TURUN YLIOPISTON JULKAISUJA
ANNALES UNIVERSITATIS TURKUENSIS

SARJA - SER. B OSA - TOM. 327

HUMANIORA

COMPARISON STRUCTURE ANALYSIS

by

Atte Tenkanen

TURUN YLIOPISTO
UNIVERSITY OF TURKU
Turku 2010

From the School of History, Cultural Research and Art Studies
Department of Musicology
University of Turku
FIN-20014 Turku
Finland

Supervisors:

Erkki Huovinen Visiting Associate Professor
School of Music
University of Minnesota

Professor Marcus Castrén
The DocMus Department of Doctoral Studies
in Musical Performance and Research
Sibelius Academy, Helsinki

Reviewers:

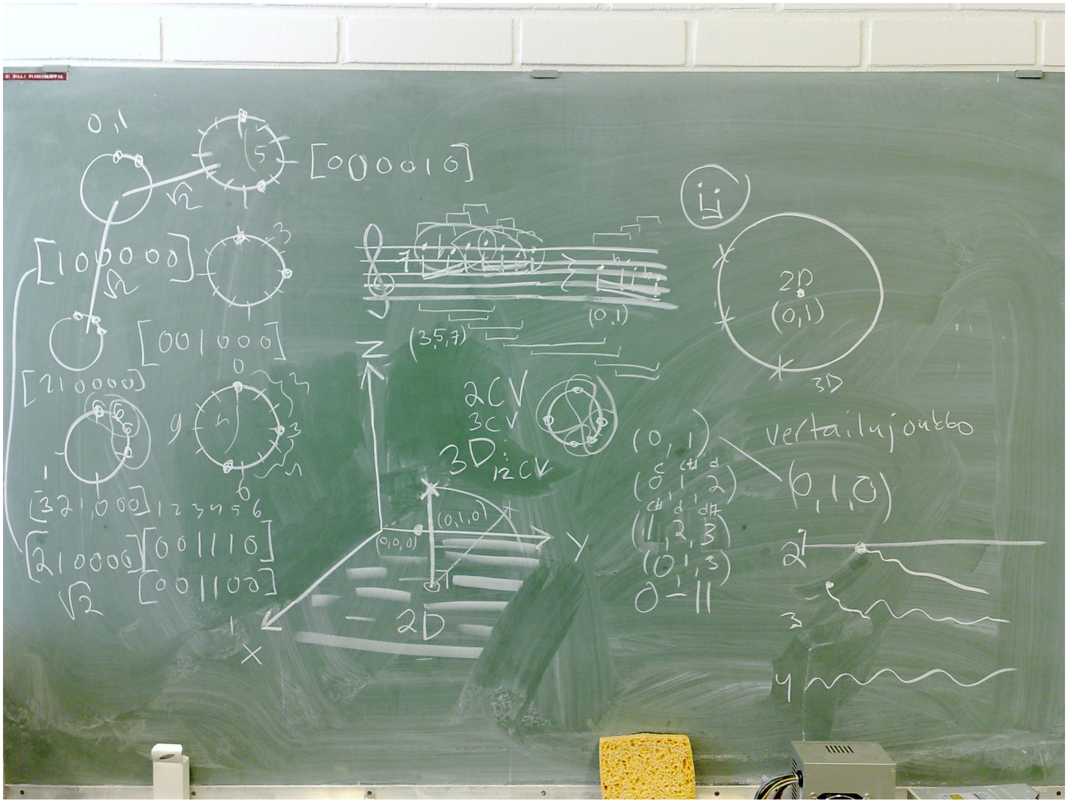
Darrell Conklin Ikerbasque Research Professor
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad del País Vasco
San Sebastián, ESPAÑA

Dr Tuukka Ilomäki
Sibelius Academy, Helsinki

Opponent:

Darrell Conklin Ikerbasque Research Professor
Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad del País Vasco
San Sebastián, ESPAÑA

ISBN 978-951-29-4283-1 (PRINT)
ISBN 978-951-29-4284-8 (PDF)
ISSN 0082-6987
Painosalama Oy – Turku, Finland 2010



A blackboard at the University of Turku.

Although computers are today more and more used in scientific research and in music research as well, the traditional blackboard still plays an important role in university studies.

Acknowledgements

This kind of research cannot be conducted alone in an isolated researcher's room, although it was mainly carried out at Turku University's Department of Musicology during the years 2003-2010. I wish to express my sincere gratitude to everyone I have worked with in Europe, the United States and China during these years.

First, I would like to express my gratitude to my supervisor and friend, Professor Erkki Huovinen, for his guidance and support throughout this study. You taught me to believe in research. Writing the first article together formed a very important period in the process. I would also like to thank my second supervisor, Professor Marcus Castrén, whose model has inspired me through these years: during this process, I have eventually begun to understand your thesis, which I merely used to admire years ago. The following persons have had an important influence on my work: my friend and organ improvisation teacher, Mr Olli Linjama, whose MIDI-recorded improvisations provided the original impulse for this research; and, Professor Emeritus of church music, Paavo Soinne, the supervisor of my master's thesis at the Sibelius Academy, who so often reminded me to consider postgraduate research – I just needed a little push to get started! My friend, statistician Mr Tommi Viitanen has taught me to program a computer with R. Without that indispensable skill, this project would have been a 'mission impossible'. My friend and collaborator, Fernando Gualda, with whom I have not only written papers and presentations using Skype (!) but have also spent a nice time in London, Paris and Turku! My warmest gratitude to both of you! Last but not least, I am indebted to Professor Darrell Conklin and Dr Tuukka Ilomäki, the reviewers of the thesis, for their valuable comments.

I pestered my teachers at the Department of Mathematics a lot. I would like to express my gratitude to Professors Jarkko Kari, Kari Ylinen, Iiro Honkala and Marko Mäkelä for their excellent help with my mathematical problems. I am especially grateful to the Lecturer in Mathematics, Kauko Lindström; your blackboard has a miraculous ability to supply students with mathematical comprehension! I warmly thank Ms Lena Helenius and Mr Erik Hieta for the excellent revision of the language of my articles. I want to thank my colleague, Dr Tuire Kuusi, who kindly offered me research material for the fifth article. I would like to thank Mr Timour Klouche, the editor of articles II and III, for his insightful questions, comments and suggestions during their completion process. I want to thank Professor Jukka Sarjala for managing all the necessary tasks for completing this project at the Department of Musicology. I also want to express my gratitude to my other colleagues, Yrjö Heinonen, Tiina Koivisto and Jopi Harri, for their valuable feedback in our seminars. I thank Marjaana Virtanen and Hanna Kaikko for their company and friendship during the long working hours and for nice conversations during the lunch and coffee breaks in our department.

Many friends and colleagues in Finland and abroad have helped me while preparing the thesis. I warmly thank my friends in Beijing and Shanghai who helped me a lot during the final stages of this project, Zhang Lan, Hu Teng Zhi, Zhang Qi Hui, Prima Alam and Stephen McIlroy. I especially thank Chen Nan and Jiang Shuo, together with their hospitable families, who offered me an opportunity to prepare the final two sections of the Introduction in their friendly 'full board': I was happy to finish the concluding chapter of my thesis on the train between Beijing and Shanghai. I thank my friend Tommi Stordell for his company, long talks about life during these years and taking part in my conference trip to New Haven. My father Tuomo and mother Raita have made this 'journey' possible through their emotional and financial support. Thank you! To my brother Marko and his family, I am thankful that we have been able to share our life experiences with each other during these years. Since so many people supported me, I would like to ask pardon for forgetting someone.

Finally, my warmest gratitude belongs to my family, my lovely wife Salla and my children, Niilo, Tuomas, Aino, Maria, Eeva, Elias, Linnea and Ada, who have stood by me the whole time. Salla, I thank you with all my heart for being here with me and sharing in my dreams and daily life as well as all the difficult moments during this process. Let's continue enjoying music and research!

My research was funded by the Finnish Cultural Foundation (Huhtamäen rahasto, together with Varsinais-Suomen rahasto), the Oskar Öflund Foundation and the Turku University Foundation. I thank them for supporting my dissertation financially.

Abstract

This study presents an automatic, computer-aided analytical method called Comparison Structure Analysis (CSA), which can be applied to different dimensions of music. The aim of CSA is first and foremost practical: to produce dynamic and understandable representations of musical properties by evaluating the prevalence of a chosen musical data structure through a musical piece. Such a comparison structure may refer to a mathematical vector, a set, a matrix or another type of data structure and even a combination of data structures. CSA depends on an abstract systematic segmentation that allows for a statistical or mathematical survey of the data. To choose a comparison structure is to tune the apparatus to be sensitive to an exclusive set of musical properties. CSA settles somewhere between traditional music analysis and computer aided music information retrieval (MIR). Theoretically defined musical entities, such as pitch-class sets, set-classes and particular rhythm patterns are detected in compositions using pattern extraction and pattern comparison algorithms that are typical within the field of MIR.

In principle, the idea of comparison structure analysis can be applied to any time-series type data and, in the music analytical context, to polyphonic as well as homophonic music. Tonal trends, set-class similarities, invertible counterpoints, voice-leading similarities, short-term modulations, rhythmic similarities and multiparametric changes in musical texture were studied.

Since CSA allows for a highly accurate classification of compositions, its methods may be applicable to symbolic music information retrieval as well. The strength of CSA relies especially on the possibility to make comparisons between the observations concerning different musical parameters and to combine it with statistical and perhaps other music analytical methods.

The results of CSA are dependent on the competence of the similarity measure. New similarity measures for tonal stability, rhythmic and set-class similarity measurements were proposed. The most advanced results were attained by employing the automated function generation – comparable with the so-called genetic programming – to search for an optimal model for set-class similarity measurements. However, the results of CSA seem to agree strongly, independent of the type of similarity function employed in the analysis.

Keywords: Comparison Structure Analysis, computational musicology, comparison structure, feature vector, musical segmentation, time-series analysis, similarity measure, musical surface, music information retrieval

Tiivistelmä

Tämä tutkimus esittelee uuden musiikkianalyyttisen metodin, *vertailurakenneanalyysin* (VRA, engl. *Comparison Structure Analysis*, CSA), jonka avulla voidaan analysoida musiikin eri ulottuvuuksia, kuten harmoniaa tai rytmiä. VRA:n ideana on mitata tietyn ennalta valitun musiikillisen rakenteen, vaikkapa jonkin sävelasteikon, vallitsevuutta musiikin kullakin ajanhetkellä. Tämä edellyttää kolmea asiaa. Ensiksi, intuitiivisesti tai muulla tavoin valittu musiikillinen piirre, jota tässä kutsutaan yleisesti *vertailurakenteeksi*, on esitettävä matemaattisessa muodossa, esimerkiksi matemaattisen avaruuden vektorina. Vertailurakenne voidaan muodostaa myös useiden eri tyyppisten, musiikin eri ulottuvuuksiin liittyvien tietorakenteiden yhdistelmänä. Toiseksi, analysoitava musiikillinen data, esimerkiksi musiikista muodostetut sävelluokat (C:stä H:hon), on pystyttävä ryhmittelemään vastaavantyyppisiksi objekteiksi. Lisäksi tarvitaan vielä matemaattinen funktio, joka kykenee mittaamaan valitun vertailurakenteen ja musiikista ryhmiteltyjen segmenttien välistä samankaltaisuutta tai vastaavasti, etäisyyttä. Toisin sanoen, VRA:ssa verrataan valittua vertailurakennetta, esimerkiksi diatonista asteikkoa, kaikkiin musiikista segmentoituihin vastaavantyyppisiin objekteihin. Mittaustulokset saadaan lukuarvoina yleensä välillä 0–1, jossa arvo 1 voi – mittausfunktion luonteesta riippuen – tarkoittaa joko täydellistä samankaltaisuutta tai suurinta mahdollista etäisyyttä. Havainnollisena analyysin kohteena voisimme kuvitella länsimaista taidemusiikkia edustavan sävellyksen, jossa siirrytään keskiaikaisesta diatonisesta musiikista historiallisesti ja tyylillisesti kohti 1900-luvun atonaalista musiikkia. Mikäli tässä tapauksessa vertailurakenteena käytettäisiin mainittua diatonista asteikkoa, VRA paljastaisi musiikissa korvinkin havaittavan ei-diatonisoitumisen. Tulosten esittämisellä esimerkiksi ajallisia muutoksia esittävin mittauskäyrin tai luokittelua havainnollistavin keskiarvopistein on merkittävä asema analyysissa. VRA sijoittuu perinteisen musiikkianalyysin ja tietokonetta hyödyntävien musiikin sisältöhakuun (music information retrieval, MIR) keskittyvien tekniikoiden välimaastoon. Sen avulla voidaan tunnistaa ja mitata perinteiselle musiikkianalyysille tyypillisia kohteita kuten karakteristisia rytmejä, sävelluokkajoukkoja, joukkoluokkia, tonaliteetteja ja käänteiskontrapunkteja soveltamalla MIR:lle tyypillisiä segmentointi- ja vertailualgoritmeja.

Vertailurakenneanalyysin suurimmaksi haasteeksi on osoittautunut musiikillisten segmenttien muodostamiseen tarvittavan automaattisen algoritmin kehittäminen. Voidaan näet osoittaa, että sama musiikillinen data on useimmiten mahdollista segmentoida – musiikillisesti mielekkäästi – monella eri tavalla. Silloin, kun kyse on harmoniaan liittyvistä objekteista, tehtävä on erityisen haastava, sillä tällöin musiikin säveltapahtumia joudutaan tarkastelemaan niin ajallisessa kuin vertikaalisessakin suunnassa. Musiikin tonaalisuudessa ja sävelluokkasisällössä tapahtuvien muutosten analysoimista varten tässä tutkimuksessa kehitettiinkin kaksi erilaista segmentointialgoritmia, jotka muodostavat musiikillisesta datasta osin limittäisiä sävelluokkajoukkoja. Metodien erilaisuudesta huolimatta ‘herkkyysanalyysillä’ voitiin osoittaa, että molemmat menetelmät ovat hyvin vähän riippuvaisia syötetyn datan luonteesta; niiden avulla saadut tulokset olivat hyvin samankaltaisia. VRA:lla saatuja tuloksia voidaan edelleen tarkastella myös tilastollisen merkitsevyyden näkökulmasta.

Koska VRA:lla pystytään havaitsemaan musiikin eri dimensioissa tapahtuvia muutoksia, tämän johdannaisena voidaan tutkia myös sitä, missä määrin jokin sävellys on tyylillisesti koherentti verrattuna johonkin toiseen sävellykseen eli kummassa muutokset ovat tarkasteltavan ominaisuuden

suhteen keskimäärin pienemmät ja kummassa suuremmat. Lisäksi VRA tarjoaa mahdollisuuden musiikin luokitteluun saatujen mittausarvojen perusteella: mitä enemmän musiikillisia parametrejä ja useampia vertailurakenteita analyysissa hyödynnetään, sitä tarkemmin sävellyksiä voidaan luokitella. Niinpä VRA:n keinoja voidaan tulevaisuudessa kuvitella käytettävän myös musiikin sisältöhakuun (MIR). Tällaisessa tapauksessa vertailurakenne tai -rakenteet voitaisiin 'laskea' musiikillisesta datasta suoraan jollakin matemaattisella menetelmällä – kuten pääkomponenttianalyysilla – etukäteen suoritettavan intuitiivisen valinnan sijaan.

Tutkimuksen tuloksiin lukeutuvat myös useat VRA:n tarpeisiin kehitetyt samankaltaisuusmittarit. Näistä mielenkiintoisin lienee sävelluokkajoukkojen välisen samankaltaisuuden mittaamiseen kehitetty funktio *expcos*, joka löytyi ns. *geneettisen ohjelmoinnin* avulla. Mainitussa kokeessa tietokoneella generoitiin arviolta n. 800 000 samankaltaisuusmittaria, joiden tuottamia tuloksia verrattiin ihmisten tekemiin samankaltaisuusarvioihin. Niistä n. 450 osoittautui käyttökelpoiseksi. Sensitiivisyysanalyysi osoitti, että em. funktio paitsi korreloi voimakkaammin empiiristen samankaltaisuusarvioiden kanssa, on VRA:ssa myös robustimpi kuin kenties tunnetuin samaan tarkoitukseen kehitetty funktio, *REL* (David Lewin, 1980). Käytännössä tällä ei ole kuitenkaan merkitystä: *REL* toimii VRA:ssa aivan yhtä hyvin kuin *expcos*.

