Evaluation of Pitch-Class Set Similarity Measures for Tonal Analysis (Literature Review)

Nicholas Harley - Masters Thesis UPF

March 28, 2014

Contents

1 Introduction

The objectives of the present work are principally concerned with descriptive modelling of tonality using PC-set theory. In particular, the analytical potential of set-class similarity measures will be assessed and evaluated through comparison with similar work. 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.

Chapters 2-5 comprise a 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 introduced the basic concepts of PC-set theory including set-class similarity measures. Chapter 4 gives a basic description of MDS 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.

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 (cited in Martorell 2013, pp. 6): "... 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 from just one type of description is unlikely to a definitive definition. 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 resembles 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 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) (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 so many approaches are based, whilst certainly possessing cognitive 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 analyitcal 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 Pitch-Class Set Theory

One such method available for systematic description of the musical surface is Pitch class set theory. PC-set 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 sequence of pitches. This chapter will outline the basic principles. A glossary of terms is included in 6 to assist with the number of new terms that will be introduced.

3.1 Pitch Class

Pitch-class set theory uses octave equivalence. A pitch-class (PC) is a number indicating the position of a note within the octave. For example, in Western equal temperament (TET) the octave is divided in to 12 steps. Each of these notes is given a number from 0-11 (Table 1).

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 and any segment of music can be represented as a PC-set.

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 (11-PC, Example 4). From these two transformations it is possible to define three types of SC: Tn, TnI and I, although I-types are not commonly used.

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 348 distinct Tn types.
Inversional (I):	All PC-sets that can be transformed to each other by inversion belong to the same SC. There are 197 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 220 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 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). There are 6 distinct interval-classes.

Table 2: Forte's Prime form and numbering convention

PC-set	{1,4,9}
Prime Form (TnI)	{0,3,7}
Prime Form (Tn)	{0,4,7}
Forte Name (TnI)	3-11
Forte Name (Tn)	3-11B

3.3 Vector Analysis

3.3.1 Membership and Inclusion

Two concepts that are crucial in PC-set 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 denoted Q $\subset$ S and means that all members of 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	$\{0,4,9\} \subset \{0,1,4,5,9\}$

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	200

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).

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 (Example 9).

Subset vectors form the basis of the majority of analysis performed by PC-set 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 PC-Set 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 approaches 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 the different approaches to similarity measures can only be approached by narrowing the focus to a specific type. Here we will focus 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 measure usually involve the comparison of the SC's 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 the 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 PC-set theoretical literature. Vector Type indicates whether the measure compares ICVs or nCVs. 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 Appendix ? (NOT INCLUDED FOR PEER REVIEW).

3.4.3 Castrens Criteria

In addition to his categorisation, Castrén (1994) proposes several criteria which a good similarity relation should meet. Later, these criteria will be used in assessing the specific capabilities of 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

Table 4: 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	nCV	ANY	nC
	T%REL	nCV	ANY	TOTAL
CASTREN	RECREL	nCV	ANY	TOTAL
	SATSIM	nCV	ANY	nC
	CSATSIM	ICV	ANY	nC
	TSATSIM	nCV	ANY	TOTAL
BUCHLER	AvgSATSIM	nCV	ANY	TOTAL

- 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.

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. Here some relevant studies are discussed.

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 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 exposure at an early age. 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. 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 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.

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 PC-set 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 sf. 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 PC-set Theory for Analysis

PC-set theory as means for descriptive modelling of tonality is not widely known outside of highly theoretical circles and the use of PC-set 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 comparison sets 7-1 (chromaticism) and 7-35 (diatonicism) using the REL distance. They claim that the visual results of their analysis "reflect pertinent aspects of our listening experience" (**?**, 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. A single SC can be represented by highlighting the segment or alternatively each segment can be shaded according to its REL distance from a comparison SC. He emphasises that class-scape is an exploratory tool rather than an automated analysis system.

4 Multidimensional Scaling

Multidimensional scaling (MDS) is a numerical visualisation technique that, given a matrix of pairwise 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 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 two values: Stress and r^2.

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.
r^2	r^2 is the percentage variability of the data being explained by the solution

By plotting stress against r^2 for a number of dimensionalities is possible to observe the point at which additional dimensions do not significantly improve the solution (the "elbow"). Ultimately, 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 constricts 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 comparisons. 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 any 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.).

6 PC-set Theory Glossary

- PC: pitch-class
- set: Unordered PC-set
- SC: Set-class
- nC: Cardinality-class
- IC: Interval-class, 2C
- Tn: Transposition
- I: Inversion
- TnI: Transposition/Inversion
- Tn-Type: Transpositional SC-type
- I-Type: Inversional SC-type
- TnI: Transpositional/inversional SC-type
- Prime Form: PC-set representing all members of an SC
- ICV: Interval-class vector
- nCV: n-class subset vector
- nC%V: n-subset class percentage vector
- SATV: Saturation vector
- CSATV: Cyclic saturation vectors
- DV: Difference vector
- #X: Set cardinality
- #nCV: Vector cardinality

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.

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.

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.

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.

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.

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

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

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.

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.

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.