SYSTEMATIC MULTI-SCALE SET-CLASS ANALYSIS

Agustín Martorell
Universitat Pompeu Fabra
agustin.martorell@upf.edu

Emilia Gómez
Universitat Pompeu Fabra
emilia.gomez@upf.edu

ABSTRACT

This work reviews and elaborates a methodology for hierarchical multi-scale set-class analysis of music pieces. The method extends the systematic segmentation and representation of Sapp's 'keyscapes' to the description stage, by introducing a set-class level of description. This provides a systematic, mid-level, and standard analytical lexicon, which allows the description of any notated music based on fixed temperaments. The method benefits from the representation completeness, the compromise between generalisation and discrimination of the set-class spaces, and the access to hierarchical inclusion relations over time. The proposed *class-matrices* are multidimensional time series encoding the pitch content of every possible music segment over time, regardless the involved time-scales, in terms of a given set-class space. They provide the simplest information mining methods with the ability of capturing sophisticated tonal relations. The proposed *class-vectors*, quantifying the presence of every possible set-class in a piece, are discussed for advanced explorations of corpora. The compromise between dimensionality and informativeness provided by the class-matrices and class-vectors, is discussed in relation with standard content-based tonal descriptors, and music information retrieval applications.

1. INTRODUCTION

Pitch-class set theory has been used in music analysis practice since decades. However, its general applicability to post-tonal music has contributed, and still contributes, to be perceived as for specialists only. This apparent difficulty is far from real, and just a matter of the application context. The systematic and objective nature of the theory, together with the compactness of the basic representations, constitutes a powerful and flexible descriptive framework suited for any kind of pitch-based music. [1] This description level is purposeful for several music information retrieval (MIR) applications, such as structural analysis, similarity, pattern finding, classification, and generation of content-based metadata. More interestingly, it provides a means for approaching complex topics, such as similarity, in alternative and insightful musically-grounded scenarios. In addition, the basic descriptors are trivial to compute, and they can be readily exploited by standard information mining techniques.

The remaining of this work is organised as follows. Section 2 introduces the basic set-theoretical concepts, and contextualise them in terms of our systematic analysis endeavour. Section 3 describes the computational approach. Sections 4 and 5 discuss the method in several application contexts. Section 6 summarises the proposed method, and points to future extensions.

2. BACKGROUND

2.1 Set-class description

Pitch class [1] is defined, in the twelve-tone equal tempered system (TET), as an integer representing the residue class modulo 12 of a pitch, that is, any pitch is mapped to a pitch class by removing its octave information. A *pitch-class set* (henceforth *pc-set*) is a set of pitch classes without repetitions in which the order of succession of the elements in the set is not of interest. In the TET system, there exist $2^{12} = 4096$ distinct pc-sets, so a vocabulary of 4096 symbols is required for describing any possible segment of music. Any pc-set can also be represented by its intervallic content [5]. Intervals considered regardless of their direction are referred to as *interval classes*. The total count of interval classes in a pc-set can be arranged as a six-dimensional data structure called an *interval vector* [4].

Relevant relational concepts for analysis are the *set-class equivalences*, whereby two pc-sets are considered equivalent if and only if they belong to the same *class*. As pointed out by Straus, equivalence is not the same thing as identity, rather it is a link between musical entities that have something in common. This commonality underlying the surface may eventually lend unity and/or coherence to musical works [12]. In this respect, the class equivalences can be conceived as *all or nothing* similarity measures between two pc-sets. In the context of pc-sets, the number of pitch classes in a set is referred to as its *cardinality*. This is perhaps the coarsest measure of similarity. Despite its theoretical relevance, cardinality is too general a notion of similarity to be of use in many analytical situations. Among the many equivalence systems in the set-theoretical

[1] In which the concepts of octave equivalence and fixed temperaments are applicable. Although the pitch relations of interest may be quite different, depending on the temperament and the applied context, any discrete pitch organization of the octave can be handled by the general mathematical framework. In this work, we bound to the twelve-tone equal temperament.

literature, three of them are particularly useful:

1. *Interval vector equivalence* (iv-equivalence), which groups all the pc-sets sharing the same interval vector. There exist 197 different iv-types.

2. *Transpositional equivalence* (T_n-equivalence), which groups all the pc-sets related to each other by transposition. There exist 348 distinct T_n-types.

