TOWARDS MODELING TEXTURE IN SYMBOLIC DATA

Mathieu Giraud
LIFL, CNRS
Univ. Lille 1, Lille 3

Florence Levé
MIS, UPJV, Amiens
LIFL, Univ. Lille 1

Florent Mercier
Univ. Lille 1

Marc Rigaudière
Univ. Lorraine

Donatien Thorez
Univ. Lille 1

{mathieu, florence, florent, marc, donatien}@algomus.fr

ABSTRACT

Studying *texture* is a part of many musicological analyses. The change of texture plays an important role in the cognition of musical structures. Texture is a feature commonly used to analyze musical audio data, but it is rarely taken into account in symbolic studies. We propose to formalize the texture in classical Western instrumental music as melody and accompaniment layers, and provide an algorithm able to detect homorhythmic layers in polyphonic data where voices are not separated. We present an evaluation of these methods for parallel motions against a ground truth analysis of ten instrumental pieces, including the first movements of the six quatuors op. 33 by Haydn.

1. INTRODUCTION

1.1 Musical Texture

According to Grove Music Online, texture refers to *the sound aspects of a musical structure*. One usually differentiates *homophonic* textures (rhythmically similar parts) and *polyphonic* textures (different layers, for example melody with accompaniment or countrapuntal parts). Some more precise categorizations have been proposed, for example by Rowell [17, p. 158 – 161] who proposes eight "textural values": orientation (vertical / horizontal), tangle (interweaving of melodies), figuration (organization of music in patterns), focus vs. interplay, economy vs. saturation, thin vs. dense, smooth vs. rough, and simple vs. complex. What is often interesting for the musical discourse is the *change of texture*: J. Dunsby, recalling the natural tendency to consider a great number of categories, asserts that *"one has nothing much to say at all about texture as such, since all depends on what is being compared with what"* [5].

Orchestral texture. The term *texture* is used to describe orchestration, that is the way musical material is layed out on different instruments or sections, taking into account registers and timbres. In his 1955 *Orchestration* book, W. Piston presents seven types of texture: orchestral unison, melody and accompaniment, secondary melody, part writing, contrapuntal texture, chords, and complex textures [15].

In 1960, Q. R. Nordgren [13] asks: *"Is it possible to measure texture?"*. He proposes to quantify the horizontal and vertical relationships of sounds making up the texture beyond the usual homophonic/polyphonic or light/heavy categories. He considers eight features, giving them numerical values: the number of instruments, their range, their register and their spacing, the proportion and register of gap, and doubling concentrations with their register. He then analyzes eight symphonies by Beethoven, Mendelssohn, Schumann and Brahms with these criteria, finding characteristic differences between those composers.

Non-orchestral texture. However, the term texture also relates to music produced by a smaller group of instruments, even of same timbre (such as a string quartet), or to music produced by a unique polyphonic instrument such as the piano or the guitar. As an extreme point of view, one can consider texture on a monophonic instrument: a simple monophonic sequence of notes can sound as a melody, but also can figure accompaniment patterns such as arpeggiated chords or Alberti bass.

Texture in musical analysis. Studying texture is a part of any analysis, even if texture often does not make sense on its own. As stated by J. Levy, *"although it cannot exist independently, texture can make the functional and sign relationships created by the other variables more evident and fully effective"* [10]. Texture plays a significant role in the cognition of musical structures. J. Dunsby attributes two main roles to texture: the illusion it creates and the expectation it arouses from the listeners towards familiar textures [5]. J. Levy shows with many examples how texture can be a sign in Classic and Early Romantic music, describing the role of accompaniment patterns, solos and unison to raise the attention of the listener before important structural changes [10].

1.2 Texture and Music Information Retrieval

Texture was often not as deeply analyzed and formalized as other parameters (especially melody or harmony). In the field of Music Information Retrieval (MIR), the notion of texture is often used in audio analysis, reduced to timbral description. Any method dealing with audio signals is somewhat dealing with timbre and texture [3, 9]. Based on audio texture, there were for example studies on segmentation. More generally, the term "sound texture" can be used to describe or synthesize non-instrumental audio signals, such as ambient sounds [18, 19].

Among the studies analyzing scores represented by symbolic data, few of them take texture into account. In 1989, D. Huron [7] explains that the three common meanings about the texture term are the volume, the diversity of elements used and the "surface" description, the first two being more easily formalizable. Using a two-dimensional space based on onset synchronization and similar pitch motion, he was able to capture four broad categories of textures: monophony, homophony, polyphony and heterophony. He found also that different musical genres occupy a different region of the defined space.