VRA:n avulla musiikkia tarkastellaan ikään kuin jonkinlaisena tilastollisena sävelmassana, eikä se niin muodoin kykene kertomaan siitä, miten analysoitava musiikki on yksityiskohtien tasolla sävelletty; perinteiset musiikkianalyysimenetelmät pureutuvat tehtävään paremmin. Toisaalta, tämä ei ole VRA:n tarkoituskaan vaan päinvastoin, sen avulla sävellysten muodosta pystytään muodostamaan laajoja yleiskuvia, jotka ovat useimmiten havaintokykymme ulottumattomissa.

Vertailurakenneanalyysi on hyvin joustava menetelmä. Mikään ei nimittäin estä tarkastelemasta musiikin eri dimensioista saatuja mittaustuloksia keskenään ja näin etsimästä niiden välisiä yhteyksiä. Lisäksi menetelmän periaatteita voitaisiin kuvitella käytettävän yleisemminkin, esimerkiksi linnunlaulun muodon tarkasteluun tai vaikkapa jokipuron solinasta löytyvien toistuvien jaksojen havainnointiin. VRA:n periaatteita voidaankin soveltaa mihin tahansa numeerisesti diskreettiin muotoon saatettuun aikasarjaan.

Avainsanat:

Vertailurakenneanalyysi, tietokoneavusteinen musiikkianalyysi, piirrevektori, vertailurakenne, musiikin automaattinen segmentointi, aikasarja, samankaltaisuusmittari, musiikin pintataso, musiikin sisältöhaku

Definitions and abbreviations

Cardinality

The number of elements of a finite set is a natural number (non-negative integer), and is called the cardinality of the set. In this text, cardinality refers to the number of elementary units; for example, pitch classes or inter-onset-intervals in a vector or an unordered set (in which the order of items is irrelevant).

Correlation

In this presentation, the concept of correlation refers to the well-known statistical concept of the Pearson correlation coefficient (r) that varies between -1 and 1. It is a measure of the strength of a linear association between two variables.

Genetic programming

In artificial intelligence, genetic programming is an evolutionary algorithm-based methodology inspired by biological evolution to find computer programs that perform a user-defined task.

IOI

Inter-onset-interval (IOI) is the time between the attack-points of successive events or notes.

Middleground feature

In this study, a middle-ground feature refers to a musical feature in which the components are not straightforwardly connected to each other in the surface (foreground) level of the music. In other words, they cannot be derived in a straightforward manner from some symbolic representation of music, such as a list of MIDI events.

Pattern matching

Pattern matching is the act of checking for the presence of the constituents of a given pattern in research data, in which the pattern is rigidly specified.

Pitch-class set

Pc-set belongs to the main concepts of pitch-class set theory. It is an unordered collection of pitch classes ($pc \in \{0, ..., 11\}$). It is denoted by expressions such as '{5,11,2,4}'.

Scalar

In linear algebra, real numbers are called scalars. Scalar is a quantity used to multiply vectors in the context of vector spaces.

Set class

Set class is a collection of pitch-class sets that are considered equivalent under some canonical operations. In the present study, all set classes are related to each other by transposition and referred to as 'T_n-classes', in contrast to Forte's 'T_n/I-type' set classes, each of which includes a collection of pitch-class sets that are equivalent under transposition and/or inversion.[1] Following Castrén [2] and others, inversionally related set classes are here distinguished by adding the suffix ‘A’ or ‘B’ to the customary Forte names. The reason for using a T_n-classification instead of T_n/I-classification is that it makes the difference, for example, between the set-class representatives of the minor- and major-chords 3-11A[0,3,7], 3-11B[0,4,7], which has been proved to be useful in Comparison Structure Analysis. The subscript denotes the transposition interval in semitones modulo 12.

Statistical significance

In statistics, a result is called statistically significant if it is unlikely to have occurred by chance. The amount of evidence required to accept that an event is unlikely to have arisen by chance is known as the significance level or critical p-value.

Time-series

Time-series is a sequence of time ordered observations, measured typically at successive times and spaced at uniform time intervals.

Tonal stability

Tonal stability refers in the present study to the degree of closeness between tonal objects like pitch-class sets.

Vector

A geometric entity endowed with both length and direction. It is represented as an ordered set of n numbers $(a_1, a_2, ..., a_n)$. Vectors may be added together and multiplied (‘scaled’) by numbers, called *scalars*. A collection of vectors may form a mathematical structure called *vector space*, which is, however, not considered in the present study.

List of original publications

This thesis is based on the following publications, which will be referred by Roman numerals in the text.

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

II Tenkanen, A.: Tracking Features with Comparison Sets in Scriabin's Study op. 65/3. First International Conference of Mathematics and Computation in Music (MCM2007), Berlin. Timour Klouche and Thomas Noll (Eds). Springer series on *Communications in Computer and Information Science* 37, 211–219 (2009) **

III Tenkanen, A., Gualda, F.: Multiple Approaches to Comparison Structure Analysis. Ed. Timour Klouche. SIM & Springer: *Computational Musicology* 1. (accepted for publication)

IV Tenkanen, A.: Evaluating Tonal Distances between Pitch-Class Sets and Predicting Their Tonal Centres by Computational Models. Second International Conference of Mathematics and Computation in Music (MCM2009), New Haven. E. Chew, A. Childs, C-H. Chuan (Eds.). Springer series on *Communications in Computer and Information Science* 38, 245–257 (2009) **

V Tenkanen, A.: Searching for Better Measures: Generating Similarity Functions for Abstract Musical Objects. *Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems* (ICIS 2009), Shanghai, Vol. 4, 472–476 (2009) ***

Other publications related to the topic of the thesis:

Tenkanen, A.: A Linear Algebraic Approach to Pitch-Class Set Genera. First International Conference of Mathematics and Computation in Music (MCM2007), Berlin. Timour Klouche and Thomas Noll (Eds). Springer series on *Communications in Computer and Information Science,* Vol. 37, 521–530 (2009)

Tenkanen, A. and Huovinen, E.: Laajan musiikillisen korpuksen säveltasorakenteiden tietokoneavusteinen tutkimus (Studying the pitch organisational features of a large compositional corpus). *Suomen musiikintutkijoiden symposiumin satoa*, 161–165, Jyväskylä (2005)

Contents

1 Introduction

If our analytical problem is so well defined that we can solve it either with the computer or by hand, we today more often resort to the computer. The calculator is used because it counts faster than we do and because it does our tasks accurately. As the amount of data or the complexity of the task increases, the horizon of manual alternatives soon dissipates and no other possibility except the computational approach remains.

1.1 Computer-assisted music analysis

Similar to many other research fields, the field of music analysis has changed quite dramatically after the advent of the personal computer. Old analytical methods are still in use, but there are more and more computational approaches that are capable of confirming or questioning manually-achieved results and which, in addition, give rise to and help solve new research questions. Today, computer-assisted music analysis can borrow methods from such areas as statistics, artificial intelligence, pattern matching, information theory, cognitive psychology, linear programming (optimisation), mathematical and numerical modelling, etc. The possibilities seem to be unlimited. However, the goals of music analysis remain the same: we hope to gain new insights into music, to understand 'more fully how it makes its effect' [3], to hear music differently and, as a performer, to make performance more effective and meaningful [4]. We are still interested in musical form, internal coherence in a composition, compositional details, tonalities in music and so on. But, as a spokesman for the objectivist paradigm [4], we, unfortunately, also fall into the same trap – mentioned by Nicholas Cook – of turning 'into a quasi-scientific discipline in its own right' by formulating 'sophisticated analytical methods more or less as an end in itself' [5].

Since the 1950s, countless numbers of different kinds of stand-alone computer applications have been developed for music analysis [6], hardly any of them survived for long. At the same time, such *programming environments* as Humdrum, Matlab and OpenMusic have established their position in the research field. This reflects, for its part, the nature and needs of analytical research in music: certain musicological hypotheses can be encountered only with tools flexible enough to deal with them.

1.2 Manual vs. computational approach

In principle, there is not a big difference between manual and computer-assisted analysis, since it is not the computer that formulates research questions but the analyst her- or himself. The careful study of music is in both cases an important part of the analytical process. Equally, both approaches may include pre-processing stages such as segmentation and, in the computer-aided approach, the conversion of a score into numbers or some other symbolic form. They both also often involve an iterative process requiring a series of trial and error assessments, a back and forth movement between the changing of input parameters (or presuppositions) and the evaluation of results. Henceforth, we are talking about analysis that focuses on the *symbolic representation* of music instead of sampled audio, since the objects of the present study belong to the former category. However, the conclusions in this section may fit into an analysis of audio signals as well. One concrete difference between the two representations lies in the number of samples; audio data requires a huge number of samples compared with symbolic data.

In order to illustrate the differences between the manual and the automatic method and the raw analytical power of the computer, we consider the often-cited [7, 8] motivic analyses of the first movement of Brahms *String Quartet op. 51 nr. 1* by Forte (1983) and Huron (2001) [9, 10]. In his article, Forte emphasised the importance of motivic constructions in Brahms' music as the determinants of 'musical gestures at levels beyond the scale of the foreground'. For him, the motif was primarily an intervallic event, distinct from any particular pitch manifestation. He considered the scale-like 'α-motif' {C,D,Eb} along with some of its derivatives as the most prominent motif in the movement. Two decades later, David Huron reassessed Forte's study and the *significance* of its interval features. By taking advantage of the computer, Huron found that the α-pattern is no more common in op. 51 nr. 1 than in some other comparable quartet movements by Brahms. He concluded that it would be inappropriate to depict an interval feature independent of its rhythmic properties and determined that the pitch-interval feature of Forte's α-motif is inextricably linked to a long-short-long rhythm.

In fact, Forte considered the rhythmic features of α-motif [9, p.483], but what he could not do and what Huron could do was investigate the significance of different interval/rhythm combinations. Such an approach requires extensive calculations, which are in practice unattainable using a manual computational approach. On the other hand, Huron's computer implementation was unable to examine large-scale background features or whether the α-motif is truly the best candidate for the musical features of the movement. Such a consideration may remain unattainable when it comes to the computational approaches. Unfortunately, research results often seem to be determined by practical possibilities, what can be done and what cannot.

We might say that the more complex our data and theory is, the more we are dependent on computational resources. However, even the most outstanding analytical methods do not ensure that statistically significant results or results exhibiting formal economy or remarkable formal relationships between items – such as Huron was able to produce – are in accordance with our personal (and often inconsistent) perception or cognition. As Willard Rhodes stated at the dawn of computer-assisted music analysis for those who fear the 'dehumanisation' of analytical studies, 'there is no substitute for the aesthetic, intuitive and musical judgement of the analyst' [13].

1.3 Semi-automatic and automatic analysis

Music analysis is often regarded as a part of *systematic musicology,* which aims at inducing laws and regularities in its subdisciplines [25]. Computer-assisted music analysis, broadly understood, belongs by its very nature to the subcategory of 'formal music analysis', which – as stated by Nicholas Cook – means 'any kind of analysis that involves coding music into symbols and deducing the musical structure from the pattern these symbols make' (it should be noted that Cook makes here a distinction between 'formal analysis' and 'analysis of form') [5]. It originates from the theory of serial music first developed by Milton Babbitt and George Perle in the 1950s and 1960s [5]. From the 1960s onwards, musical *set theory* in particular offered a fertile ground for computer-aided analysis, its concepts consisting of measurable algebraic units. Computers had already been employed in music analysis for many years. Possibly, the first documented study was conducted by ethnomusicologist Bertrand H. Bronson. He classified and ordered British-American folk tunes as early as the year 1949 by using a digital IBM punch-card computer. The system was far from being automatic. The pre-processing phase consisted of coding some elements of folk-tunes, such as range, time signature, phrasal scheme, etc. with the cards. Although the process of encoding was time-consuming, the computer served as a labour-saving device by enabling him to work with masses of material numbering thousands of items [12].

Among the first pioneers who applied set theory to computer-assisted analysis was Richard Teitelbaum, who used the computer to unmask motivic relationships in the atonal music of Webern and Schönberg[11]. Similar to other pioneers, he developed new methods of his own. His analytical tools included the first geometric 'similarity function' for measuring the similarity between what later came to be called set-classes (SC). In addition, his study consisted of statistical considerations, another common approach to computational music analysis. However, the degree of automation was still quite low. To consider relations between pitch-class sets (pc-set), they had first to be segmented from a large mass of input data and, thereafter, entered into the computer by hand. As with most other computational approaches, this kind of approach can be called *semi-automatic*. A partially manual processing of the data and a back-and-forth motion between different parameter settings are highly characteristic of the analysis. However, such acts are very useful since they increase our understanding of the music: the process itself may be as important as the results.

Several stand-alone applications have been written that are meant to assist in manually-oriented music analysis; for example, for finding the set-class designation that corresponds to a list of pitch-class names or measuring the similarity between two set-classes. As noted by Anthony Pople [3], such programs function like a pocket calculator; 'they can save time, but they don't engage directly with the business of analysis, and so the empirical engagement with the score remains entirely the responsibility of the analyst.' Pople himself developed two small applications for music analysis in the 1990s, RowBrowser and SetBrowser (for Macintosh System 7). The former program allows us to generate different transformations of a twelve-tone row and the latter to build a database of pitch-class sets in order to see some relationships that exist between them.

The results are not either reproducible in semi-automatic methods, if the procedures in the manual part of the analysis are not defined explicitly. This is not the case in statistical analysis when its calculations depend on, for example, knowing exactly the note occurrences or unambiguously defining the segmentation method. Such an analysis is easy to execute *fully automatically*. For example, many key-finding algorithms are based on mechanical segmentation, in which the pitch classes for the algorithm are collected from a durationally-restricted area, such as bars or within a small window that runs across the length of the music. (e.g. [14, 15]).

The statistical analyses of large data sets are, for sure, 'objective'; they can be submitted to statistical tests, producing suitable information for particular questions. However, music happens over time, forming a dynamic system. Consequently, music analysis is commonly concerned with 'describing the structures in a piece of music and discovering how they relate to one another and change through time' [16]. For that purpose, we need three analytical acts: the act of i) segmentation (or reduction), the act of ii) comparison, and the act of iii) representation. The segmentation of music refers to producing analytical objects; for example, by limiting notes to pitch-class sets or inter-onset intervals to 'rhythmic sets'. Comparison, which is common in all types of music analysis, may consist of a test for identity or a measurement of similarity between two segments. The final aim of music theory and analysis is to describe musical structure and to communicate with the addressee. If we are able to solve the most fundamental issue – that of segmentation – successfully, we can easily produce a general and fully automatic 'analytical apparatus' consisting of the input data, the analytical device itself and the output data. Such an apparatus is comparable to a mathematical model defined by variables, a series of equations and adjustable parameters aimed at characterising the process being investigated. As a consequence, if the set of parameters are kept the same, the result should always be the same when entering the same input data.

1.4 Sketching an ideal analytical model

There is a wonderful machine mentioned in the Finnish national epic, the *Kalevala*, which brought good fortune, joy and happiness to its holder and which was capable of grinding not only grain and salt but also gold and whatever else was demanded of it. A skilful blacksmith named Ilmarinen built this magical artefact, which was called the Sampo. In the field of music analysis, we would be equally happy to have such a fabulous Sampo that could answer all of our research questions. In the future, there certainly will arise skilled blacksmiths who program more and more marvellous applications for our music analytical needs. However, for now, we have to content ourselves with merely sketching such a machine. We start by presenting some desirable requirements for it to fulfil and some capabilities we would like it to offer at a very general level. Thereafter, we consider the concrete prerequisites for these offerings. Our model should be as general as possible, independent of the musical styles, remembering at the same time that there are several ways to compose music and, therefore, there must be several methods available for analysis. We cannot assume such a tool exists in real life, since 'one size does not fit all'.