3. *Inversional and transpositional equivalence* (T_nI-equivalence), which groups all the pc-sets related by transposition and/or inversion. There exist 220 different T_nI-types (also referred to as T_n/T_nI-types).

Aside the comprehensive coverage of every possible pc-set, the compromise between discrimination and generalisation of these class-equivalence systems fits a wide range of descriptive needs, thus their extensive usage in general-purpose music analysis. From them, iv-equivalence is the most general (197 classes). It shares most of its classes with T_nI-equivalence (220 classes), with some exceptions, named *Z-relations* [4], for which the same interval vector groups pc-sets which are not T_nI-equivalent [7]. The most specific from the three systems is T_n-equivalence.

2.2 Systematic approaches to set-class analysis

To date, one of the most systematic approaches to set-class surface analysis is proposed in [6], under the concept of 'tail-segment array', whereby every note in a composition is associated with all the possible segments of a given cardinality that contains it. This segmentation is combined with certain set-class-based 'detector functions', in order to obtain summarized information from music pieces and collections. The usefulness of the method is comprehensively discussed in the context of style characterization. Some limitations of this technique are addressed in [8], by first identifying the segmentation, description and representation stages of the method, and extending systematization to all of them simultaneously. This is done by combining the exhaustive segmentation and representation of Sapp's 'keyscapes' [11], with a systematic description of the segments in terms of set-classes. The multidimensional, massive and overlapping information resulting from this method, is managed by summarising features and interfacing design, targeting specific analytical tasks.

3. MULTI-SCALE SET-CLASS ANALYSIS

This work elaborates directly upon [8], in which detailed and extended discussions can be consulted. A description of our general method follows.

3.1 Segmentation

The input to the system is a sequence of MIDI events, which can be of any rhythmic or polyphonic complexity. This signal is processed by the segmentation stage, for which two different algorithms are used: a) an approximate technique, non comprehensive but practical for interacting with the data; b) a fully systematic method, which exhausts all the segmentation possibilities.

The approximate method applies many overlapping sliding windows, each of them scanning the music at a different time-scale. The minimum window size and the number of time-scales are user parameters, and can be fine tuned as a trade-off between resolution and computational cost. The same hop size is applied for all the time-scales, in order to provide a regular grid for visualisation and interfacing purposes. Each segment is thus indexed by its centre location (time) and its duration (time-scale).

The fully systematic method is required for the quantitative descriptors in which completeness of representation is necessary. It is computed by finding every change in the pc-set content, whether the product of onsets or offsets, and segmenting the piece by considering all the pairwise combinations among these boundaries.

3.2 Description

Denoting pitch-classes by the ordinal convention (C=0, ..., B=11), each segment is analysed as follows. Let $b_i = 1$ if the pitch-class i is contained (totally or partially) in the segment, or 0 otherwise. The pc-set in the segment is encoded as an integer $p = \sum_{i=0}^{11} b_i \cdot 2^{11-i} \in [0,4095]$. This integer serves as an index for a precomputed table of set classes, [2] including the iv-, T_nI- and T_n-equivalences (discussed in Section 2.1). For systematisation completeness, the three class spaces are extended to include the so-called *trivial forms*. [3] With this, the total number of interval vectors rises to 200, while the T_nI- and T_n-equivalence classes sum to 223 and 351 categories respectively. In this work, we use Forte's cardinality-ordinal convention [4] to name the classes, as well as the usual A/B suffix for referring to the prime/inverted forms under T_n-equivalence. We also follow the conventional notation to name the Z-related classes, by inserting a 'Z' between the hyphen and the ordinal. As an example, a segment containing the pitches $\{$G5,C3,E4,C4$\}$ is mapped to the pc-set $\{$0,4,7$\}$ and coded as $p = 2192$ (100010010000 in binary). The precomputed table is indexed by p, resulting in the interval vector $\langle 001110 \rangle$ (iv-equivalence, grouping all the sets containing exactly 1 minor third, 1 major third, and 1 fourth), the class 3-11 (T_nI-equivalence, grouping all the major and minor trichords), and the class 3-11B (T_n-equivalence, grouping all the major trichords). The discrimination between major and minor trichords is thus possible under T_n-equivalence (3-11A for minor, 3-11B for major), but not under iv- or T_nI-equivalences.