Some of the features of the jSymbolic library, used for classification of MIDI files, concern musical texture [11, 12]. *"[They] relate specifically to the number of independent voices in a piece and how these voices relate to one another."* [11, p. 209]. The features are computed on MIDI files where voices are separated, and include statistical features on choral or orchestral music organization: maximum, average and variability of the number of notes, variability between features of individual voices (number of notes, duration, dynamics, melodic leaps, range), features of the loudest voice, highest and lowest line, simultaneity, voice overlap, parallel motion and pitch separation between voices.

More recently, Tenkanen and Gualda [20] detect articulative boundaries in a musical piece using six features including pitch-class sets and onset density ratios. D. Rafailidis and his colleagues segment the score in several textural streams, based on pitch and time proximity rules [2,16].

1.3 Contents

As we saw above, there are not many studies on modeling or automatic analysis of texture. Even if describing musical texture could be done on a local level of a score, it requires some high-level musical understanding. We thus think that it is a natural challenge, both for music modeling and for MIR studies.

In this paper, we propose some steps towards the modeling and the computational analysis of texture in Western classical instrumental music. We choose here not to take into account orchestration parameters, but to focus on textural features given by local note configurations, taking into account the way these may be split into several layers. For the same reason, we do not look at harmony or at motives, phrases, or pattern large-scale repetition.

The following section presents a formal modeling of the texture and a ground truth analysis of first movements of ten string quartets. Then we propose an algorithm discovering texture elements in polyphonic scores where voices are not separated, and finally we present an evaluation of this algorithm and a discussion on the results.

2. FORMALIZATION OF TEXTURE

2.1 Modeling Texture as Layers

We choose to model the texture, by grouping notes into sets of *"layers"*, also called *"streams"*, sounding as a whole grouped by perceptual characteristics. Auditory stream segregation was introduced by Bregman, who studied many parameters influencing this segregation [1]. Focusing on the information contained on a symbolic score, notes can be grouped in such layers using perceptual rules [4, 16]. The number of layers is not directly the number of actual (monophonic) voices played by the instruments. For instance, in a string quartet where all instruments are playing, there can be as few as only one perceived layer, several voices blending in homorhythmy. On the contrary, some figured patterns in a unique voice can be perceived as several layers, as in a Alberti bass.

More precisely, *we model the texture in layers according to two complementary views*. First, we consider two main roles for the layers, that is how they are perceived by the listeners: *melodic* (mel) layers (dominated by contiguous pitch motion), and *accompaniment* (acc) layers (dominated by harmony and/or rhythm). Second, we describe how each layer is composed.

- A melodic layer can be either a monophonic voice (solo), or two or more monophonic voices in homorhythmy (h), or within a tighter relation, such as (from most generic to most similar) parallel motion (p), octave (o) or unison (u) doubling. The h/p/o/u relations do not need to be exact: for example, a parallel motion can be partly in thirds, partly in sixths, and include some foreign notes (see Figure 1).

- An accompaniment layer can also be described by h/p/o/u relations, but it is often worth focusing on its rhythmic component: for example, such a layer can contain sustained, sparse or repeated chords, Alberti bass, pedal notes, or syncopation.

The usual texture categories can then be described as:

- **mel/acc** – the usual accompanied melody;

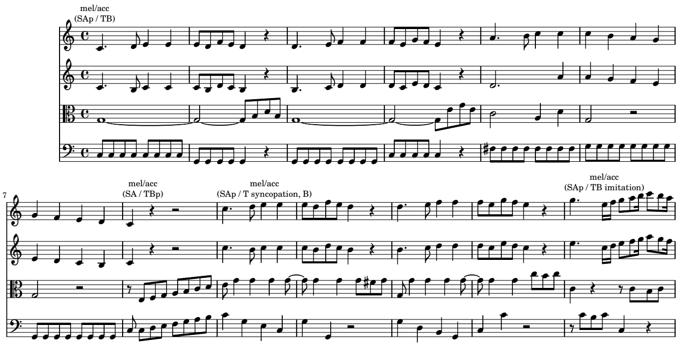
- **mel/mel** – two independent melodies (counterpoint, imitation...);

- **mel** – one melody (either solo, or several voices in h/p/o/u relation), no accompaniment;

- **acc** – only accompaniment, when there is no noticeable melody that can be heard (as in some transitions for example).