We might find an indefinite number of fascinating properties for our tool. At this point, we mention only some of the most applicable properties, which we do not rank according to their order of importance.

1. Our analytical tool allows for the repeatability of measurements.
2. It is suitable for studying different musical dimensions.
3. All parameters are likewise adjustable.
4. It allows for statistical examinations.

5. It is able to classify musical objects.
6. Music can be examined as a dynamic system as well.
7. Results can be represented illustratively.
8. Results and methods are combinable with other approaches.
9. The validity of the results is verifiable.
10. It reveals something about the rules of composition.
11. It reveals middleground connections as well as surface features.
12. Results correspond to our perception.

Repeatability is a commonly invoked criterion in empirical musicology and in scientific research in general for estimating the uncertainty of analysis and confirming the results [17]. However, experiments are seldom repeated in music analysis and, therefore, musical analyses typically stand alone: as analysts, we are not only interested in 'true' or 'correct' answers to our questions, but also how we and others hear or 'read' music as individuals. In every case, in order to duplicate the results unambiguously, all analytical procedures have to be automatic or – at least, when repeating the analysis – all prerequisites must be known.

Most music analytical methods are originally intended for investigating a particular musical dimension, such as rhythm, tonality, melody, voice-leading, and so on. However, the same methods can be applied to different dimensions if the analytical objects themselves are suitable for them. This is possible if the objects are presented as *vectors*, which is typically the case in *pattern-matching* techniques. As an example, we can measure the dissimilarity between two vectors by using the Euclidean distance, regardless of whether the vectors are derived from rhythmic or pitch information. The third property refers to adjustable parameters that may have an effect on the result. Changing one factor at a time reveals the effect that a particular factor has on the output and how sensitive the model is to parameter changes. This may be useful in corroborating the results.

Statistical and classifying approaches to music analysis are innumerable and it is axiomatic that they are a part of our approach. The latter approach is especially common in music information retrieval (MIR), which aims at organising large music collections and providing semantic structures to access as well as comparing arbitrary musical works.

Computational approaches which observe musical structures and which are more complex than simple entities, like pitches or note durations, and which, at the same time, see music as a dynamic system are rare. This is because an automatic segmentation method is required, which forms, for its part, a common bottleneck in music analysis. Although MIR is a very active research field, researchers have mostly focused on low-level features of digital audio data and, at the higher level, instead of on harmonic segmentation, on melodic segmentation [15]. This is partly because the query-based retrieval (e.g. 'by humming') belongs to the most desired application in the field. There are multiple solutions for melodic segmentation, each of which may focus on specific issues within the research project in question. However, Widmer et al state – referring to audio data – that 'precise, correct, and general solutions to problems like automatic rhythm identification or harmonic structure analysis are not to be expected in the near future – the problems are simply too hard [...]' [18].

Since the nature of music is dynamic (even though not all music is dynamic in the dramaturgical sense), automated segmentation is of considerable value. Through segmentation and comparative methods, we are able to study tendencies and inter-associations in music, whether the segments are monophonic or polyphonic in nature. By segmenting music into consecutive feature vectors, at least three comparative approaches become available: 1) *dynamic* (consecutive objects in time are compared), 2) an approach based on *self-similarities* (all objects are compared with all other objects), and 3) an approach in which all objects are compared with one permanent comparison object (see Figure 1). The latter method is here called *Comparison Structure Analysis* (CSA), which we present in Section 2.

For illustrating the musical dynamics, a graphic representation is helpful. The graph of music analysis is not only a result, but it diminishes the gap between the technical knowledge required and a person who tries to interpret the analysis. Dynamic illustrations can also be compared with other dynamic analyses of the same musical object.

Our eighth property, combinability and comparability, widens the usefulness of our magic analytical machine. For example, by combining the second and sixth property, a huge number of new viewpoints appear within our reach. We are not only able to compare our new results with earlier analyses of the same composition but, perhaps, with the results of human perception.

As our comparative example in Section 1.1 showed, investigating the significance of an analysis done by hand may be troublesome or even impossible. On the other hand, there are many tools for examining the *robustness* of our model predictions when the computer is used. As stated earlier, by changing the parameter settings we can study how sensitive our model is to them. We consider the sensitivity analysis and statistical significance in Section 2.4 and demonstrate corresponding analyses in Appendices A.1 and A.2.

The last three properties of our ideal analytical apparatus undoubtedly belong to the most challenging among others. In order to reveal the 'rules' of composition, the analytical viewpoint has to be re-oriented from observing the surface to detecting the inner dynamics of the music. As example of such an analytical application, David Cope's computer program *Gradus* is able to 'learn' a series of compositional rules of Fuxian first-species counterpoint by analysing a number of two-part models [19]. What is especially interesting here is that some of the rules that the program found are new and not mentioned earlier in the counterpoint guidebooks. At the same time, *Gradus* learns how to compose two-voice, one-against-one counterpoints. An even more challenging type of analysis, proposed in the eleventh item, is to find middle-ground connections between musical events. Unfortunately, that may be difficult for human perception as well.

A cognitive scientist might think that the last-mentioned property should instead be first on the list and, as we know, there are theories that are meant to model human musical perception or cognition, such as Lerdahl and Jackendoff's *Generative Theory of Tonal Music* (GTTM) [22]. However, the twelfth condition is certainly not imperative. As Erkki Huovinen states, ‘a rule that produces music-structural claims may be interpreted purely structurally, without physical or psychological commitments [. . .] it should be clear nowadays that the application of a compositional strategy may not always have implications for musical hearing’ [23]. To give an example, a skilful listener may recognise an invertible counterpoint in Bach’s music or even a retrograde form of a theme, but at the same time that particular skill is not necessary for the listener to enjoy the music. In every case, Bach’s music includes different kinds of contrapuntal techniques, which *are* there in music (given a formal definition of these techniques) and which can be found by using analytical techniques.

After discussing the properties of our ideal analysis model, we conclude that some of the properties are more dependent on the prerequisites than others and, at the same time, some of them open up more possibilities in new directions (considering that we are talking about the quantity of possibilities, not about their – more or less indefinable – quality). However, we have to remember that in order to attain something, we normally have to forfeit something else. For example, to examine music as a dynamic system automatic segmentation is mandatory. This may produce, for its part, segments (artefacts) that are of no consequence with respect to ‘actual musical structure’. On the other hand, computational segmentation not only allows us to repeat and confirm the analysis, but also to classify the segments, to compare them and to consider the distribution of the classified segments.

If the automated segmentation produces segments that are assigned feature vectors through the chosen representation, it is quite easy to build a multidimensional system that allows us to combine and compare several different musical dimensions. This leads to a multitude of new approaches. The problem is that there are, in theory, a combinatorial number of possible segmentations. As several scholars have stated, it seems virtually impossible to systematise the segmentation procedure in any useful way. For example, Christopher Hasty has convincingly shown that multiple segmentations of the same material may be meaningfully developed from different perspectives [24]. As there is no ideal segmentation procedure, and because we cannot define the borders of the segments unambiguously, we have to be content with the second best solution: a segmentation method that is somehow even-handed in each aspect of the composition. In Section 2.2, we discuss possible solutions to automated segmentations.

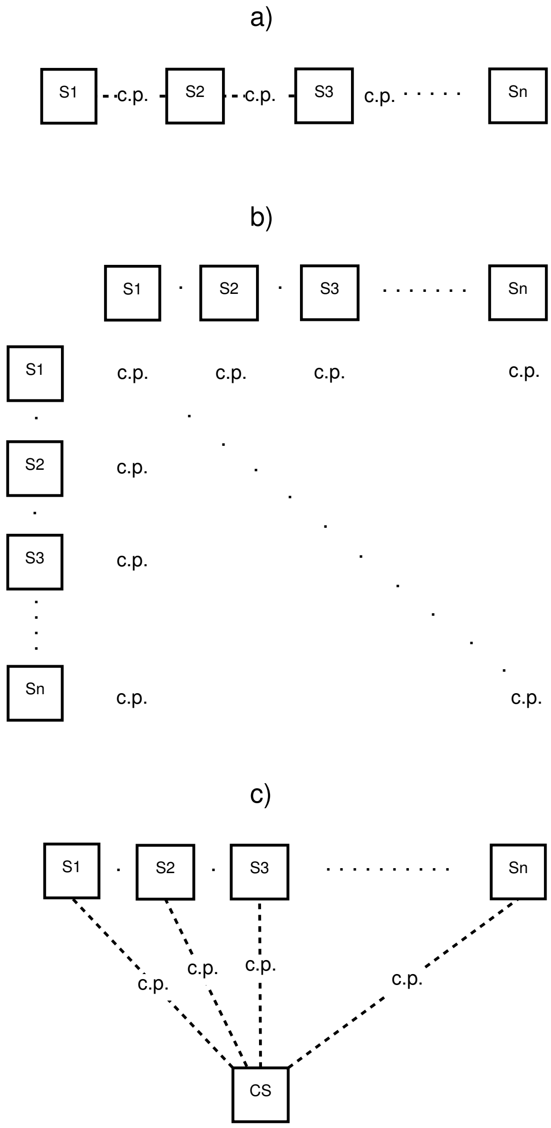


Fig. 1: Three different comparative approaches between musical structures: a) *dynamic*, b) *self-similarity matrix*, and c) *Comparison Structure Analysis*. All approaches are dependent on a particular comparison procedure (c.p.), and an automatic segmentation method that creates segments $(s_1, s_2, s_3, ..., s_n)$. Comparison Structure Analysis utilises the 'outer' *comparison structure*.

2 Comparison Structure Analysis

In this section, I formalise a flexible analytical method called *Comparison Structure Analysis* (CSA). The term *Comparison Set Analysis* is used instead in articles [I,II,VI,V]. Since CSA is dependent on the initial values of several parameters, it is *not fully* automatic before all the parameters have been set. It offers the features 1–9 and perhaps nr. 12 of our ideal analytical machine depicted in Section 1.3. The aim of CSA is first and foremost practical: to produce dynamic and understandable representations of musical properties concerning the composition under analysis. It resembles 'traditional' music analytical methods by examining the musical piece as a whole and by utilising the concepts of music analysis. Most music analytical methods require the pre-processing of musical data before the actual analytical procedures, and CSA does not make an exception. In fact, it is dependent on the 'fixed-length segmentation'. The most salient aspect of CSA is that it uses a 'sensing' device or *detector* consisting of a similarity measure and a *comparison structure* embodying the particular musical properties of the detector. The comparison structure may refer to a comparison set, set-class, pitch-class set, string, or even to a combination of different data structures. The comparison structure closely resembles the concept of *query* in text and music information retrieval. Similarity and distance functions have been used in music analysis for evaluating similarities between set classes, melodies, rhythms and 12-tone rows (see e.g. [11, 21, 34]). What is peculiar to CSA is the usage of a constant comparison structure. I will discuss this topic later.

There are several possible ways to perform a valid segmentation and many alternative similarity measures are available. However, it is good to remember that asking a good research question and finding a working application is more important than selecting the 'right' alternatives for aforementioned tasks.

2.1 Formal definition of Comparison Structure Analysis

Basically, CSA consists of two main functions, one that includes the segmentation procedures and another that consists of the comparison procedures. The segmentation function takes the encoded musical data, e.g. note-onset times, and the selected segmentation *cardinality* as its input. It returns *feature vectors* or other types of data structures, henceforth called segments. These are thereafter entered into the comparison function, which, for its part, compares the segments with the selected comparison structure, producing a sequence of similarity values. Since the comparison structure can normally be chosen from a huge number of alternatives, its selection, which may be based on statistical, mathematical or intuitive considerations, forms a crucial part of the method. Depending on the musical dimension under investigation, the actual input data may vary. In practice, the input data consists of elementary musical units; for example, a sequence of pitch classes or relative durations derived from the musical score or performance that has first been encoded to MIDI file or another symbolic representation of music. The constant cardinality of the segments guarantees the most unbiased performance in the comparison calculations. The segmentation cardinality is often a consequence of the length of the selected comparison structure, since most similarity or dissimilarity functions are meant to be used with vectors of equal length, i.e. they belong to a space of equal dimension. CSA can be formalised mathematically in the following way.

Definition 1. Let C be a set of musical scores and c ($\in C$) a score subject for analysis. We consider a set $S \subset \mathbb{N} = \{1,2,3,...\}$ of numbers called segmentation cardinalities. In practice, the elements of S are usually relatively small numbers. Let $n \in S$ be selected. A constant comparison structure cs ($\in CS$) of length $l = |cs|$ is a vector ($CS \subset \mathbb{R}^l$) or a non-empty finite set ($CS \subset \mathbb{R}$). Often $n = l$. Let $segm : C \times S \to \mathscr{M}(\mathbb{R})$ be a function that performs the segmentation by producing feature matrix $M = (m_1, m_2, ..., m_k)^T (\subset \mathscr{M}(\mathbb{R}))$, where rows $m_1, m_2, ...m_k$ are feature vectors ($\forall i : m_i \in \mathbb{R}^n$) or unordered sets ($\forall i : m_i \in \mathbb{R}$). The number of resulting segments k is defined by the segmentation procedure. Let $comp : \mathscr{M}(\mathbb{R}) \times CS \to \mathbb{R}^k$ be a function that performs the similarity calculation between M and cs. We define *Comparison Structure Analysis* method as a function $csa : C \times S \times CS \to \mathbb{R}^k$, $csa(c,n,cs) = comp(segm(c,n),cs)$.

The definition is represented here in its simplest and strictest form. The input may naturally be only a part of the score, which takes the place of the whole score. The resulting segments, for their part, are not always actual n-dimensional vectors or ordered sets but may represent unordered sets, matrices or even the combinations of different data structures. They do not necessarily represent the elements of the mathematical *vector space.*[1] By varying the segmentation cardinality n or the functions *segm* or *comp*, we can further examine how sensitive the analysis is to these parameter changes. This is important for verifying the results. For example, it is easy and justifiable to impugn the mechanical segmentation in the analysis, but in the case that different segmentation methods produce similar results, we can be sure that the segmentation method itself does not have a significant effect on the results.

It may be noted that several other useful applications are closely connected to CSA, but here they are not called CSA (see Figure 1). For example, if the feature vectors are classifiable and named, statistical analysis with the 'set-class' occurrences – without continuing to the comparison procedure itself – is possible [II, p.214].

CSA can be applied to polyphonic as well as homophonic music and its main procedures are, in principle, independent of musical styles. In this study, CSA is applied to music in which the common factor is that all the pieces are based on 12 pitch classes and encoded to MIDI files. In fact, the basic idea of comparison structure analysis can be used with any time-series type of data. However, CSA is in the present study applied only to a discrete type of data. In order to apply CSA to a continuous type of data, for example glissandos in music, more advanced tools or a huge increase in the number of 'samples' (i.e. cardinality) is required as far as the segmentation and comparison structures are concerned.

In CSA, we observe the prevalence of some musical property according to the consecutive time points of music. Such an approach offers a kind of overall picture for the music. The problem with CSA is, however, that we do not see the larger connections within the piece just by measuring the

[1] This is not a requirement for a successful CSA. A vector space is a mathematical structure formed by a collection of vectors that may be added together and multiplied by scalars. Operations of vector addition and scalar multiplication have to satisfy certain requirements, called axioms, which are not listed here.

similarity between actual musical objects. For example, the form of a musical piece can be based on a cantus firmus, but we cannot detect that cantus firmus melody by using CSA. In this sense, CSA offers merely a statistical viewpoint for music analysis. The strength of CSA relies especially on the possibility of making comparisons between the observations concerning different musical parameters and combining them with other analytical methods.