3.3 Representation

The main data structure, named *class-scape*, is the set-class equivalent of Sapp's 'keyscapes' [11]. It represents the class content of every possible segment, indexed by

[2] As formalised in [4]. See *Supplemental material* (Section 7).

[3] The null set and single pitch classes (cardinalities 0 and 1, containing no intervals), the undecachords (cardinality 11) and the universal pc-set (cardinality 12, also referred to as the *aggregate*).

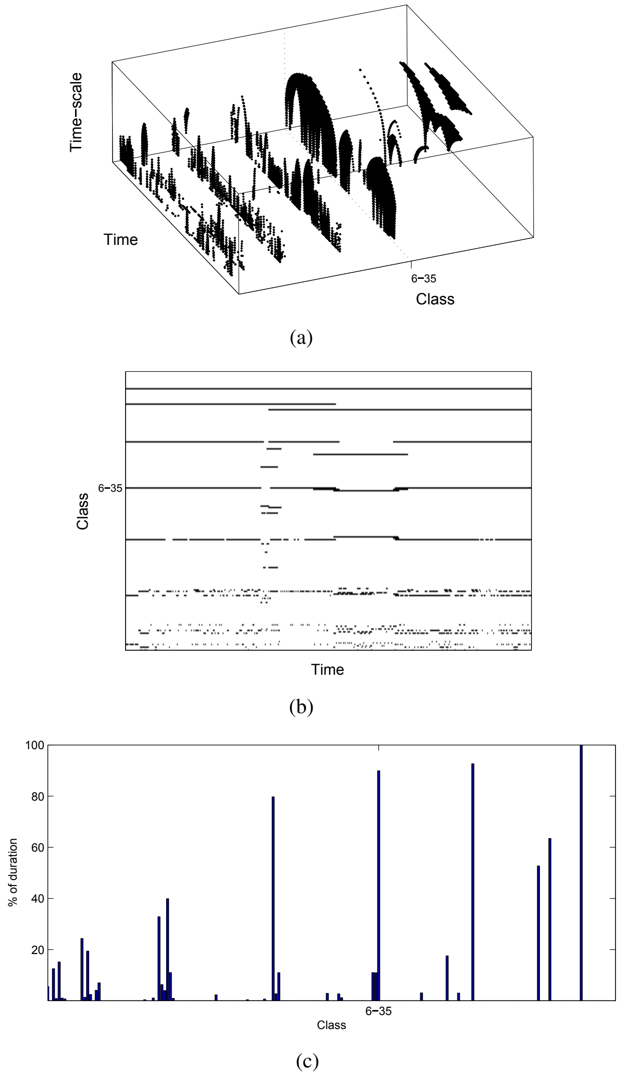


Figure 1: Debussy's *Voiles*. a) class-scape; b) class-matrix; c) class-vector.

their time position and duration. The dimensionality of the class-scapes (time, time-scale and class) is then reduced to more manageable, yet informative, data structures. The first reduction consists on projecting the class-scape to the time-class plane, which results in the concept of *class-matrix*. This is done by realising in time each point in the class-scape, thus retaining a substantial information from the lost dimension (time-scale). A further reduction summarizes the class-matrix in a single vector, named *class-vector*, by quantifying the presence of every possible class in the piece as a percentage of the piece's duration. The class-scape, class-matrix and class-vector, computed from Debussy's *Voiles* are depicted in Figure 1, with the prominent whole-tone scale (class 6-35) labelled as a reference.

4. MINING CLASS-MATRICES

In this section, we will review and elaborate upon the information conveyed by the class-matrices. Even with the loss of information, the reduction process from the class-scape to the class-matrix guarantees that every instantiation of every class is represented in the class-matrix, regardless the involved time-scales. The class-matrix represents the temporal *activation* of every possible class over time. A time *point* activated for a given class in the matrix means that it exist at least one segment containing this time point which belongs to such class. As the representation guarantees a strict class-wise separation, the class matrix constitutes a time-series of a special kind. It does not only capture evidence from every class instantiation over time, but it also informs about their set-class *inclusion relations*. The class-matrix, thus, embeds a considerable hierarchical information, allowing the analysis of the specific *constructions* of the class instantiations.

