B	6-Z29	3-11B	4-27B	6-Z19B	6-32	4-27B	5-34	5-31B
V7	6-Z25A	6-Z24A	6-Z46A	7-29A	5-31B	6-Z29	6-33B	6-Z29	4-27B	4-27B	7-32A	7-35	4-27B	5-34	5-31B
bVI	5-21B	4-20	6-30B	6-30B	5-25A	5-25A	5-32B	4-26	6-Z19B	7-32A	3-11B	5-22	6-27A	7-31A	6-27A
vi	4-26	5-32A	5-34	5-34	6-31B	6-31B	4-20	5-21A	6-32	7-35	5-22	3-11A	6-Z25B	6-Z25B	7-32A
viio	6-Z25A	6-Z24A	5-32A	6-Z50	4-28	5-31A	5-25A	5-31A	4-27B	4-27B	6-27A	6-Z25B	3-10	4-27A	4-28
vii07	7-35	7-34	5-32A	6-Z50	5-31A	6-27A	5-25A	6-27A	5-34	5-34	7-31A	6-Z25B	4-27A	4-27A	5-31A
viiO7	7-32B	7-32A	6-27A	7-31A	4-28	5-31A	6-27A	5-31A	5-31B	5-31B	6-27A	7-32A	4-28	5-31A	4-28