The formalism also enables to describe more layers, such as mel/mel/mel/mel, acc/acc, or mel/acc/acc.

Limitations. This modeling of texture is often ambiguous, and has limitations. The distinction between melody and accompaniment is questionable. Some melodies can contain repeated notes, arpeggiated motives, and strongly imply some harmony. Limiting the role of the accompaniment to harmony and rhythm is also over-simplified. Moreover, some textural gestures are not modeled here, such as upwards or downwards scales. Finally, what Piston calls "complex textures" (and what is perhaps the most interesting), interleaving different layers [15, p. 405], can not



```
 1,  75 : mel/acc (SAp / TB)
 8,  82 : mel/acc (SA / TBp)
 9,  83 : mel/acc (SAp / T syncopation, B)
13,  87 : mel/acc (SAp / TB imitation)
19,  93 : mel/acc (SAo / TBh)
20,  94 : mel/acc (SAp / TBhr)
21,  95 : imitation/acc (SA / TB)
25,  99 : imitation/acc (SA / TB)
    102 : mel/acc (S / ATB)
29, 103 : mel/acc (S / ATBhr)
30      : mel/acc (S / ATBh)
    104 : mel/acc (SAh/ TBh)
31, 105 : mel/acc (S / ATBh sparse chords)
33, 107 : acc/mel/acc (S / ATp / B)
  ...
```

Figure 1. Beginning of the string quartet K. 157 no. 4 by W. A. Mozart, with the ground truth analysis describing textures. We label as S / A / T / B (sopran / alt / tenor / bass) the four instruments (violin I / violin II / viola / cello). The first eight measures have a melodic layer "SAp" made by a parallel motion (with thirds), however the parallel motion has some exceptions (unison on c, strong beat on measures 1 and 8, and small interruption at the beginning of measure 5).

always be modeled by this way. Nevertheless, the above formalization is founded for most music of the Classical and of the Romantic period, and corresponds to a way of melody/accompaniment writing.

2.2 A Ground Truth for Texture

We manually analyzed the texture on 10 first movements of string quartets: the six quartets op. 33 by Haydn, three early quartets by Mozart (K. 80 no. 1, K. 155 no. 2 and K. 157 no. 4), and the quartet op. 125 no. 1 by Schubert. These pieces covered the textural features we wanted to elucidate. We segmented each piece into non-overlapping segments based only on texture information, using the formalism described above.

It is difficult to agree on the signifiance on short segments and on their boundaries. Here we choose to report the texture with a resolution of *one measure*: we consider only segments during at least one measure (or filling the most part of the measure), and round the boundaries of these segments to bar lines.

We identified 691 segments in the ten pieces, and Table 1 details the repartition of these segments. The ground truth file is available at www.algomus.fr/truth, and Figure 1 shows the analysis for the beginning of the string quartet K. 157 no. 4 by Mozart.

The segments are further precised by the involved voices and the h/p/o/u relations. For example, focusing on the most represented category "mel/acc", there are 254 segments labelled either "S / ATB" or "S / ATBh" (melodic layer at the first violin) and 81 segments labelled "SAp / TB" or "SAp / TBh" (melodic layer at the two violins, in a parallel move). Note that h/p/o/u relations were evaluated here in a subjective way. The segments may contain some small interruptions that do not alter the general perception of the h/p/o/u relation.

3. DISCOVERING SYNCHRONIZED LAYERS

We now try to provide a computational analysis of texture starting from a polyphonic score where voices are not separated. A first idea is to *first segment the score into layers by perception principles, and then to try to qualify some of these layers.* One can for example use the algorithm of [16] to segment the musical pieces into layers (called "streams"). This algorithm relies on a distance matrix, which tells for each possible pair of notes whether they are likely to belong to the same layer. The distance between two notes is computed according to their synchronicity, pitch and onset proximity (among others criteria); then for each note, the list of its k-nearest neighbors is established. Finally, notes are gathered in clusters. A melodic stream can be split into several small chunks, since the diversity of melodies does not always ensure coherency within clusters; working on larger layers encompass them all. Even if this approach produces good results in segmentation, many layers are still too scattered to be detected as full melodic or accompaniment layers. Nonetheless, classification algorithms could label some of these layers as melodies or accompaniments, or even detect the type of the accompaniment.