4.1 Case study: subclass analysis

An example of this analytical potential is depicted in Figure 2. It shows the comparison between the *pure diatonicisms* in Victoria's parody masses in Ionian mode [4] and Bach's preludes and fugues in major mode from the Well Tempered Clavier. This is done by first isolating the diatonic segments (activation of 7-35 in the class-matrix) of each movement, and constructing a *subclass-matrix* with the subset content of these segments. The differences can be quantified by computing the corresponding *subclass-vectors* out of the subclass-matrices, and averaging them across pieces in the corpora. This tells about what the particular diatonicisms (and only the diatonicisms) are made of. Some relevant differences stand out from the comparison. Victoria's larger usage of major and minor triads (3-11) and cadential chord sequences (5-27) stands out. On the other hand, Bach makes more prominent usage of the scalar formation 6-Z25: aside its instantiations as perfect cadences, it is recurrent in many motivic progressions, which are not idiomatic in Victoria's contrapuntal writing.

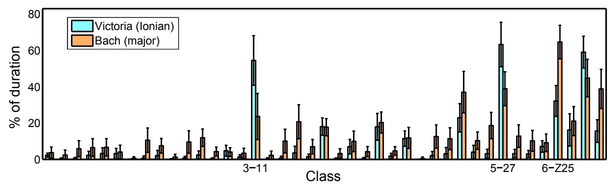


Figure 2: Diatonicism in Victoria and Bach. Mean subclass-vectors under 7-35.

4.2 Case study: structural analysis

Self-similarity matrices (SSM) are a simple standard tool used for structural analysis [3]. Classical inputs to the SSM are spectral or chroma feature time series. Some of the SSM-based methods can handle different time-scales, and some of the chroma methods allows transpositional invariance [9]. These functionalities are usually implemented at the SSM computation stage, or as a post processing. In the class-matrices, both the equivalence mappings (including their inherent hierarchies) and the multi-scale na-

[4] Including *Alma Redemptoris Mater*, *Ave Regina Caelorum*, *Laetatus Sum*, *Pro Victoria*, *Quam Pulchri Sunt*, and *Trahe Me Post Te*. See (Rive, 1969) for a modal classification.

ture of the information are *already* embedded in the feature time-series, so a plain SSM can be used for finding sophisticated recurrences. For instance, a passage comprised of a chord sequence can be recognized as similar than a restated passage with different arpeggiations and/or inversions of the chord intervals (e.g. from major to minor triads). A *vertical* chord and its arpeggiated version may not be recognized as very similar at the lowest cardinalities, but their common T_nI-sonority will certainly do at their corresponding time-scales. Moreover, any sonority containing the chords (*supersets*) will also be captured at their proper time-scales, climbing up the hierarchy until reaching the whole-piece segment, everything indexed by a common temporal axis. A quantification of similarity between variations may thus be possible at the level of embedded sonorities.

This is discussed next for large-scale recurrence finding in Webern's *Variations for piano*, op.27/I. This serial piece presents an *A*-*B*-*A*' structure, built upon several instantiations of the main twelve-tone row, at different transpositional and/or inversional levels. Figure 3 (top) depicts the class-scape of the piece, filtered by the prominent hexachordal iv-sonority ⟨332232⟩, and Figure 3 (bottom) shows the well-known (extensively analysed in literature) structure of the row instantiations, annotated according to [2]. Figure 4 depicts the output of a plain SSM, computed from three different inputs: a) the pc-set time series;[5] b) the class-matrix under T_n; c) the class-matrix under T_nI. The pc-equivalence does not capture any large-scale recurrence. The restatement of the first two phrases in A is captured by the T_n-equivalence, as these phrases are mainly related by transposition in A'. Finally, the T_nI-equivalence reveals the complete recapitulation, including the last two phrases of A, which are restated in A' in both transposed and inverted transformations. It is worth noting that the method does not limit to compare the general sonority, the ubiquitous ⟨332232⟩, but its specific construction down the subclass hierarchy. This allows the discrimination of the *B* section, built upon the same kind of row instantiations than A and A', but presented in distinct harmonisations.