2.2 Segmentation procedures

Comparison structure analysis is based decisively on the segments of the same cardinality produced by the automatic segmentation. Unlike manual approaches, in which the pre-processing phase may depend on the intuitive clustering of notes or on rules that produce hierarchical segmentation (like in GTTM [22]), CSA depends on an abstract systematic segmentation that allows for a statistical or mathematical survey of the data. If we can assume that manual methods are based on 'musical thinking', CSA is then based on the algorithmic 'projection' of the musical material. The practical solutions for segmentation in CSA vary depending on the musical dimension under investigation. These have been discussed in [I,III]. In practice, we do not assume a single ideal solution for all purposes, but instead consider several possible methods depending on the analytical aim and application. Similar procedures can be found in the fields of text and music information retrieval as well as in bioinformatics (DNA sequence analysis).

Implementing a satisfactory segmentation solution for the computer has proven to be difficult, if not impossible. In his *General Computational Theory of Musical Structure (GCTMS)*, Cambouropoulos aims at reaching a structural description of a musical surface, independent of any specific musical style or idiom [29]. He suggests that a powerful theory of music for building a computer application should address the following points: 1) explication, 2) generality and 3) induction. The first point assumes that the theory describes musical activities and tasks in an explicit and consistent manner. The second one refers to its generality in stylistic issues and the third to the theory's capacity for making generalisations by analysing existing musical works. All these points assume a fully automated method of analysis. However, when discussing Nattiez' *paradigmatic analysis,* Cambouropoulos states that the relations between 'segmentation, similarity and categorisation are quite complex especially when it comes down to developing a computational model'. A lot of energy has been spent finding workable solutions to pattern processing algorithms that aim at producing musically meaningful patterns (see, for example, [30]). Nevertheless, this is just the first phase in music analysis and actual analytical computer programs that consider the composition as a whole are rare. In the event that we are able to develop a general and comprehensive segmentation method that produces musically 'meaningful' units, it is not recommended that they be used with CSA, which produces the most accurate results with the constant segmentation cardinality.

At first glance, the fixed-length segmentation may look alien to music analysis. However, it has proven to work well in certain tasks. Neve and Orio present an overview of different approaches for extracting melodic units to match relevant documents with queries, from fixed-length and data-driven segmentations to perceptually- or theoretically-based segmentations [27]. They carried out the comparisons using the Cranfield model for 2310 MIDI files (for more on the general principles of the Cranfield model see, for example, [28]). According to their experiments, simple approaches to segmentation with overlapping units gave better performances than approaches based on music perception or music theory. A simple approach, one that does not filter out any information, improves recall without degrading precision. However, since the aim of the application is different than in the present study, such an experiment may only give some overall support to the hypothesis considering the different segmentations.

CSA can be employed for treating several musical dimensions at the same time, which may especially be effective in classification tasks: as the dimension of the feature space grows, the search area becomes more selective. The problem is that each musical dimension may need a segmentation procedure of its own. In order to facilitate the comparisons between the measurements of different musical dimensions, the segments and the similarity values can be *anchored to onsets* and, thereafter, by averaging the values, to bars. This is just a practical solution that, unfortunately, loses part of its information and one that does not remove the fact that the segmented borders of the different dimensions differ from one another. The sensitivity analysis presented in Appendix A.1 provides a more focused basis for evaluating the competency of the fixed-length segmentation used in CSA.

Let us briefly consider the segmentation of harmonic and rhythmic dimensions in a composition. A musical rhythm can be derived, for example, from a particular melodic line or from each unique note onset of the whole score as a sequence of IOIs or their IOI-ratios. Such a sequence can be thereafter easily partitioned into overlapping rhythmic strings of a particular cardinality (see Figure 2). These segments can then be evaluated against a comparison string by applying a *substring distance* or other similarity measure [III]. Other types of feasible approaches for choosing the rhythmic onsets are certainly available. For example, the selection of onsets might be based on ‘weights’ by calculating the number of vertical notes or taking the durational situation of the onset into account.

The segmentation for harmonic analysis requires a more complicated treatment. One possible solution, which produces pitch-class set clusters – fulfilling a particular cardinality – for each onset time, may comply with the following rules presented in article [III]:

1. For each onset time, aim at defining an associated segment (or segments) of the selected cardinality.
2. Vertical groupings are preferable.
3. If the number of pitch classes at the onset exceeds the selected cardinality, assign all possible pitch-class combinations to the onset.
4. It is preferable to add missing pitch classes from the onsets that are closest temporally.
5. If there are alternatives before and after the onset that are equally viable, it is preferable to search for them systematically either from one or the other direction, but not from both directions, in order to diminish evolving ‘noise’ from the resulting data.
6. To be even-handed, several alternatives of equal value may have to be assigned to the same onset.

Such a ‘clustering segmentation’ has also been illustrated in Figure 2, in which the pitch classes have been clustered with the tetrachordal pitch-class sets. For this example, I preferred searching for the missing pitch classes first before an onset (see the fifth rule). In order to illustrate the effect of the third rule, let us consider the case in which the vertical harmony consists of five pitch classes and the segmentation cardinality has been defined as 4. Since the number of combinations is the binomial coefficient $\binom{n}{k} = \frac{n!}{k!(n-k)!}$, where n is the number of objects from which you can choose and k is a number to be chosen, we get 5 different pitch-class sets for that particular note onset time. Following the sixth rule, we have to define, on the whole, four segments for the fourth onset. This may seem arbitrary but, when looking more closely at the segments, we notice that they are almost the same and consist of only five pitch classes in total. Huovinen and Tenkanen present another segmentation method for the set-class CSA termed ‘tail segmentation’ [I]. In Appendix A.1, I compare the results produced by these two different segmentation methods, which both generate overlapping segments.

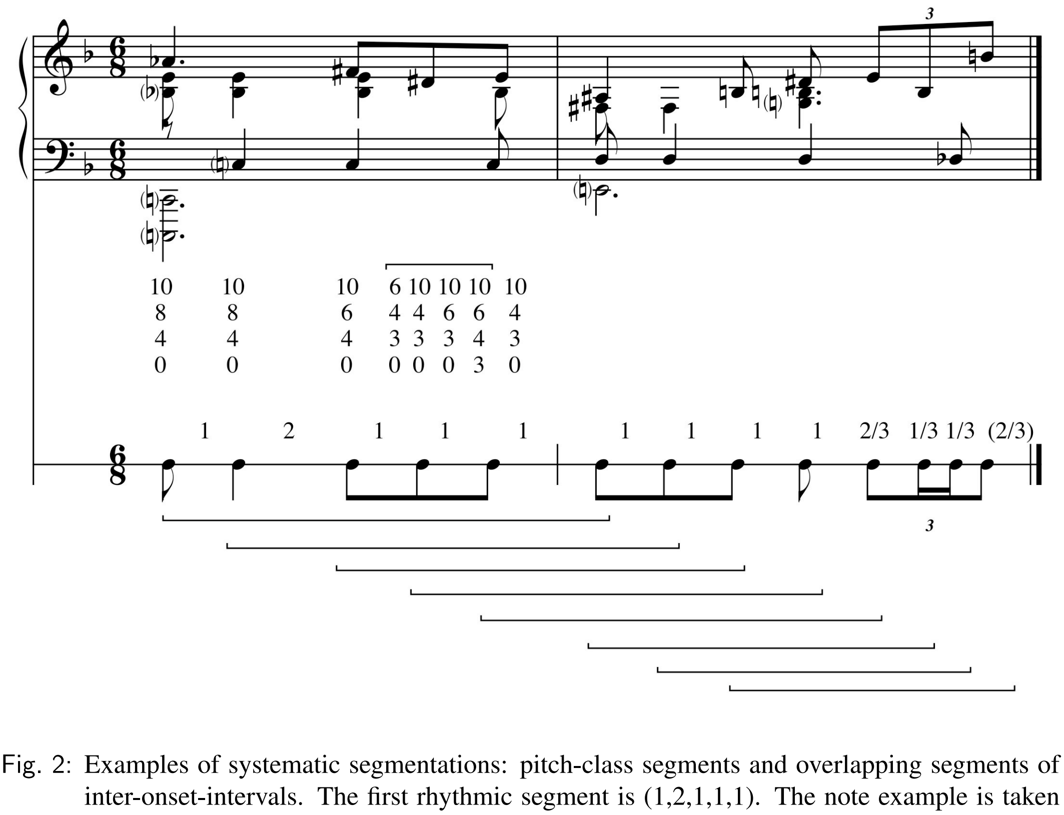


Fig. 2: Examples of systematic segmentations: pitch-class segments and overlapping segments of inter-onset-intervals. The first rhythmic segment is (1,2,1,1,1). The note example is taken from Alban Berg’s *Wozzeck, Intermezzo.*

The generated segments themselves enable us to make statistical observations about their types or classes and, further, about the distribution of these classes. In the right situation, this may produce useful results [II]. Labelling set classes may be useful if the aim of the analysis is to find particular set-class occurrences. However, this does not tell us anything about the relations between the segments. CSA offers a better means for understanding such relationships.

There is one other justification for an approach based on similarity measurements: even if we might not be able to name (and classify) the resulting segments – as in the case of rhythm segments – they can be utilised for similarity comparisons and classification using numerical evaluations. To make this clear, there is a finite and passable number of set classes for pitch-class sets: 224 pieces of T_nI-types and 352 pieces of T_n-types (see, for example, [26]). However, this is not the case for many other types of musical objects. For example, the number of possible 'rhythm-classes' is infinite, growing in combination with the length of the rhythms. In order to illustrate the problem, let us sketch a simple method for classifying rhythmic strings, which is based on the order of the length of the inter-onset intervals in each rhythmic 'string', i.e. the rhythmic information is reduced to a minimum. In this system there are three classes for rhythmic strings of length 2: 11, 12 and 21, since 11==22 and 13 classes of three-element rhythms: {111}, {112, 121, 211}, {122, 212, 221} [111==222==333], {123, 132, 213, 231,312, 321} [322==211][323==212==313][331==221]. It follows that, for example, dactylic rhythm is always marked as 211. It can be shown that there are 75 rhythmic tetra-classes, 541 penta-classes and 4683 hexa-classes, etc. The number of classes in the different groups follows the so-called 'ordered Bell numbers'. Classifying rhythms in a composition might be sensible if the segmentation cardinality is kept lower than or equal to 5.

2.3 Similarity measurements and the comparison structure

Similarity is often defined as a partial identity; that is, two entities are similar if they share some, but not necessarily all, properties [31]. According to Cambouropoulus, similarity depends on context and is, by definition, relational. Therefore, it cannot be objective since 'context affects perception of similarity, categorisation and, even, the perception of entities per se'. CSA is based on mathematical similarities between the musical objects within the composition. CSA provides numerical results according to which an analyst makes her or his conclusions. In CSA, the similarity is handled only with relation to the musical material itself and, thus, we do not see it as indispensable to consider the issue of perceptual similarity in relation to CSA.

As stated previously, the fixed-length segmentation is inextricably linked with similarity measurements in CSA and does not focus on compositional details or the occurrences of a particular type of features but, instead, on the 'fuzzy' prevalence of a certain abstract musical structure in a composition: The values given by a similarity measure are normally scaled between 0 and 1. The segmentation method produces measurable units, e.g. vectors or unordered sets which are handled like vectors in similarity calculations. CSA may reveal something about the shape of the overall form of the piece in relation to the comparison structure that works as a constant sensing detector in analysis (see Figure 6 in [II] and Figure 7a in Appendix A.1). Instead of searching for the equivalent occurrences of a particular musical object, we are interested in the degree of similarity between the segments and the comparison structure. For that purpose, a similarity measure is applied. Similarity measures are used in many research areas, since defining the similarity of objects is crucial in data analysis and decision-making process. They are employed, for example, to separate relevant and irrelevant data items and to classify data objects. A similarity measure is a function, $sim(\bar{x}_1, \bar{x}_2)$, that takes two patterns (often vectors) as arguments and yields a single value (similarity value). Two patterns are said to be similar according to the similarity measure if the similarity value is high or dissimilar if the similarity value is low.

Many similarity measures are based on distance calculation between the entities, i.e. they are in fact dissimilarity measures. In this case, the more similar two objects are the closer they are and, correspondingly, the more dissimilar two objects are the greater their distance is from one another. Typically a *distance metric* is required when using a dissimilarity measure. A distance metric is a well-defined concept which embodies some essential properties of what we would consider a 'well-behaving' set of values describing the similarity of objects. Distances define a metric if they satisfy the next four requirements: 1) non-identical points must have a positive distance (non-negativity), 2) the distance from x to y must equal the distance from y to x (symmetry), 3) the distance between points x and y cannot exceed the sum of the distance from x to z and the distance from z to y (triangle inequality) and, 4) if the distance from one point to another point is zero, it implies that the points are the same (identity). There are distance functions that fulfil the requirements 2–4, but not the first requirement. If the distance between two distinct points would be zero, we talk about a *pseudometric*.[2] Four requirements correspond with our intuition about the relations associated with the closeness of two sets. Chen et al. have recently provided a formal definition for a *similarity metric*. They also show the relationship between a similarity metric and a distance metric and present general solutions to normalise a given similarity metric or distance metric [32]. In practice, it is easy to numerically prove whether a measure is not a metric by searching for a counter-argument, but it may turn out to be more laborious to mathematically prove that it is a metric. This problem relates especially to the condition of triangle inequality.

Ilomäki, who used a metric to analyse the behaviour of similarity measures for twelve-tone rows [34], states that even if a similarity measure could not be defined as a metric, which is the case with many of the functions used in the present study (e.g. *REL* [I], *costotal* [II], *expcos* [V]), it might still be proven to be compositionally or analytically useful. On the other hand, although the metric function ensures well-defined measurements between certain mathematical properties of the objects being compared, we cannot say that such a function models human perception well, which might in fact be, for its part, a desired property of the similarity function. Shmulevich et al. argue that the successful implementation of a music recognition system *must* incorporate perceptual information and error criteria [33]. A part of their tonality algorithm that utilises both rhythm and pitch information applies to empirically derived Krumhansl-Kessler key profiles, which are taken to reflect cognitive pitch-class hierarchies in a tonal context [38].

I consider different types of similarity measures and the results they offer as analytical examples in Appendix A.1, in which both set-class similarities and tonal stabilities are measured. It seems that

[2] A distance metric is formally defined as:

Distance Metric. Given a set X, a real-valued function $d(x,y)$ on the Cartesian product $X \times X$ is a distance metric if for any $x,y,z \in X$, it satisfies the following conditions:

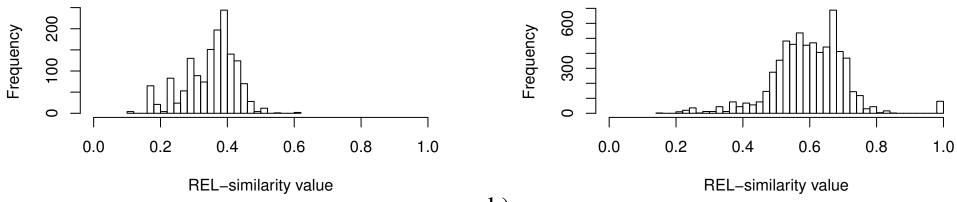
1. $d(x,y) \geq 0$ (non-negativity),
2. $d(x,y) = d(y,x)$ (symmetry),
3. $d(x,z) \leq d(x,y) + d(y,z)$ (triangle inequality),
4. $d(x,y) = 0$ if and only if $x = y$ (identity of indiscernibles).

there are no significant differences between the results produced by different similarity functions in CSA (see also [I:29, III, IV:254]). The characteristic differences of the segment vectors appear to be so robust – independent of the measure – that the resulting curves are often depicted almost parallel to one another. The term ‘similarity function’ is used from this point onwards for all types of measures whether they represent similarity, distance or other kinds of measures.

As stated previously, similarity measures have been used in the field of pitch-class set theory and music information retrieval. There are still some obvious issues concerning the mathematically-defined measures. For example, none of the similarity functions that are meant to measure similarity between set classes take the dissonant quality of the interval classes into consideration, even though they rely on them. However, this might be possible by weighting the interval classes by empirical estimations. In the end, the performance of such a function should be examined by empirical tests. One practical method for finding a function which models human perception is to generate a huge number of new functions by employing *genetic programming* and to compare their performance with the empirical results [V].