The second idea, that we will develop in this paper, is to *detect directly noteworthy layers from the polyphonic data*. Here, we choose to focus on perceptually significant relations based on homorhythmic features. The following paragraphs define the notion of *synchronized layers*, that is sequences of notes related by some homorhythmy relation (h/p/o/u), and show how to compute them.

3.1 Definitions: Synchronized Layers

A *note* n is given as a triplet $(n.pitch, n.start, n.end)$, where $n.pitch$ belongs to a pitch scale (that can be defined diatonically or by semitones), and $n.start$ and $n.end$ are two positions with $n.start < n.end$. Two notes n and m are *synchronized* (denoted by $n \equiv_h m$) if they have the same start and the same end.

A *synchronized layer* (SL) is a set of two sequences of consecutive synchronized notes (in other words, these sequences correspond to two "voices" in homorhythmy). Formally, two sequences of notes $n_1, n_2 ... n_k$ and $m_1, m_2 ... m_k$ form a synchronized layer when:

- for all i in $\{1, \ldots, k\}$, $\quad n_i.start = m_i.start$

	tonality	length	mel/acc	mel/mel	acc/mel/acc	acc/mel	mel	acc	others	h	p	o	u
Haydn op. 33 no. 1	B minor	91m	38	0	8	1	0	0	5	19	21	1	0
Haydn op. 33 no. 2	E-flat major	95m	37	0	2	4	0	0	7	34	13	0	0
Haydn op. 33 no. 3	C major	172m	68	0	0	0	3	13	6	29	50	1	0
Haydn op. 33 no. 4	B-flat major	90m	25	0	1	0	0	0	6	16	6	0	0
Haydn op. 33 no. 5	G major	305m	68	0	3	4	7	0	5	56	45	6	0
Haydn op. 33 no. 6	D major	168m	58	0	1	3	15	0	29	43	42	0	2
Mozart K. 80 no. 1	G major	67m	36	4	6	0	2	0	3	5	33	3	0
Mozart K. 155 no. 2	D major	119m	51	0	0	0	1	0	0	21	32	4	1
Mozart K. 157 no. 4	C major	126m	29	0	3	6	2	0	7	18	22	2	0
Schubert op. 125 no. 1	E-flat major	255m	102	0	0	0	20	2	0	54	8	46	2
		1488m	512	4	24	18	50	15	68	295	272	63	5

Table 1. Number of segments in the ground truth analysis of the ten string quartets (first movements), and number of h/p/o/u labels further describing these layers.

- for all i in $\{1, \ldots, k\}$, $\quad n_i.end = m_i.end$
- for all i in $\{1, \ldots, k-1\}$, $\quad n_i.end = n_{i+1}.start$

This definition can be extended to any number of voices. As p/o/u relations have a strong musical signification, we want to be able to enforce them. One can thus restrain the relation $\equiv_h$, considering the pitch information:

- we denote $n \equiv_\delta m$ if the interval between the two notes n and m is δ. The nature of the interval δ depends on the pitch model: for example, the interval can be diatonic, such as in "third" (minor or major), or an approximation over the semitone information, such as in "3 or 4 semitones". Some synchronized layers with $\equiv_\delta$ relations correspond to parallel motions;
- we denote $n \equiv_o m$ if notes n and m are separated by any number of octaves;
- we denote $n \equiv_u m$ where there is an exact equality of pitches (unison).

Given a relation $\equiv \in \{\equiv_h, \equiv_\delta, \equiv_o, \equiv_u\}$, we say that a synchronized layer respects the relation $\equiv$ if its notes are pairwise related according to this relation. The relation $\equiv_h$ is an equivalence relation, but the restrained relations do not need to be equivalence relations: Some $\equiv_\delta$ relations are not transitive.

For example, in Figure 1, there is between voices S and A (corresponding to violins I and II), in the first two measures:

- a synchronized layer ($\equiv_h$) on the two measures;
- and a synchronized layer ($\equiv_{\text{third}}$) on the two measures, except the first note.

Note that this does not correspond exactly to the "musical" ground truth (parallel move on at least the first four measures) because of some rests and of the first synchronized notes that are not in thirds.

A synchronized layer is *maximal* if it is not strictly included in another synchronized layer. Note that two maximal synchronized layers can be overlapping, if they are not synchronized. Note also that the number of synchronized layers may grow exponentially with the number of notes.