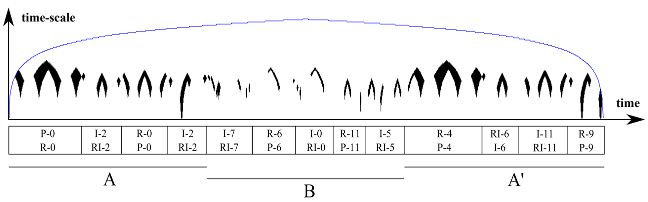
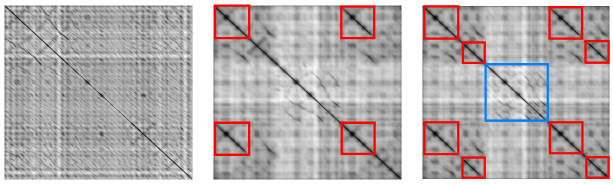


Figure 3: Webern's op.27/I. Top: class-scape filtered by ⟨332232⟩; Bottom: structure.

A relevant advantage of the pc-set-based spaces, with respect to *continuous* ones,[6] is that music can be analysed in terms of different class systems at no extra computational cost. Being *finite and discrete* spaces (4096 classes at most for the TET system), the whole equivalence systems, including their inner metrics, can be precomputed.

[5] In some respect, the discrete *equivalent* of the chroma features.

[6] Such as chroma features, a *finite, but continuous* space.



(a) (b) (c)

Figure 4: SSM from Webern's op.27/I. a) pc-equivalence; b) T_n-equivalence); c) T_nI-equivalence).

The mapping from pc-sets to set-classes, as well as the distances between any pair of music segments, can thus be implemented by table indexing. Once the pc-set of each possible segment has been computed (which constitutes the actual bottleneck of the method), the rest of the process is inexpensive, and multiple *set-class lenses* can be changed in real time, allowing fast interactive explorations of the massive data. This feature, alongside with a variety of filtering options for visual exploration, can be tested with our proof-of-concept set-class analysis tool.[7]

5. MINING CLASS-VECTORS

In this section, we will review and elaborate upon the information conveyed by the class-vectors. For each class, the corresponding value in the vector accounts for the relative duration of the piece which is *interpretable* in terms of the specific class, that means, the proportion of time points which are contained in some (at least one) instance of the class. A dataset of class-vectors, thus, can be exploited in a variety of ways. Finding specific sonorities in large datasets can be combined with the extraction of the actual segments from the MIDI files. This can be exploited in varied applications, ranging from corpora analysis to music education.

A dataset of class-vectors was computed from 13480 MIDI tracks, including works by Albéniz, Albinoni, Alkan, Bach, Beethoven, Brahms, Bruckner, Busoni, Buxtehude, Byrd, Chopin, Clementi, Corelli, Couperin, Debussy, Dowland, Frescobaldi, Gesualdo, Guerrero, Haydn, Josquin, Lasso, Liszt, Lully, Mahler, Morales, Mozart, Pachelbel, Palestrina, Satie, Scarlatti, Shostakovich, Schumann, Scriabin, Soler, Stravinsky, Tchaikovsky, Telemann, Victoria and Vivaldi. It also includes anonymous medieval pieces, church hymns, and the Essen folksong collection.

5.1 Case study: query by set-class

A simple but useful application is querying the dataset for a given set-class sonority. It can be used, for instance, to find pieces with a relevant presence of *exotic* scales. Table 1 shows 10 retrieved pieces with a notable presence (relative duration) of the sonority 7-22, usually referred to as the Hungarian minor scale.[8] Both monophonic and polyphonic pieces are retrieved, ranging different styles

[7] See *Supplemental material* (Section 7).

[8] Sometimes also called Persian, major gypsy, or double harmonic scale, among other denominations.

and historic periods, as the unique requisite for capturing a given sonority it its existence as a temporal segment.

retrieved piece	7-22 (%)
Scriabin - Prelude op.33 n.3	68.61
Busoni - 6 etudes op.16 n.4	63.22
Essen - 6478	62.50
Liszt - Nuages gris	42.41
Essen - 531	36.67
Scriabin - Prelude op.51 n.2	31.74
Lully - Persee act-iv-scene-iv-28	29.73
Alkan - Esquisses op.63 n.19	28.87
Satie - Gnossienne n.1	28.15
Scriabin - Mazurka op.3 n.9	24.61