Since the main purpose of CSA is to elucidate the overall trends of abstract musical structures and the relations between the different CSAs, and not to approximate human musical cognition empirically, such empirical validity of the similarity measure is not included among the most desirable properties of CSA. However, in some cases CSA may reflect easily observable properties of the piece (without forgetting that music perception is a highly relative concept): some properties are ‘more observable’ than others and, in addition, some listeners are more sensitive to or able to perceive subtler details than others. The connection between the $REL(X, 7-1)$ of *Improvisation nr 756* by Olli Linjama [III] reflects the incremental changes in the prevalence of the highly chromatic set class 7–1[0,1,2,3,4,5,6] within the piece, and the auditory effect of the improvisation might be obvious to everybody independent of his or her musical background.

I already mentioned the preference for the constant cardinality of the segments. Even though there are similarity functions that allow measurings between the sets or vectors of a different cardinality, it is intuitively clear that if the segments are of several different cardinalities we cannot be sure to what extent the similarity values are a consequence of the segment cardinality and to what extent they are a consequence of the different character of the segments being compared. This is because the similarity value range, for the best-known and most useful similarity measures, varies according to the difference between the two cardinalities of the segments being compared (Figure 3). For example, if the maximum similarity value between two *equal* segments is 1, it is obvious that the similarity value is lower than 1 in those cases in which they are of a different cardinality, despite the possible characteristic similarity between the segments. In the case of the *REL*-function ([37], [V]), the maximum value between some set class of cardinality 3 and another set class of cardinality 6 is 0.60 whereas between the equal hexachords or trichords it is always 1. Furthermore, the larger the cardinality difference between the CS and a segment the more probable it is that the similarity value between them will be smaller. Though *REL* takes the total subset-contents of set classes into account, it looses its power of expression little by little when the cardinality difference increases between the set classes. Be keeping the cardinalities between the vectors under comparison the same, all similarity values remain in relation to the same distribution of values. Thus, the probability of such statements as '*REL*(x,y)>0.7' remains constant and the results are easier to interpret.



a) b)

Fig. 3: The *REL*-value distributions between a) all the set classes of cardinality 3 and 6 and b) between all the set classes of cardinality 6 and 6. The value ranges are 0.11-0.60 and 0.14-1.00, respectively.

To give an example, let us consider the harmonic analysis within three set classes, 6–32[0,2,4,5,7,9], 6–1[0,1,2,3,4,5], and 3–9[0,2,7], i.e. the diatonic hexachord, chromatic hexachord and the diatonic (fifth-generated) trichord. The diatonic trichord 3–9 is a subset of the diatonic hexachord 6–32 and both sets can be generated by adding pure fifths consecutively. So they are 'relatives' in their character and structure. When the similarities between the first two sets of the same cardinality and the first and the last set of the same characteristics are compared by using the *REL*-function, the resulting similarity values are almost the same, rel(6–32,6–1)=0.48 and rel(6–32,3–9)=0.47. Let us then consider a CSA in which we employ the diatonic hexachord 6–32 as our comparison set-class and extract our imaginary composition according to the 'musically meaningful' segments of different cardinalities (Figure 4). In the hypothetical case that the composition consists of a section of chromatic clusters 6–1 and another section of diatonic triads 3–9, the CSA forms quite a steady value curve, even though these two sections are obviously different in character. The result would be a little bit better if the latter section would consist of diatonic *tetrachords* 4–23[0,2,7,9] or diatonic *pentachords* 5–35[0,2,5,7,9]; this part would then prove to be more similar to the diatonic comparison set-class, with the resulting similarity values being 0.66 or 0.84, respectively. By using the fixed-length segmentation of cardinality 6, the latter section might consist of diatonic hexachords or set-classes related to it, giving it a similarity value of 1 or close to 1. In that case, the CSA would reveal the different characters of chromatic and diatonic sections immediately (0.48∼1.00). Isaacson's IcVSIM, which calculates the population standard deviation of the interval-class difference vector of two sets, with the value of 0 referring to maximum similarity [35], might produce more acceptable results in this case (c.f. Figure 4), but outputs odd results in some special cases: for example, IcVSIM(3–10,6–30)=0 represents the maximum similarity between set classes because of their similar interval-class vector profile [35, p. 19-22].

At the same time, we also prefer using the same cardinality between the segments and the comparison structure, which allows us to detect exact matches. Increasing the cardinality difference decreases the separating capacity and, if either the comparison structure or segments are one-dimensional, the separating capacity is totally lost.

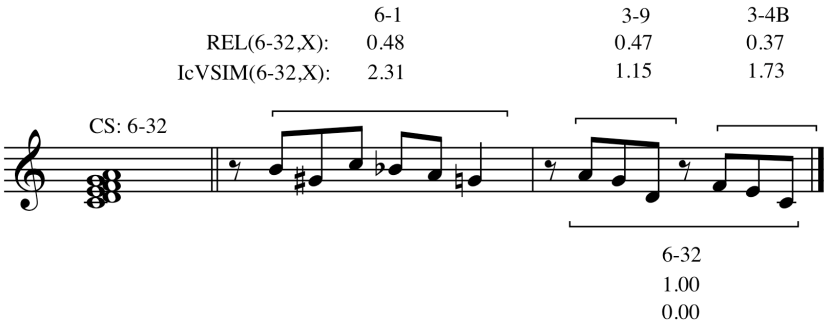


Fig. 4: An imaginary composition that has been segmented into musically meaningful pitch-class sets. These segments have, furthermore, been transformed into set classes that are compared with the diatonic comparison set-class 6-32 using the *REL*- and IcVSIM-functions.

How one selects a suitable comparison structure depends on the aim of the analysis and, perhaps, on technical boundary conditions. Some possible principles for this issue have been presented in article [I] and article [III]. Choosing a comparison structure involves tuning the apparatus to be sensitive to an exclusive set of musical properties. Thus, the selection of the comparison structure is a restrictive act: 'local events are thus examined as if they were being observed through a filter, which enables the observer to see only objects of a certain colour' [I]. The comparison structure may be chosen deterministically; for example, by searching for the most 'significant' segment of the segment space (c.f. 'germinal motive' in [20]). But the comparison structure may be chosen intuitively or perceptually as well, without any mathematical or statistical consideration. We might also use a combination of data structures as the comparison structure; for example, by deriving the properties of different musical dimensions from a particular bar. This kind of procedure would have been available in the case of Sibelius analysis (see [III]), in which the dynamic model (c.f. Figure 1a) was applied with overlapping 5-bar long windows.

The complexity of the calculation varies depending on the application and the aims of the CSA. A short but complicated example of CSA based on the tail segmentation is illustrated in Figure 3 in article [II]. Depending on the purpose of the analysis, it may be applicable to average the CSA values over each notated bar, yielding one value that represents the entire bar. In some cases, for example in an analysis like that of Bach's chorale (c.f. article [I], Figure 3), it might be more meaningful to calculate the means for each phrase and not for the bar. In order to produce reliable bar-by-bar 'curves', we have to take care of the density of segments per bar, thereby legitimising the results. One solution to this issue is to apply the 'cardinality correction', presented in article [I]. In the next subsection, I discuss briefly the reliability of CSA from another viewpoint.

2.4 Testing the reliability and the statistical significance of CSA

Mathematical models of reality are always subject to a variety of reliability issues since they do not permit a straightforward understanding of the relationship between input factors and output results. In order to investigate this path from data to predictions, we have to approach the problem by way of a detour. Usually, this is executed by using *sensitivity analysis*: we aim to determine how 'sensitive' the model is to changes in the value of the parameters of the model and to changes in the structure of the model. This is of fundamental importance for checking the quality of a given model and the robustness and reliability of the results it produces. If a small change in an input parameter results in relatively large changes in the output, the output is said to be sensitive to that parameter. This may mean that the parameter has to be determined very accurately or that the alternative model has to be redesigned for low sensitivity. Sensitivity tests not only help the researcher understand the dynamics of a system by providing information on factors that mostly contribute to the output variability, but they can also indicate which parameter values are reasonable to enter into the model.

Parameter sensitivity is, in practice, performed as a series of tests in which the researcher sets different parameter values to see how a change in the parameter causes a change in the results. A common approach is that of changing one factor at a time (assuming that the factors are independent). In this way, the sensitivity analysis allows the researcher to determine what level of accuracy is necessary for a particular parameter to make the model sufficiently useful and valid. For example, if the tests reveal that our CSA model is *insensitive* with respect to the segmentation method and segmentation cardinality, but not with respect to the comparison structure, we can be sure that the model somehow reflects the effect of the comparison structure.

If some of the factors are interdependent, the situation is more complex. In that case, we may calculate each of the possible results by employing all of the input combinations of the factors. If this is not possible because of the large number of combinations, we can apply a probabilistic method called *Monte Carlo analysis*, in which the input combinations are drawn randomly from the input distributions. The resulting distribution of outputs is then interpreted as an approximation of the desired probability distribution. From this result, the mean or median and the variance can be reported and decisions can be made by means of them. Experimenting with a wide range of values offers insights into the behaviour of a system in extreme situations. In CSA, we normally fix the comparison structure first, which then may determine the segmentation cardinality. It may happen that the only parameters we can vary are the segmentation method and the similarity function.

Articles [I, II, III, VI] offer several illustrative examples and some derivative applications of CSA. However, I have not thus far systematically investigated how sensitive CSA is to the changes in the input parameters and what kind of effect they may have on analytical results. Many segmentation methods are proposed for melodic segmentation (e.g. [27]), but are much rarer for harmonic and polyphonic segmentation: We are especially interested in the effect of changing the segmentation method in set-class and tonality-based CSA. I have considered the sensitivity in connection with CSA in Appendix A.1.

Another important question to consider is whether the results of CSA are *statistically significant*. Calculations in CSA always produce some values and we may ask whether they have occurred by chance. The amount of evidence required to accept that an event is unlikely to have arisen by chance is known as the *significance level*, expressed as a *p-value*. For that, we define the so-called null hypothesis H_0, which might be falsified using a test of observed data. Thus, the p-value is the probability that, if the null hypothesis were true, the sampling variation would produce an estimate that is further away from the hypothesis value than our data estimate. For CSA, we could choose the null hypothesis H_0 'the composer has selected the notes of this piece randomly' (furthermore, we may think that the derived comparison values have occurred by chance) and an alternative hypothesis H_1 'the composer's intentions are controlled by some constraints and rules'. In order to make actual evaluations, we fit the models to the data: we may view the data as our 'glimpse of reality' and the model as the 'theory' to try to explain how that aspect of reality came to be. The notion of 'model' could refer to a probability distribution (e.g., normal, beta) as a model of random behaviour in a population, or to fitting an equation to a random data (a 'statistical model'). The 'goodness' of the fit has to do with both the amount of information (data) available to fit (or 'train') the model and the appropriateness of the model to the data.

The most popular levels of significance are 5% (0.05), 1% (0.01) and 0.1% (0.001). If a test of significance gives a p-value lower than that selected, it can be said that either the null hypothesis is false or an unusual event has occurred. The lower the significance level, the stronger the evidence required. Choosing the level of significance is an arbitrary task, but often a level of 5% is chosen for no better reason than that it is conventional. If the resulting p-value is 0.01, we may say that 'there is only one chance in a hundred this could have happened by coincidence'. For the test, we have to decide on the relevant test statistic, a numerical summary of a comparable set of data (e.g. mean, variance etc.) that can be used to perform a hypothesis test. A statistic is calculated from the sample. For CSA, we may produce the test statistic by producing a set of 'random compositions' by generating the notes of composition randomly, and thereafter, derive the distribution of comparison values from the random compositions. The properties of the random distribution are compared against the properties of the actual composition under study. To begin with, we assume that the hypothesis about the population parameter is true. If the difference between them is small, the null hypothesis is accepted, and if the difference between them is large, it is rejected.

Although we certainly know that the compositions in the present study are not created using random procedures, it is useful to test whether our analytical procedures produce anything other than random results. For that purpose, and as an example of a statistical significance test, a sample analysis is presented in Appendix A.2.

2.5 Variation of applications

In CSA, the average values for bars illustrates the prevalence of features similar to those of a particular comparison structure. The set-class-based analysis of Olli Linjama's *Improvisation nr. 756*, presented in article [III], represents the most illustrative examples of the basic form of CSA. In this analysis, the descending diatonic and ascending chromatic line clearly and understandably illustrates the overall harmonic trends according to the two comparison set-classes 7–35[0,1,3,5,6,8,10] and 7–1[0,1,2,3,4,5,6]. CSA is a harsh but flexible method, which has been applied to distinct musical entities or sets of compositions as well [I, II]. Moreover, we can distinguish at least four different application viewpoints for CSA:

1. The most common case in which CSA is used as such.

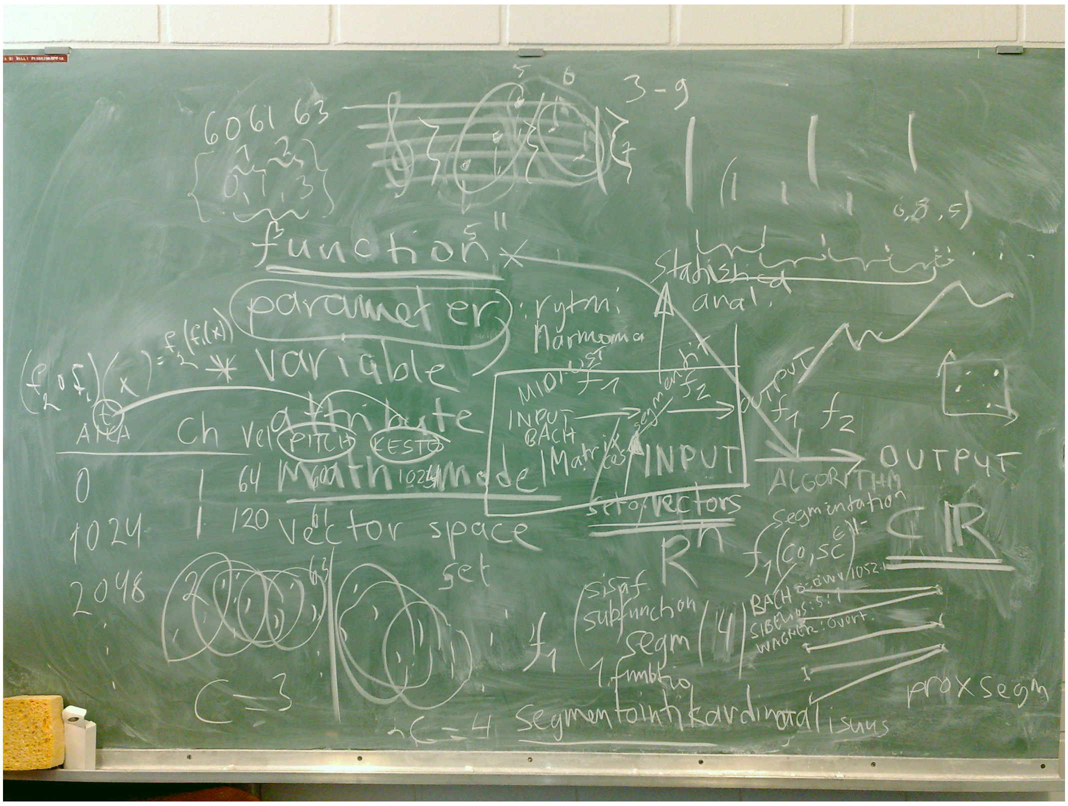
2. Correlation between two CSAs.
3. A statistical test comparing CSA with results produced in other ways.
4. Derivatives of CSA, in which parts of CSA are applied to another type of analysis.

We have presented several different applications of CSA and its derivatives in articles [I, II, III]. Those analyses that utilise CSA as such include examples in similarity measurements between set classes, rhythmic strings, pitch-class sets (tonality induction) and melodic segments (experiments on voice-leading and invertible counterpoint). In addition, papers [I] and [II] present examples of the classification of compositions or composer styles based on CSA. The value distribution of CSA may also offer a viewpoint on the material *coherence* of a piece. In order to obtain reasonable results, relative comparisons for evaluating the material coherence of a particular piece are assumed.