3.2 Detection of a Unique Synchronized Layer

A very noticeable textural effect is when *all* voices use the same texture at the same time. For example, a sudden striking unison raises the listener's attention. We can first check if all notes in a segment of the score belong to a unique synchronized layer (within some relation). For example, we consider that all voices are in octave doubling or unison if it lasts at least two quarters.

3.3 Detection of Maximal Synchronized Layers

In the general case, the texture has several layers, and the goal is thus to extract layers using *some* of the notes. Remember that we work on files where polyphony is not separated into voices: moreover, it is not always possible to extract voices from a polyphonic score, for example on piano music. We want to extract maximal synchronized layers. However, as their number may grow exponentially with the number of notes, we will compute only the start and end positions of maximal synchronized layers.

The algorithm is a kind of 1-dimension interval chaining [14]. The idea is as follows. Recursively, two voices $n_1, \ldots, n_k$ and $m_1, \ldots, m_k$ are synchronized if and only if $n_1, \ldots, n_{k-1}$ and $m_1, \ldots, m_{k-1}$ are synchronized, n_k and m_k are synchronized and finally $n_{k-1}.end = n_k.start$. Formally, the algorithm is described by the following:

Step 1. *Compute a table with left-maximal SL.* Build the table *leftmost_start*$_\equiv[j]$ containing, for each ending position j, the leftmost starting position of a SL respecting $\equiv$ ending in j. This can be done by dynamic programming with the following recurrence:

$$\textit{leftmost\_start}_\equiv[j] = \begin{cases} \min\{\textit{leftmost\_start}_\equiv[i] \mid i \in S_\equiv(j)\} \\ \quad \text{if } S_\equiv(j) \text{ is not empty} \\ j \quad \text{if } S_\equiv(j) \text{ is empty} \end{cases}$$

where $S_\equiv(j)$ is the set of all starting positions of synchronized notes ending at j respecting the relation $\equiv$:

$$S_\equiv(j) = \left\{ n.start \;\middle|\; \begin{array}{l} \text{there are two different notes } n \equiv m \\ \text{such that } n.end = j \end{array} \right\}$$

Step 2. *Output only (left and right) maximal SL.* Output (i, j) with $i = \textit{leftmost\_start}_\equiv[j]$ for each j, such that $j = \max \{j_o \mid \textit{leftmost\_start}_\equiv[j_o] = \textit{leftmost\_start}_\equiv[j]\}$

```
                  1                   2                   3                   4                   5                   6
1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4
SA- SA-  SA---  SA- SA- SA---- SA-  SB-                 AB             TB       AT        SB---- SA----              AB       AT
              TB              SA  SA                    AB-             SA                AT  SB-
                                    SB-                                                     ST--
          7                   8                   9                   10                  11                  12
5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
    TB              SA- SA-  SA---  SA- SA- SA---- SA-  SB-  TB-  TB        AB              AT      AT        SB ST--------
      TB--                        TB              SA  SA                    AB-                               SB- SB
                                                        SB-                  SA                               ST  AT-
```

Figure 2. Result on the parallel move detection on the first movement of the string quartet K. 157 no. 4 by Mozart. The top lines display the measure numbers. The algorithm detects 52 synchronized layers respecting the $\equiv_p$ relation. 39 of these 52 layers overlap layers identified in the truth with p/o/u relations. The parallel motions are further identified by their voices (S / A / T / B), but this information is not used in the algorithm which works on non-separated polyphony.

The first step is done in $O(nk)$ time, where n is the number of notes and $k \leq n$ the maximal number of simultaneously sounding notes, so in $O(n^2)$ time. The second step is done in $O(n)$ time by browsing from right to left the table *leftmost_start*$_\equiv$, outputing values i when they are seen for the first time.

To actually retrieve intervals, we can store in the table *leftmost_start*$_\equiv[j]$ a pair (i, ℓ), where ℓ is the list of notes/intervals from which the set of SL can be built (this set may be very large, but not ℓ). The time complexity is now $O(n(k + w))$, where $w \leq n$ is the largest possible size of ℓ. Thus the time complexity is still in $O(n^2)$. This allows, in the second step, to filter the SL candidates according to additional criteria on ℓ.

Note finally that the definition of synchronized layer can be extended to include consecutive notes separated with rests. The same algorithm still applies, but the value of k rises to the maximum number of notes that can be linked in that way.