Table 1: Retrieved pieces: 7-22

5.1.1 Query by combined set-classes

The strict separation of classes in the class-vectors, allows the exploration of any class combination, whether common or unusual. For instance, the first movement of Stravinsky's *Symphony of psalms* is retrieved by querying for music containing substantial diatonic (7-35) and octatonic (8-28) material, certainly an uncommon musical combination. The class-vector also reveals the balance between both sonorities, as 30.18 % and 29.25 % of the piece duration, respectively. As discussed in Section 4, the class-matrices allow the hierarchical analysis of specific sonorities. The class-vectors, on the other hand, summarise the information in a way in which it is not possible, in general, to elucidate the subclass content under a given class. However, if the queried sonorities have a substantial presence (or absence) in the piece, the class-vectors alone can often account for some hierarchical evidence. Table 2 shows 10 retrieved pieces, characterised by a notable presence of the so-called *suspended* trichord (3-9), [9] constrained to cases of mostly diatonic contexts (7-35). This situation, as reflected in the results, is likely to be found in medieval melodies, early counterpoint, or works composed as reminiscent of them. It is worth noting that the 3-9 instantiations appear in quite different settings, whether in monophonic voices, as a combination of melody and tonic-dominant drones, and as actual suspended (voiced) chords.

retrieved piece	3-9 (%)	7-35 (%)
Anonymous - Angelus ad virginem 1	56.79	100
Anonymous - Instrumental dances 7	50.41	100
Lully - Persee prologue-3c	47.11	100
Lully - Phaeton acte-i-scene-v	45.95	90.81
Lully - Persee prologue-3b	44.36	100
Anonymous - Ductia	43.82	100
Anonymous - Danse royale	35.58	100
Anonymous - Cantigas de Santa Maria 2	32.06	100
Anonymous - Instrumental dances 9	27.43	100
Frescobaldi - Canzoni da sonare-11	26.69	81.34

Table 2: Retrieved pieces: mostly 7-35 with 3-9.

As non-existing sonorities may also reveal important characteristics of music, the dataset can be queried for combinations of present and absent classes. For instance, the sonority of fully diatonic (7-35) pieces depends on whether they contain major or minor trichords (3-11) or not. Retrieved pieces in the latter case (diatonic, not triadic) are mostly medieval melodies or early polyphonic pieces, prior to the establishment of the triad as a common sonority.

These results point to interesting applications related with music similarity, such as music recommendation and music education. We find of particular interest the potential of retrieving pieces sharing relevant tonal-related properties, but pertaining to different styles, composers, or historical periods. Music similarity is, to a great extent, a human construct, as it depends on cultural factors and musical background. It would thus be possible to *learn* to appreciate non familiar similarity criteria, which could be suggested by music discovery or recommendation systems.

5.2 On dimensionality and informativeness

In feature design, the ratio between the size of the feature space and the informativeness of description is a relevant factor. The class content of a piece, as described by its class-vector, have 200, 223 or 351 dimensions, depending on the chosen equivalence (iv, T_nI or T_n). Compared with other tonal feature spaces, these dimensions may seem quite large. However, the benefits of class vectors are the systematicity, specificity and precision of the description. Several relevant differences with respect to other tonal-related features are to be noticed. A single class-vector, computed after a fully systematic segmentation, accounts for:

1. Every different segment in the piece, regardless of their time position or duration. No sampling artefacts of any kind are introduced.

2. Every possible sonority among the set-class space, which is *complete*. Every instantiation of every class is captured and represented.

3. An objective and precise description of the *set-class sonority*. No probabilities or estimations are involved.

4. A description in (high level) music theoretical terms, readable and interpretable by humans. Set-classes constitute a standard lexicon in music analysis.

5. An objective quantification of every possible sonority in terms of relative duration in the piece. No probabilities or estimations are involved.

6. A content-based, model-free, description of the piece. Neither statistics nor properties learned from datasets are involved.

7. In cases of large presence or notable absences of some sonorities, an approximation to the hierarchical inclusion relations (fully available through the class-matrices only).

In contrast, the most common tonal piecewise and labelwise feature (global key estimation) conveys:

[9] A major trichord with the third degree substituted by the fourth.

1. A single label for the whole piece, often misleading for music which modulates.

2. 24 different labels, but actually two different sonorities (major and minor), non representative of a vast amount of music.