Multidimensional CSA offers a wide variety of correlation examinations between different CSAs of the same composition. I include an example of correlation analysis in the sensitivity analysis (see Appendix A.1. However, in that case, the correlation is not measured between different musical dimensions). Since CSA is represented as a value vector, a comparison between it and another type of numerical analysis is available in the form of statistical tests. The number of derivative applications seems to be numerous. Papers [I, II, III] include examples of music analyses concerning short-term modulations, multiparametric changes in musical texture and studies based on segment distributions. The similarity measures and segmentation methods used in CSA may also be applicable to symbolic music information retrieval. Nothing hinders us from applying comparative methods between two different compositions, a procedure that allows us to find similar sections between compositions. In practice, this kind of alignment produces an $n \times m$-matrix of similarity values: the similar sections can be found in diagonals according to the segment comparisons (c.f. Figure 16 in article [III]). Yet, one possible derivative of CSA is available that has not been considered in this study: in case of a certain musical dimension, say tonality, we might be interested in using a particular sequence of comparison pitch-class sets, instead of using only one set. This would offer an analytical tool to detect, for example, chaconne chains in a composition.

CSA settles somewhere between traditional music analysis and computer aided MIR: theoretically defined musical entities, such as particular rhythm patterns, pitch-class sets and set-classes, are detected in compositions using pattern extraction and pattern comparison algorithms that are typical within the field of MIR. All algorithms in the present study were implemented using R, a programming environment for statistical computing. R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques. R is released under GNU Public License and is available for free for various computer systems. The breadth of R packages continues to grow. At the time of this writing, the number of packages was nearly 2000 and the suggested number of R users 250,000. See R's website *http://www.r-project.org/*.

3 Conclusions



In order to sketch summaries, the blackboard was often used during the research process.

3.1 Summary of the publications

This study consists of five articles, two of which are collaborative works. The first article [I] was written as part of a very intensive colloquial team effort in the summer of 2005 (Huovinen 50% / Tenkanen 50%). Huovinen wrote most of the final text and Tenkanen developed most of the algorithms for the article. We introduce Comparison Structure Analysis, termed as Comparison Set Analysis, in the third section (in fact, the idea of CSA was presented earlier in [36]). The third article [III] was written with Fernando Gualda, who wrote sections 3.4.–3.4.2. The first three articles concentrate on the applications of CSA. In the last two articles, I propose new similarity measures for set-class- and tonality-based CSA, since they have an important function in CSA.

Paper I

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

This article presents a series of methods for analysing the average characteristics of pitch-class material on the surface level of musical works. All the methods rely on 'tail segmentation', the partitioning of a musical work into a large number of overlapping pitch-class sets of equal cardinality. The resulting data can be used as a means of scanning a piece of music in order to detect changes in the local prevalence of some chosen pitch class–related feature. For example, scanning the musical surface for each individual interval class in turn results in a 'stretched interval-class vector' represented by six curves. We apply CSA for investigating both diatonic and chromatic trends (with comparison set-classes 7–35[0,1,3,5,6,8,10] and 7–1[0,1,2,3,4,5,6] in several pieces as well as for classifying several composers according to these two comparison set-classes. We also consider Debussy's prelude *Ce qu'à vu le vent d'ouest* by putting the set-class-based CSA to the test in order to give support to Olli Väisälä's hypothesis considering the normative harmony of the piece.

Paper II

Tracking Features with Comparison Sets in Scriabin's Study op. 65/3.
Tenkanen, A.: *First International Conference, MCM 2007, Berlin. Revised Selected Papers (Communications in Computer and Information Science, Vol. 37)*, Springer, Timour Klouche and Thomas Noll (Eds), 211–219 (2009)

The second article is a case study showing how set-class-based CSA works in a spontaneous situation. Four compositions were proposed for the analysis workshop of the First International Conference, MCM 2007 in Berlin, of which the author chose Scriabin's *Prelude op. 65 nr. 3* (other alternatives included Webern's 'Sehr schnell' from *Variations Op.27*, Xenakis' *Keren* and Schumann's *Träumerei.*). A new similarity measure (*costotal*) for set-class comparisons is introduced. The first step towards multidimensional CSA is taken by applying both rhythmic and set-class-based CSA. The resulting graph (Figure 6) is an illustrative example of the block-type form of the composition revealed by CSA. The number of occurrences of Scriabin's 'Promethean' chord, associated with SC 6–34A [0,1,3,5,7,9], are detected using large-scale analysis with 22 piano pieces. The analysis shows how the 'Promethean' chord-type structures were already prevalent in the composer's works six years before composing the orchestral piece *Prométhée* (1910), which gave the name to the chord. The results matched well with the notions presented by the different analysts showing how Scriabin made a gradual transition from tonality to atonality over years.

Paper III

Multiple Approaches to Comparison Structure Analysis
Tenkanen, A., Gualda, F.: Ed. Timour Klouche. *SIM & Springer: Computational Musicology*, Vol.1. (accepted for publication)

In this article, we examine how CSA works with different musical dimensions. In addition to set-class- and rhythm-based CSA, we apply CSA to melodic analysis. Some derivative methods for classifying compositions, measuring short-term modulations and multidimensional textural changes in music are presented. We introduce a new rhythmic similarity measure, termed *substring distance* to detect the rhythm of the main theme in the first movement of Mozart's *Symphony nr. 40*. The study reveals the often-emphasised contrast between the main and sub-theme groups. Our analysis considering the short-term modulations in C.P.E. Bach's two *Litanies* is in line with the composer's opinion on the harmonic differences between the compositions. It is also worth mentioning a simple but effective method for detecting invertible counterpoints in polyphonic texture. We applied the approach to J.S. Bach's *Invention nr. 5 Eb major*, a short 2-voice piece. By utilising the same method for a more extensive piece or a set of pieces we might obtain new insights into the composer's systematic usage of the invertible counterpoint.

Paper IV

Evaluating Tonal Distances between Pitch-Class Sets and Predicting Their Tonal Centres by Computational Models.
Tenkanen, A.: Mathematics and Computation in Music 2009, New Haven. E. Chew, A. Childs, C-H. Chuan (Eds.). *Springer series on Communications in Computer and Information Science* 38, 245–257 (2009)

CSA is dependent on a viable similarity measure. Our analysis of Debussy's *Ce qu'à vu le vent d'ouest* in paper [I] showed that set-class-based CSA cannot answer our question about the most probable referential set in the piece since set classes do not carry transpositional information. Four algorithmic models presented in this paper will respond to that issue. Each of the models, except for the *cofrel*-function, are developed partly using the existing models. The same models can be used either for finding the hypothetical tonal centres of pitch-class sets or for measuring 'tonal stability' between pitch-class sets. Both approaches can be applied in CSA. In the first case, CSA would predict the tonal centres, with the results being discrete values between 0 and 11 (from C to B). The models are applied as 'tonal stability' measures to a piece by Alban Berg from the opera Wozzeck. *Cofrel* is purely a mathematical model, developed by the author.

Paper V

Searching for Better Measures: Generating Similarity Functions for Abstract Musical Objects
Tenkanen, A.: *Proceedings of the IEEE International Conference on Intelligent Computing and Intelligent Systems* (ICIS 2009), Shanghai, Vol. 4, 472–476 (2009)

In this paper, I present a new and effective approach to finding a suitable similarity measures for set-class analysis. An automated function generation, which resembles the so-called *genetic programming*, found approximately 450 new similarity functions that satisfy the required mathematical conditions needed in measurements. Some of the generated functions produced stronger correlations with empirical data (gathered by Tuire Kuusi) than *REL* (David Lewin, 1980 [43]), perhaps the most successful model in past studies. As a satisfying by-product, the results hint at the fact that there may be a connection between the perceived closeness of pitch-class sets and Roger Shepard's universal cognitive models, according to which the probability of generalisation declines exponentially according to the metric distances between the stimuli if a physical parameter space is mapped onto an individual's psychological space.

3.2 General conclusions and future work

The possibilities offered by computer-assisted music analysis are numerous and differ in many ways from those of the manual methods, even though the aims of both approaches may overlap. The origin of computer-assisted analysis dates back to the 1940s, but its status has increased after the debut of the personal computer, when it attained a firm position from which it has subsequently enriched the field with new viewpoints. Today, programming environments offer flexible tools for music analysis. The computer enables extensive and complex calculations, which are in practice unattainable by human capacity alone. For example, whenever two or more musical parameters are taken into consideration the advantages of the computer tools for music analysis become evident. On the other hand, compared with the analysis done by hand, a computational method requires precisely defined constraints, conditions and function rules, which may lead to inflexible and narrow applications.

In the present study, I have formalised a fully automatic analytical method for music analysis called Comparison Structure Analysis, which extensively utilises similarity measurements between musical entities. As opposed to many earlier 'pocket calculators' that were applied to set-class analysis, its main components (i.e. automated segmentation, the similarity measure and the comparison structure) enable us to consider the whole composition at the same time. Since comparison structure analysis can be applied to several different musical dimensions, the similarity measures have been employed to pitch-class set, set-class, rhythmic and melodic segments.

Comparison structure analysis can be seen as a particular type of pattern matching application: the same ideas can be used for any time-series type data. The effectiveness of CSA can be condensed into the concept of the comparison structure, the selection of which forms a crucial part of the method. The comparison structure embodies a particular musical property or properties, which are scanned through a piece. Whereas in text information retrieval we try to find exact matches for a 'query,' or in DNA analysis a particular DNA-sequence, in CSA we evaluate the fuzzy prevalence of the selected comparison structure through the musical composition producing intuitively accessible representations of musical surfaces in the different forms of graphs. These may represent dynamic trends according to the selected comparison structure or centroids (mean points) if CSA is applied to a classification task. The trend curves, for their part, may reveal the formal divisions of musical compositions.

In Section 1.3, I discussed some desirable properties of the computer-assisted music analysis application at a very general level. The 'fixed-length' segmentation, which creates a kind of algorithmic 'projection' of the musical data, forms the most critical part of the method. However, in Appendix A.1 I show that two quite different segmentation methods for harmonic analysis can create highly similar results, i.e. CSA seems not to be unduly sensitive to the segmentation method. This is of fundamental importance for checking the quality of a given model as well as the robustness and reliability of the results it produces. Naturally, we do not assume a single segmentation solution for all purposes, since the needs vary depending on the musical dimension. However, by anchoring the segments and the similarity values to the onsets of the musical data and thereafter to bars, we are able to compare the trends of several different musical dimensions. The number of derivative applications is numerous. In addition, the same results can offer viewpoints on different types of conclusions as well. For example, the basic type of CSA produces a trend curve that may be informative in and of itself. On the other hand, it may also depict the relative coherence of the piece (with respect to other similar pieces) according to the particular comparison structure that is utilised. This kind of analysis was not, however, put to the test in our studies. In addition, nothing hinders us from using comparative methods between two different compositions; a procedure that allows us to find similar sections between compositions. The sample analyses offered not only new viewpoints on music analysis, but they were also able to confirm the previous observations of other analysts.

Several new similarity measures were introduced for tonal stability, rhythmic and set-class similarity measurements. The most advanced results were obtained by employing automated function generation – comparable with the so-called genetic programming – to search for an optimal model for set-class similarity measurements. However, the results of CSA seem to agree strongly, independent of the type of similarity function employed in the analysis: the characteristic differences of the segment vectors appear to emphasise robustness in the results.

Comparison structure analysis is a flexible tool for many different kinds of analytical considerations, but it still misses some desirable properties: by detecting only the 'surface data' it does not reveal anything about the middle-ground connections in music and, in addition, CSA does not answer the important and interesting question about the rules of composition. It is known, however, that the latter types of programs are currently being developed [19].

The analysis examples presented in this thesis are still quite restricted and simple. In the future, more advanced approaches may be applied by segmenting the composition systematically in several dimensions, by employing CSA for the different dimensions and by investigating correlations between the dimensions. In order to extend CSA by combining it with other methods, such applications as *transfer entropy* or *Granger causality*, which detect the asymmetrical direction of information flow between two coupled processes, may be used [39, 40]. For example, CSA can first be utilised to detect invertible counterpoints of polyphonic music [III] and, thereafter, we may use transfer entropy for measuring the causality between two or more counterpoints. This might reveal something about the importance between the different voices in polyphonic music. We may also use the average lengths of the notes of the themes for the same purpose, since we know that in classical polyphony the themes of shorter notes are more probably determined by the themes of longer notes, rather than vice versa (Figure 5). Such approaches may open new perspectives for the order of the composition stages and for studying the development of the style of a particular composer. For example, it is known that the importance of strict contrapuntal techniques and invertible counterpoint increased in the compositions of Johann Sebastian Bach during his career.

Another type of excursion might be to study how analytical results correspond to music listeners' perceptions, since CSA may reflect easily observable properties of the piece. Even more imaginative applications could be developed in other research fields, for example in bird song analysis.

Comparison structure analysis opens up several possibilities for a music analyst. However, as in the case of any research tool, good research questions are still worth more than their weight in gold.

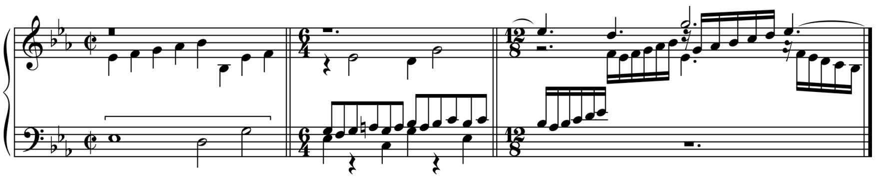


Fig. 5: Three extracts from Johann Sebastian Bach's fugue, BWV 552. It is probable that the counter themes of shorter note durations have been constructed after the theme of long notes – Eb-D-G.