4. RESULTS AND DISCUSSION

We tested the proposed algorithm to look for synchronized layers respecting $\equiv_\delta$ relation (constant pitch interval, including parallel motion) on the ten pieces of our corpus given as `.krn` Humdrum files [8]. Although the pieces are string quartets, we consider them as non-separated polyphonic data, giving as input to the algorithm a single set of notes. The algorithm finds 434 layers. Figure 2 shows an example of the output of the algorithm. Globally, on the corpus, the algorithm labels 797 measures (that is 53.6% of the length) as synchronized layers.

Evaluation against the truth. There are in the truth 354 layers with p/o/u relations: mainly parallel moves, and some octave doubling and unisons. As discussed earlier, these layers reported in the truth correspond to a musical interpretation: they are not as formalized as our definition of synchronized layer. Moreover, less information is provided by the algorithm than in the ground truth: when a parallel motion is found, the algorithm cannot provide at which voice/instrument it appears, since we worked from polyphonic data with no voice separation.

Nevertheless, we compared the layers predicted by the algorithms with the ones of the truth. Results are summarized on Table 2. A computed layer is marked as true positive (TP) as soon as it overlaps a p/o/u layer of the truth.

356 of the 434 computed synchronized layers are overlapping the p/o/u layers of the truth, thus 82.0% of the computed synchronized layers are (at least partially) musically relevant. These 356 layers map to 194 p/o/u layers in the truth (among 340, that is a sensitivity of 58.0%): a majority of the parallel moves described in the truth are found by the algorithm.

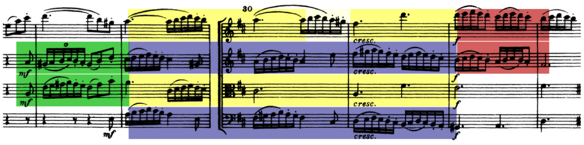


Figure 3. Haydn, op. 33 no. 6, m. 28-33. The truth contains four parallel moves.

Merged parallel moves. If one restricts to layers where borders coincide with the ones in the truth (same start, same end, with a tolerance of 2 quarters), the number of truth layers found falls from 194 to 117. This is because the algorithm often merge consecutive parallel moves. An example of this drawback is depicted on Figure 3. Here a melody is played in imitation, resulting in parallel moves involving all voices in turn. The algorithm detects a unique synchronized layer, which corresponds to a global perception but gives less information about the texture. We should remember here that the algorithm compute boundaries of synchronized layers and not actual instances, which would require some sort of voice separation and possibly generate a large number of instances.

False positives. Only 78 false positives are found by the algorithm. Many false positives (compared to the truth) are parallel moves detected inside a homorhythmy $\equiv_h$ relation between 3 ou 4 voices. In particular, the algorithm detects a parallel move as soon as there are sequences of repeated notes in at least two voices. This is the case in in op. 33 no. 4 by Haydn which contains many homorhythmies in repeated notes, for which we obtain 30 false positives. Even focusing on layers with a real "move", false positive could also appear between a third voice and two voices with repeated notes. Further research should be carried to discard these false positives either in the algorithm or at a later filtering stage.

	hits length	hits	TP	FP	truth-overlap	truth-exact
Haydn op. 33 no. 1	40m (44%)	37	32 (86.5%)	5	14/22 (63.6%)	7/22
Haydn op. 33 no. 2	21m (22%)	17	15 (88.2%)	2	7/13 (53.9%)	7/13
Haydn op. 33 no. 3	73m (42%)	48	44 (91.7%)	4	27/51 (52.9%)	15/51
Haydn op. 33 no. 4	19m (21%)	47	17 (36.2%)	30	5/6 (83.3%)	3/6
Haydn op. 33 no. 5	235m (77%)	58	47 (81.0%)	11	27/51 (52.9%)	11/51
Haydn op. 33 no. 6	63m (37%)	24	21 (87.5%)	3	19/44 (43.2%)	11/44
Mozart K. 80 no. 1	45m (67%)	27	26 (96.3%)	1	20/36 (55.6%)	14/36
Mozart K. 155 no. 2	76m (64%)	46	44 (95.7%)	2	27/37 (73.0%)	15/37
Mozart K. 157 no. 4	62m (49%)	52	39 (75.0%)	13	15/24 (62.5%)	8/24
Schubert op. 125 no. 1	163m (64%)	78	71 (91.0%)	7	33/56 (58.9%)	20/56
	797m (54%)	434	356 (82.0%)	78	194/340 (57.1%)	111/340

Table 2. Evaluation of the algorithm on the ten string quartets of our corpus. The columns TP and FP show respectively the number of true and false positives, when comparing computed parallel moves with the truth. The columns truth-overlap shows the number of truth parallel moves that were matched by this way. The column truth-exact restricts these matchings to computed parallel moves for which borders coincide to the ones in the truth (tolerance: two quarters).