3. An *estimation* of the key: not only because of the inherent ambiguity of tonality, but also because the (most often) limited tonal *knowledge* of the algorithms.

4. A description in (high level) music theoretical terms, but conveying very little musical information (e.g. at compositional level).

5. No quantification, just a global label. At most, including an indicator of *confidence* (in the descriptor terms), usually the key strength.

6. A description based on specific models (e.g. profiling methods or rule-based), which do not generalize. Some models are trained from specific datasets, biasing the actual *meaning* of the descriptor.

7. No access to the (very rich) hierarchical relations of the piece's tonality.

With this in mind, it seems to us that a piecewise description in 200 dimensions is a reasonable trade-off between size and informativeness. Considering the somewhat sophisticated tonal information conveyed by the class-vectors, they may constitute a useful complementary feature for existing content-based metadata.

6. CONCLUSIONS

The proposed systematic methodology for multi-scale set-class analysis is purposeful for common music information retrieval applications. An appropriate mining of the class-matrices can bring insights about the hierarchical relations among the sets, informing about the specific construction of the class sonorities. In combination with simple recurrence finding methods, the class-matrices can be used for music structure analysis of complex music, beyond the scope of mainstream tonal features. The proposed class-vectors, as piecewise tonal summaries, convey a rich information in terms of every possible class sonority. They can be mined for querying tasks of some sophistication. Their compromise between dimensionality and informativeness, point to potential advances in music similarity and recommendation applications. The examples in this work show that set-classes can inform about very different music compositions, ranging simple folk tunes, early polyphony, common-practice period, *exotic* or uncommon scales, and atonal music. Besides our ongoing musicological analyses, and current research with chroma-based transcriptions from audio, future work may explore the potential of these methods in actual classification and recommendation systems.

7. SUPPLEMENTAL MATERIAL

The interactive potential of the methods discussed in this work can be tested by our multi-scale set-class analysis prototype for Matlab, freely available from http://agustin-martorell.weebly.com/set-class-analysis.html. A comprehensive table of set-classes, and a growing dataset of class-vectors, are also available at this site.

8. ACKNOWLEDGEMENTS

This work was supported by the EU 7th Framework Programme FP7/2007-2013 through PHENICX project [grant no. 601166].

9. REFERENCES

[1] M. Babbit: "Some Aspects of Twelve-Tone Composition," *The Score and I.M.A. Magazine*, Vol. 12, pp. 53–61, 1955.

[2] N. Cook: *A Guide to Music Analysis*, J. M. Dent and Sons, London, 1987.

[3] J. Foote: "Visualizing Music and Audio Using Self-Similarity," *Proceedings of the ACM Multimedia*, pp. 77–80, 1999.

[4] A. Forte: "A Theory of Set-Complexes for Music," *Journal of Music Theory*, Vol. 8, No. 2 pp. 136–183, 1964.

[5] H. Hanson: *The Harmonic Materials of Modern Music: Resources of the Tempered Scale*, Appleton-Century-Crofts, New York, 1960.

[6] E. Huovinen and A. Tenkanen: "Bird's-Eye Views of the Musical Surface: Methods for Systematic Pitch-Class Set Analysis," *Music Analysis*, Vol. 26, No. 1–2 pp. 159–214, 2007.

[7] D. Lewin: "Re: Intervallic Relations between Two Collections of Notes," *Journal of Music Theory*, Vol. 3, No. 2 pp. 298–301, 1959.

[8] A. Martorell and E. Gómez: "Hierarchical Multi-Scale Set-Class Analysis," *Journal of Mathematics and Music*, Online pp. 1-14, 2014.

[9] M. Müller: *Information Retrieval for Music and Motion*, Springer, Berlin, 2007.

[10] T. N. Rive: "An Examination of Victoria's Technique of Adaptation and Reworking in his Parody Masses - with Particular Attention to Harmonic and Cadential Procedure," *Anuario Musical*, Vol. 24, pp. 133–152, 1969.

[11] C. S. Sapp: "Visual Hierarchical Key Analysis," *Computers in Entertainment*, Vol. 4, No. 4 pp. 1–19, 2005.

[12] J. N. Straus: *Introduction to Post-Tonal Theory*, Prentice-Hall, Upper Saddle River, NJ, 1990.