References

[1] Rahn, J.: *Basic Atonal Theory*. New York: Schirmer Books (1980)

[2] Castrén, M.: *RECREL. A Similarity Measure for Set-Classes*. PhD dissertation. Helsinki: Sibelius Academy (1994)

[3] Pople, A.: Modeling Musical Structure. In Eric Clarke and Nicholas Cook (eds.), *Empirical Musicology: Aims, Methods, Prospects* (New York and Oxford: Oxford University Press), 127–156 (2004)

[4] Winold, A.: Music Analysis: Purposes, Paradigms, and Problems. *Journal of Music Theory Pedagogy* 7, 29–40 (1993)

[5] Cook, A.: *A Guide to Musical Analysis*. New York. W. W. Norton and Co (1987)

[6] Davis, D. S.: *Computer Applications in Music: A Bibliography*. Supplement 1. Madison: A-R Editions (1993)

[7] Cook, N.: Computational and Comparative Musicology. In: Clarke, E. and Cook, N. (Eds.). *Empirical Musicology: Aims, Methods, Prospects*, 103-126. Oxford: Oxford University Press (2004)

[8] Anagnostopoulou C., Share C., Conklin, D.: Xenakis' Keren: A Computational Semiotic Analysis. In Makis Solomos, Anastasia Georgaki, Giorgos Zervos (eds.), *Definitive Proceedings of the International Symposium Iannis Xenakis*, Athens 2005 (2006)

[9] Forte, A.: Motivic Design and Structural Level in the First Movement of Brahms's String Quartet in C Minor, *Musical Quarterly*, 69, 471–502 (1983)

[10] Huron, D.: What is a musical feature? Forte's analysis of Brahms's opus 51 no 1, revisited, *Music Theory Online*, 7(4) (2001)

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

[12] Bronson B. H.: Mechanical Help in the Study of Folk Song. *The Journal of American Folklore* 62(244), 81–86 (1949)

[13] Rhodes, W.: The Use of the Computer in the Classification of Folk Tunes. In proceedings of *International Folk Music Council Conference* (IFMC1964), Budapest, 339–343 (1965)

[14] Sapp, G.S.: Visual Hierarchical Key Analysis. *ACM Computers in entertainment*, 4(4), 1–19 (2005)

[15] Eerola, T. and Toiviainen, P.: *MIDI Toolbox: MATLAB Tools for Music Research*. Jyväskylä: Department of Music, University of Jyväskylä (2004)

[16] Conklin D.: Melodic analysis with segment classes. *Machine Learning*, 65(2-3), 349–360 (2006)

[17] Clarke, E. and Cook N.: *Empirical Musicology. Aims, Methods Prospects*. New York: Oxford University Press (2004)

[18] Widmer G., Dixon S., Knees P., Pampalk E., Pohle T.: From Sound to "Sense" via Feature Extraction and Machine Learning: Deriving High-level Descriptors for Characterising Music. In P. Polotti and D. Rocchesso (eds.), *Sound to Sense: Sense to Sound: A State-of-the-Art in Sound and Music Computing*. Logos Verlag, 161-190 (2007)

[19] Cope, D.: A Musical Learning Algorithm. *Computer Music Journal* 28(3), 12–27 (2004)

[20] Buteau, C. and Vipperman, J.: Representations of motivic spaces of a score in OpenMusic. *Journal of Mathematics and Music* 2(2), 117–134 (2008)

[21] Toussaint, G. T.: A comparison of rhythmic similarity measures. *Proceedings of the 5th International Conference on Music Information Retrieval* (ISMIR04). Universitat Pompeu Fabra, Barcelona, 242–245 (2004)

[22] Lerdahl, F. and Jackendoff, R.: *A Generative Theory of Tonal Music*. Mass.: M.I.T. Press (1983)

[23] Huovinen, E.: The Justification of Music-structural Claims, *Philosophy of Music Education Review*, 16(1), 21–39 (2008)

[24] Christopher, H.: Segmentation and Process in Post-Tonal Music. *Music Theory Spectrum* 3, 54–73 (1981)

[25] Parncutt, R.: Systematic Musicology and the History and Future of Western Musical Scholarship. *Journal of interdisciplinary music studies*, 1 (1), 1–32 (2007)

[26] Castine, P.:, *Set Theory Objects: Abstractions for Computer-Aided Analysis and Composition of Serial and Atonal Music*. Berlin: Peter Lang, (1994)

[27] Neve G., Orio N.: A Comparison of Melodic Segmentation Techniques for Music Information Retrieval, *Lecture Notes in Computer Science*, Springer, 49–56 (2005)

[28] Kagolovsky, Y. and Moehr, J. R.: A New Look at Information Retrieval Evaluation: Proposal for Solutions. *Journal of Medical Systems*, 28 (1), 103–115 (2004)

[29] Cambouropoulos, E.: *Towards a General Computational Theory of Musical Structure*. Ph.D, The University of Edinburgh (1998)

[30] Cambouropoulos, E., Crawford, T., Iliopoulos, C.S.: Pattern Processing in Melodic Sequences: Challenges, Caveats and Prospects, Special Issue on Pattern Processing in Music Analysis and Creation, *Computers and the Humanities* 35(1), 9–21 (2001)

[31] Cambouropoulos, E.: How similar is similar? *Musicae Scientiae*, Discussion Forum 4B, 7–24 (2009)

[32] Chen, S., Ma, B. and Zhang, K.: On the Similarity Metric and the Distance Metric. *Theoretical Computer Science* 410, 2365–2376 (2009)

[33] Shmulevich, I., Yli-Harja, O., Coyle, E., Povel, D.-J. and Lemström, K.: Perceptual Issues in Music Pattern Recognition: Complexity of Rhythm and Key Finding. Special Issue on Pattern Processing in Music Analysis and Creation, *Computers and the Humanities* 35(1), 23–35 (2001)

[34] Ilomäki, T.: *On the Similarity of Twelve–Tone Rows*. Studia Musica 30, Sibelius Academy, Helsinki (2008)

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

[36] Tenkanen, A. and Huovinen, E.: Laajan musiikillisen korpuksen säveltasorakenteiden tietokoneavusteinen tutkimus (Studying the pitch organisational features of large compositional corpus). *Suomen musiikintutkijoiden symposiumin satoa*, p. 161–165. Jyväskylä 2005. http://www.jyu.fi/musica/symposium/symposiumin_satoa_2005.pdf (2005)

[37] Kuusi, T: *Set-Class and Chord: Examining Connection between Theoretical Resemblance and Perceived Closeness*. Sibelius Academy, Helsinki (2001)

[38] Krumhansl, C.L.: *Cognitive Foundations of Musical Pitch*. Oxford University Press, Oxford (1990)

[39] Schreiber, T.: Measuring Information Transfer. *Physical Review Letters* 85, 461–464 (2000)

[40] Kulp, C.W. and Schlingmann, D.: Using Mathematica to Compose Music and Analyze Music with Information Theory. First International Conference of Mathematics and Computation in Music (MCM2007), Berlin. Timour Klouche and Thomas Noll (Eds). Springer series on *Communications in Computer and Information Science,* Vol. 37, 441-448 (2009)

[41] Quinn, I.: 'Listening to Similarity Relations', *Perspectives of New Music*, 39(2), 108–158 (2001)

[42] Väisälä, O.: New Theories and Fantasies on the Music of Debussy: Post-Triadic Prolongation in "Ce qu'a vu le vent d'ouest" and Other Examples, in Allen Cadwallader (ed.), *Essays from the Third International Schenker Symposium* (Hildesheim: Georg Olms Verlag), 165–95 (2006)

[43] Lewin, D.: A Response to a Response: on Pcset Relatedness. *Perspectives of New Music* 18(1–2), 498–502 (1979–80)

[44] Shepard, R.N.: Toward a Universal Law of Generalization for Psychological Science. *Science* 237, 1317–23 (1987)

A Appendices

The question of how sensitive the CSA is to parameter changes (e.g. different segmentation methods) and of whether the results attained by CSA are statistically significant are not discussed in publications [I-V]. Because of this, I consider these issues in the next two Appendices, A.1 and A.2. I also consider the 'normative harmony' of Debussy's prelude *Ce qu'à vu le vent d'ouest* anew. That issue was discussed earlier in article [I], in which the set-class-based CSA was applied to validate such a hypothesis. However, the problem in the previous approach was that the set classes are not transpositionally fixed, while Väisälä's pitch-class collection is. Therefore, the same piece is considered once again in Appendix A.3 by using the tonality-oriented CSA, which gives more support – at least – to the tonic of the normative harmony of the piece.

A.1 Testing the sensitivity of CSA

I demonstrate here a set of systematic calculations testing the effect of parameter changes on both set-class- and tonal-stability-based types of CSA. The parameters observed were the segmentation cardinality (here cardinalities 3-5), the similarity function and the segmentation method. Three different similarity measures were included in the set-class-based analysis

- *REL* [I]
- *expcos*, i.e. function (8) in [V]
- *si* (in [11])

and for the tonal-stability based CSA

- *parnc* [IV]
- *cofrel* [IV]
- *kk-prob* [IV]

The segmentation methods included:

- the tail segmentation [I]
- the clustering segmentation, presented in Section 2.2. and in [III]

By executing all the parameter combinations we get altogether 18 ($3 \cdot 3 \cdot 2$) resulting similarity curves for one composition and one type of CSA (SC- or tonal-based CSA) when only one comparison structure (set-class or pc-set) is applied. Figure 6 illustrates such a set of set-class-based CSAs for Debussy's *Ce qu'à vu le vent d'ouest* from the first book of *Préludes*. The SC 4–26 [0, 3, 5, 8] is used as a comparison set-class. The values of each curve have been *normalised* to the value range of 0-1 and thereafter averaged over bars. In order to illustrate the deviations of the individual resulting curves better, a reference curve was produced by calculating a median standard deviation value for each bar separately. This gives 71 standard deviation values for Debussy's composition, since it includes 71 bars of music. In fact, this kind of an approach might be utilised with CSA more generally: taking the median curve as the resulting output for CSA produces a certain redundancy in the overall reliability of the calculation. Mean values could have been calculated as well, but the mean is influenced by outliers while the median is robust.

The resulting curves seem to agree strongly: Even the weakest correlation between the set-class curves is still statistically significant (cor=0.46, $p<0.0001$). For this special case, the first curve is produced using parameters *REL*, a clustering segmentation and 3 as the segmentation cardinality, and another curve is produced using expcos, a tail segmentation and 5 as the segmentation cardinality.

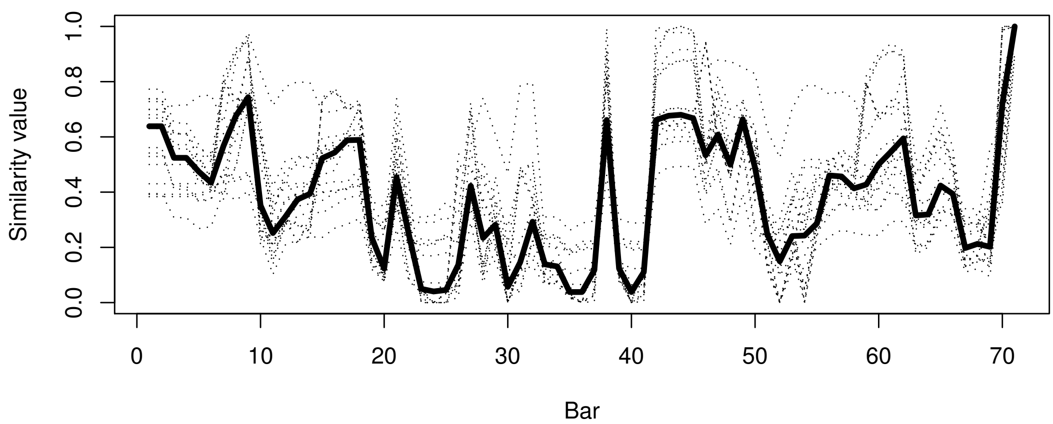


Fig. 6: Two different segmentation methods, three segmentation cardinalities (3-5) and three similarity measures *REL*, *expcos* and *si* were employed to produce eighteen set-class-based CSA's $csa(X, 3...5, 4-26)$ of Debussy's *Ce qu'à vu le vent d'ouest*.

In order to study the sensitivity of CSA, for example with regard to the particular segmentation method, we have to consider only half of the curves – those that are produced by employing either the tail or the clustering segmentation. The less the set of curves deviate, the more robust the segmentation method. However, it must be noted that this kind of calculation expresses only the relative differences between different arguments, i.e. whether the average deviation is larger when using the tail segmentation than when using the clustering segmentation. I used again three different cardinalities and three similarity measures, but, in addition, all of the 128 T_n-type set classes of cardinalities 3-5 (3–1, ..., 5–Z38B) as comparison structures with five different compositions:

- J.S. Bach's Chorale *Es ist genug*
- Bach's *Invention nr. 5*

- C.P.E. Bach's *Die Neue Litanie*
- Debussy's prelude for piano *Ce qu'à vu le vent d'ouest*
- Alban Berg's *Invention on a Key* from Wozzeck.

This gives 5760 ($3 \cdot 3 \cdot 128 \cdot 5$) comparison curves for studying the robustness of either of the main segmentation methods. For that purpose, the standard deviation was again calculated for each bar. Table 1a reports the median of (median) sd-values regarding both the tail segmentation and the clustering segmentation methods for Debussy's prelude, which are 0.127 and 0.095, which in turn means that the resulting curves deviate less when the clustering segmentation is used.

Analogously, I also carried out tests of the robustness of the three similarity measures in the face of changes across the other parameters (segmentation method, segmentation cardinality) – again, separately for each of the five compositions, seen also in Table 1a. Furthermore, the standard deviation values for each composition have also been calculated by applying the tonal stability measures (Table 1b) with the corresponding 128 pitch-class sets, i.e. the prime form transpositions of the 128 set classes as comparison structures.

Table 1. The effect of different parameters on CSA in five compositions. The smaller the median for the standard deviation values, the more robust the parameter is when applied to the particular composition.

a)

SC-analysis	Tail	Clust	*REL*	*expcos*	*si*
Es ist genug	0.147	**0.135**	0.168	**0.166**	0.185
Invention	0.151	**0.136**	0.157	**0.147**	0.167
Litanie	0.129	**0.114**	0.128	**0.113**	0.167
"Ce qu'à vu ..."	0.127	**0.095**	0.113	**0.101**	0.135
Wozzeck	0.111	**0.078**	0.112	**0.104**	0.127

b)

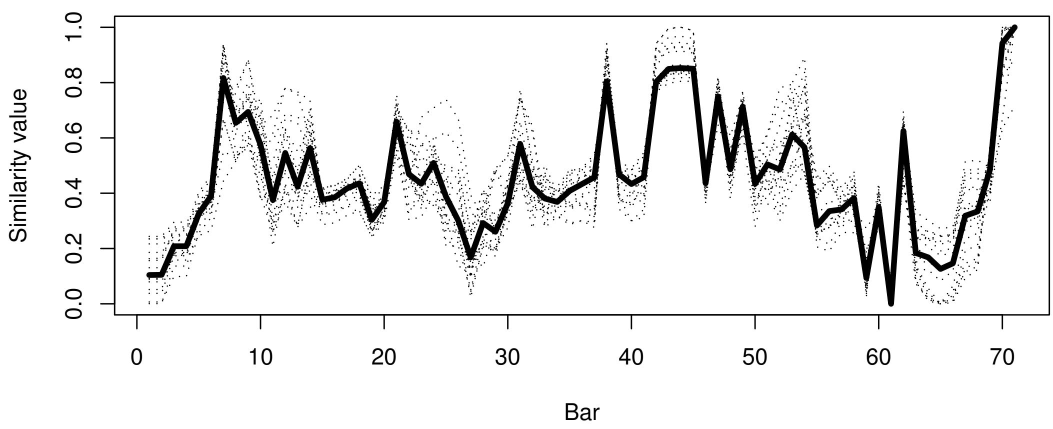
Tonal analysis	Tail	Clust	*parnc*	*cofrel*	*kk-prob*
Es ist genug	**0.105**	0.126	0.083	0.083	0.083
Invention	0.110	**0.109**	**0.056**	0.060	0.060
Litanie	**0.068**	0.077	0.058	**0.053**	0.054
"Ce qu'à vu ..."	0.110	**0.106**	0.078	**0.073**	0.074
Wozzeck	**0.094**	0.103	0.067	**0.066**	0.067

The clustering segmentation and the exponential similarity measure *expcos* seem to be more robust than their alternatives as far as the set-class-based CSA is concerned (see Table 1a). Tonality-based CSA displays greater variability on this issue (Table 1b) but, on the other hand, the standard deviation values are also smaller. This is most decidedly due to the fact that pc-sets with the transpositional information represent more characteristic objects compared with their set-class representatives.

Analogously to the previous set-class-based analysis (shown in Figure 6), Figure 7a gives a set of 18 tonality-oriented CSA curves for Debussy's *Ce qu'à vu le vent d'ouest*. As the reader will remember, these analyses are now based on the segments of the composition taken as transpositionally determined pitch-class sets and on calculations using the tonal stability measures and indicates much less deviation between the comparison curves than SC-based analysis. The deviation is also dependent on the nature of the comparison structure. In the previous set-class-based analysis of Debussy's prelude (calculated for Table 1a, fourth row), the most stable set-class proved to be the structurally, highly characteristic whole-tone pentachord 5–33[0,2,4,6,8] – of which the interval-class vector is [040402]), while the SC 5–26B[0,3,4,6,8] – with a rather plain interval-class vector [122311] – produced the results that deviated the most. In the tonality-based CSA, these included the diatonic pentachord 5–35[0,2,4,7,9] and the symmetric tetrachord 4–25[0,2,6,8], respectively. The deviations in the tonality curve (Figure 7a) correlate strongly (cor=-0.57, $p<0.001$) with the pitch-class cardinality per bar (Figure 7b), i.e. the lower the bar-cardinality the greater the deviation between curves. Tonality-based CSA is thus sensitive to the number of distinct pitch classes in the bars. This correspondence can be best seen at bar 25, where the pc-cardinality per bar decreases to 2 (compare Figures 7a and 7b). Despite this fact, all the curves themselves correlate strongly ($p<0.001$) and, thus, each of them reflects the same trends in the piece as seen in Figures 6 and 7a. These results also support to Ian Quinn's finding that the various numerical models characterized as pc-set-class similarity relations agree with each other to a high degree, despite major differences in their internal workings [41]. The tonality curves are even more interrelated, with the smallest correlation value being 0.72 (c.f. Figure 7a).