5. CONCLUSION AND PERSPECTIVES

We proposed a formalization of texture in Western classical instrumental music, by describing melodic or accompaniment “layers” with perceptive features (h/p/o/u relations). We provided a first algorithm able to detect some of these layers inside a polyphonic score where tracks are not separated, and tested it on 10 first movements of string quartets. The algorithm detects a large part of the parallel moves found by manual analysis. We believe that other algorithms implementing textural features, beyond h/p/o/u relations, should be designed to improve computational music analysis. The corpus should also be extended, for example with music from other periods or piano scores.

Finally, we believe that this search of texture, combined with other elements such as patterns and harmony, will improve algorithms for music structuration. The ten pieces of our corpus have a *sonata form* structure. The tension created by the exposition and the development is resolved during the recapitulation, and textural elements contribute to this tension and its resolution [10]. For example, the medial caesura (MC), before the beginning of theme S, has strong textural characteristics [6]. Textural elements predicted by algorithms could thus help the structural segmentation.

6. REFERENCES

[1] A. S. Bregman. *Auditory scene analysis*. Bradford, Cambridge, 1990.

[2] E. Cambouropoulos. Voice separation: theoretical, perceptual and computational perspectives. In *Int. Conf. on Music Perception and Cognition (ICMPC)*, 2006.

[3] R. B. Dannenberg and M. Goto. *Handbook of Signal Processing in Acoustics*, chapter Music Structure Analysis, pages 305–331. Springer, 2008.

[4] D. Deutsch and J. Feroe. The internal representation of pitch sequences in tonal music. *Psychological Review*, 88(6):503–522, 1981.

[5] J. M. Dunsby. Considerations of texture. *Music and letters*, 70(1):46–57, 1989.

[6] J. Hepokoski and W. Darcy. The medial caesura and its role in the eighteenth-century sonata exposition. *Music Theory Spectrum*, 19(2):115–154, 1997.

[7] D. Huron. Characterizing musical textures. In *Int. Computer Music Conf. (ICMC)*, pages 131–134, 1989.

[8] D. Huron. Music information processing using the Humdrum toolkit: Concepts, examples, and lessons. *Computer Music J.*, 26(2):11–26, 2002.

[9] A. Klapuri and M. Davy. *Signal Processing Methods for Music Transcription*. Springer, 2006.

[10] J. M. Levy. Texture as a sign in classic and early romantic music. *J. of the American Musicological Society*, 35(3):482–531, 1982.

[11] C. McKay. *Automatic music classification with jMIR*. PhD thesis, McGill University, 2010.

[12] C. McKay and I. Fujinaga. jSymbolic: A feature extractor for MIDI files. In *Int. Computer Music Conf. (ICMC)*, pages 302–305, 2006.

[13] Q. R. Nordgren. A measure of textural patterns and strengths. *J. of Music Theory*, 4(1):19–31, Apr. 1960.

[14] E. Ohlebusch and M. I. Abouelhoda. *Handbook of Computational Molecular Biology*, chapter Chaining Algorithms and Applications in Comparative Genomics. 2005.

[15] W. Piston. *Orchestration*. Norton, 1955.

[16] D. Rafailidis, A. Nanopoulos, Y. Manolopoulos, and E. Cambouropoulos. Detection of stream segments in symbolic musical data. In *Int. Conf. on Music Information Retrieval (ISMIR)*, pages 83–88, 2008.

[17] L. Rowell. *Thinking about Music: An Introduction to the Philosophy of Music*. Univ. of Massachusetts, 1985.

[18] N. Saint-Arnaud and K. Popat. Computational auditory scene analysis. chapter Analysis and Synthesis of Sound Textures, pages 293–308. Erlbaum, 1998.

[19] G. Strobl, G. Eckel, and D. Rocchesso. Sound texture modeling: a survey. In *Sound and Music Computing (SMC)*, 2006.

[20] A. Tenkanen and F. Gualda. Detecting changes in musical texture. In *Int. Workshop on Machine Learning and Music*, 2008.