The slight differences in sd-value between the compositions reflect not only the differences in their harmonic character and compositional texture but also the length of the pieces. For example, C.P.E. Bach's lengthy *Litanie die Neue,* with its 560 bars, represents an almost similar musical texture (see Figure 18 in [III]) as that of J.S. Bach's chorale *Es ist genug* (20 bars), but all the standard deviation values are smaller in the former case. Longer compositions with more samples naturally offer more robust results for sensitivity analysis. As far as the segmentation methods in both harmony-oriented CSA are concerned, the differences between the results appear to be more cosmetic than statistically significant. This is good news since, thus far, the segmentation issues have been one of the most controversial aspects of this method.

a)



b)

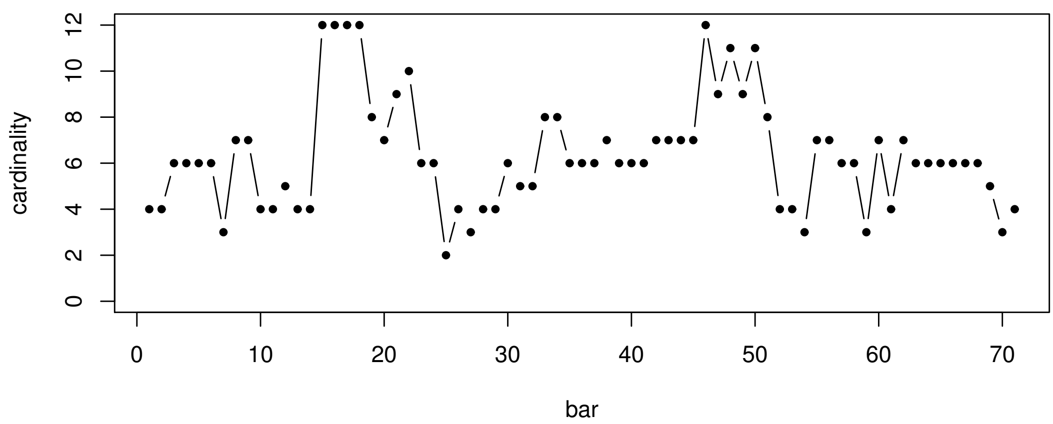


Fig. 7: a) Two different segmentation methods, three segmentation cardinalities (3-5) and three similarity measures *parnc*, *cofrel* and *kk-prob* were used to produce eighteen tonal-stability based CSA's for Debussy's *Ce qu'à vu le vent d'ouest*, $csa(X, 3...5, \{1, 3, 6, 10\})$.
b) The number of distinct pitch classes in bars.

A.2 Testing the statistical significance of CSA

Although there is a lot of potential for the use of statistics in musicology, I have not thus far applied statistical tests together with CSAs to investigate the interesting issue of whether or not the similarity values actually represent statistically significant results: How do we know for certain that the CSAs really reflect the properties of the compositions? Could similar curves be produced randomly? In order to demonstrate such an approach, I proceed to analyse Debussy's *Ce qu'à vu le vent d'ouest*, again using set-class comparisons. Some sort of time-series approach (e.g. autocorrelation, probabilities of longer segment combinations etc.) would certainly seem most applicable to CSA. However, this time the sequence information is ignored and instead the distribution of values averaged over bars is considered. In fact, these *mean values* over bars indirectly reflect consecutive *repetitive structures* that are common in compositions but uncommon in random processes. This effect is illustrated in Figures 8a-d.

Figure 8a represents CSA $expcos(X,4,4\text{–}25[0,2,6,8])$ of Debussy's composition before the averaging procedure. The sharp-edged profile of the similarity curve with highly characteristic sections results from the usage of the clustering segmentation method. A lot of information is lost by averaging the values over bars, represented in Figure 8b. However, the profile of the similarity curve remains characteristic, because the repetitive segments – associated with shared values – are still able to produce extremely high mean values. In order to find a model to which we can compare our data, in order to determine its level of statistical significance, I generate a random distribution of similarity values. By replacing the original pitches of the composition with random pitches (while at the same time preserving other information, i.e. note onset times and note durations), we get a 'random composition' which is as long as the original and has as many notes. The resulting curve of the CSA in this case resembles random noise, shown in Figure 8c. Averaging the comparison values over bars produces a resulting curve that is quite plain and uncharacteristic, seen in Figure 8d. This is expected, since it is improbable that the randomisation process produces repetitive structures, which are required for particularly high or low mean values. When two Figures, 8b and 8d, are compared, it is intuitively obvious that the CSA of the original composition, represented in Figures 8a-b, cannot be the result of a random process.

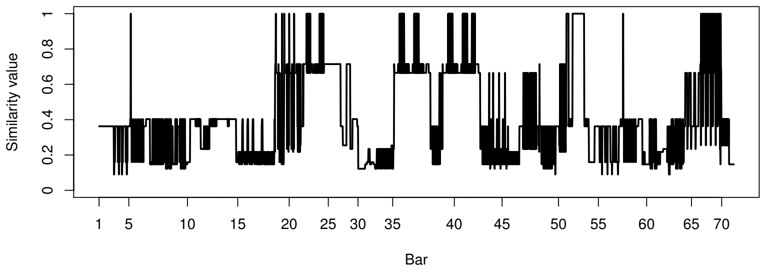
In order to *approximate* value boundaries that represent 5% level of significance, the random composition was generated 1000 times (remembering that, for CSA, each random composition needs a segmentation of its own) and all the resulting $expcos(X,4,4\text{–}25[0,2,6,8])$-values (71000 average values in total) were used to generate the random distribution, plotted as a histogram in Figure 9. When we are testing at the 5% level of significance, both tails will contain 2.5% of the values. The lower and upper-boundaries were calculated from the distribution using the R-program's *quantile*-function, which produces sample quantiles corresponding to the given probabilities 2.5% and 97.5%. These boundaries vary according to the comparison set-class. In our calculation, the lower limit of 0.185 corresponds to a probability of 2.5% and the upper limit of 0.384 to a probability of 97.5%. The boundaries have been marked in Figures 8b, 8d and 9.

The highest and the lowest values found in the averaged curve of the original composition (Figure 8b) are absent in the analysis of the random composition. The averaged values estimated for the random composition mainly stay between the evaluated borders, which provides evidence that the selection of the p-value (0.05) is reasonable. For each bar x of an analysed composition, the null hypothesis that 'the composer may have selected the notes in bar x randomly,' will be rejected if the observed mean CSA value for this bar is less than the lower boundary or greater than the upper boundary. Between these limits, we do not reject the null hypothesis. This does not mean, however, that those sections of the composition in which the similarity values settle between the two limits are composed randomly. In order to study these sections, we could test another property; for example, the number of either ascending or descending consecutive values, which are more improbable in a random than actual composition.

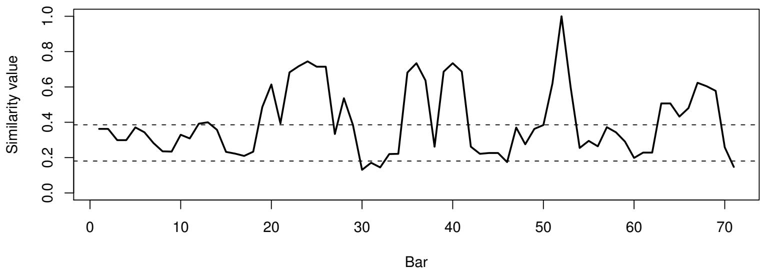
A final note of clarification. The SC 4–25[0,2,6,8] was selected as the comparison structure in the present demonstration because, among all 43 T_n-type tetrachords, it proved to produce a resulting curve that passes the upper limit most often; it did so in 29 out of 71 bars. This occurs even more often than in the case of another subset-class set of the whole tone scale, SC 4–24[0,2,4,8], which was shown to emerge as pre-eminent in Debussy's composition [I]. The upper limit for SC 4–24 is much higher, being 0.431. One reason for the difference (0.384 vs. 0.431) is that SC 4–25 -segments are two times rarer because of their symmetric structure. In 1000 random pieces, the number of SC segments were 60252 (4–24) and 29961 (4–25).

The significance test, which is based on pitch-class density estimation, may work within most compositional styles. However, it would be interesting to test if such evidence of significance could be detected in cases of 12-tone or totally serial compositions.

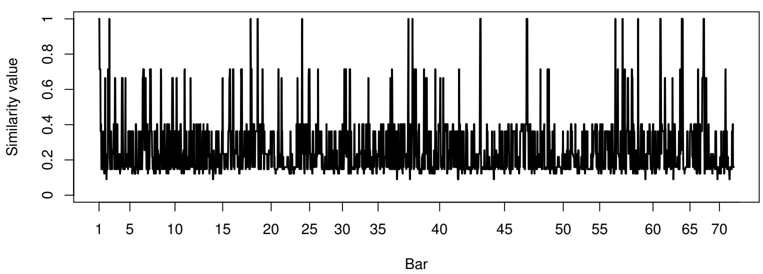
a)



b)



c)



d)

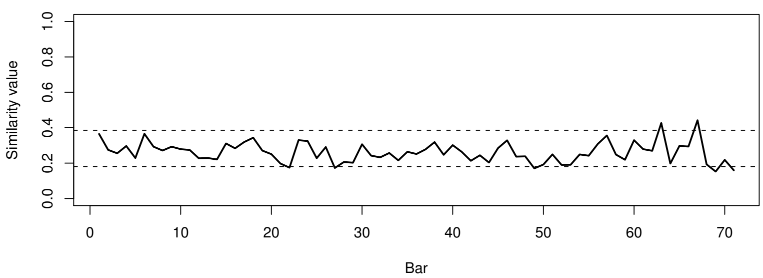


Fig. 8: CSA $expcos(X,4,4\text{–}25)$ of Debussy's *Ce qu'à vu le vent d'ouest* (a,b) and a 'random composition' (c,d).

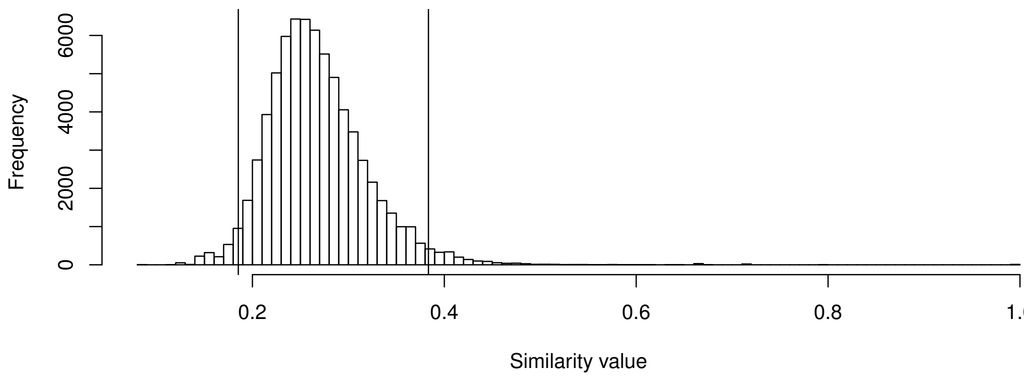


Fig. 9: Randomly generated distribution of $expcos(X, 4, 4–25)$ -values.

A.3 Once again: Reconsidering Debussy's *Ce qu'à vu le vent d'ouest*

In Appendix A.1, Figure 7a, I illustrate the sensitivity analysis in accordance with CSA by applying a constant comparison pc-set {1,3,6,10} to Debussy's prelude *Ce qu'à vu le vent d'ouest*. According to Väisälä, that pc-set {F#,A#,C#,D#}, termed an β-chord, forms a normative *reference harmony* in the composition [42]. However, our analysis in article [I] could not show quantitative evidence for the prevalence of SC 4–26. The main problem in employing set-class-based CSA for validating such hypotheses is that the set classes are not transpositionally fixed, whereas Väisälä's pitch-class collection is transpositionally fixed. Next, I briefly reconsider our analysis by applying both statistical pitch-class calculations and the tonality-oriented CSA to Debussy's prelude.

For the purposes of this analysis, the distribution of pitch classes was calculated by counting each pitch-class occurrence in each bar only once. The resulting distribution is seen in Figure 10, which seems to emphasise pitch classes 1, 3, 6 and 9, to the extent that the four most common pitch classes are considered. This hints at the fact that Väisälä's interpretation might be possible also in a statistical sense since the β-chord is the only collection among his other referential sets, six tetrachords and one pentachord all together, of which the tonal base is anchored to F#.

Then I studied which four-member pc-set would produce the smallest average stability value in the piece. I executed the tonality-based CSA by using the *parnc*, *cofrel* and *kk-prob* functions as similarity measures, the clustering segmentation and all 495 unique four-member, pitch-class-set combinations ({0,1,2,3},{0,1,2,4}, ...,{8,9,10,11}) as comparison pc-sets. The segmentation cardinality was 4 (that is a 'safe' selection since the median cardinality per bar is 6, with 6.33 being the average mean). I assigned the median value out of 71 bar-based comparison values to each comparison structure. Thereafter, the comparison pc-sets that produced the smallest stability value by evaluating each similarity measure separately were chosen. These were {1,2,6,7}, {1,3,8,10} and {0,1,2,6}, respectively.

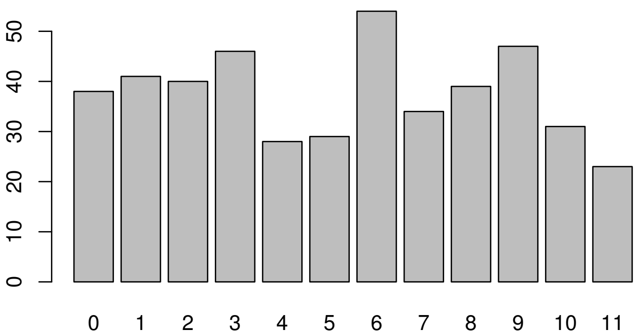


Fig. 10: The distribution of pitch classes counted in each bar separately.

Although the tonality curves of the earlier analysis strongly agree (Figure 7a), the pc-sets in the present analysis seem to offer different interpretations. The previous similarity curves are based on a huge number of segments, which statistically show much more robust results than the present comparison pc-sets found by applying the three similarity functions of different types. The question is: Could we somehow interpret these particular comparison-set candidates which represent some sort of intangible, abstract and average property of harmony in a piece? They seem not to reflect the pc-distribution in a straightforward manner since, for example, pc7 and pc10 that are present in some of the pc-sets with smallest stability values are not emphasised in the pc-distribution. Instead, they provide a more structural viewpoint for the average tonality of the piece. *cofrel* is based on the circle of fifths and this may have an effect on the fact that the pc-set that it proposes forms a consecutive sequence of fifths, i.e. pitch classes 1, 8, 3, 10. Thus, it includes the open fifths of the dominant F# major and the relative minor chord of F# major. The comparison set {1,2,6,7} proposed by *parnc* includes two open fifths that might be seen as representing a tensive *Neapolitan* N-I relation, with F# taking the place of I. *kk-prob*-function, for its part, weights more the significance of the major and minor third [IV]. The proposed set {0,1,2,6} looks to be more difficult to interpret. However, the open fifth pc6-pc1 is available in this set too and when its hypothetical 'tonal centre' is evaluated by using the means presented in [IV], F# appears to be the most apparent candidate for the average tonal centre of the set and for the whole piece.

A problem with CSA is that it does not take the rhythmic and articulative context into account; for example, by weighting structurally more important places. However, even though we could not estimate the best referential set for the piece in a 'Schenkerian' sense, our new findings give more support to Väisälä's interpretation about the tonic of the normative harmony than our previous analyses.