COSIATEC and SIATECCompress
Meredith, David

Published in:
Music Information Retrieval Evaluation eXchange (MIREX 2013)

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

? Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
? You may not further distribute the material or use it for any profit-making activity or commercial gain
? You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from vbn.aau.dk on: januar 07, 2019
ABSTRACT

Three versions of each of two greedy compression algorithms, COSIATEC and SIATECCOMPRESS, were run on the JKU Patterns Development Database. Each algorithm takes a point-set representation of a piece of music as input and computes a compressed encoding of the piece in the form of a union of translational equivalence classes of maximal translatable patterns. COSIATEC iteratively uses the SIATEC algorithm to strictly partition the input set into the covered sets of a set of MTP TECs. On each iteration, COSIATEC finds the “best” TEC and then removes its covered set from the input dataset. SIATECCOMPRESS runs SIATEC just once to get a list of MTP TECs and then selects a subset of the “best” TECs that is sufficient to cover the input dataset. Both algorithms select TECs primarily on the basis of compression ratio and compactness.

1. INTRODUCTION

In this paper, I present two greedy compression algorithms, COSIATEC and SIATECCOMPRESS, designed specifically to compute structural descriptions (i.e., analyses) of pieces of music. Both algorithms are based on the SIA and SIATEC algorithm described by Meredith, Lemström and Wiggins [8]. Each algorithm takes a point-set representation of a musical piece as input and computes a compact encoding of the piece in the form of a set of translational equivalence classes of maximal translatable patterns. COSIATEC generates a strict partitioning of the input dataset, whereas the sets of pattern occurrences computed by SIATECCOMPRESS may share points (i.e., notes).

Both algorithms are founded on the hypothesis that the best ways of understanding a piece of music are those that are represented by the shortest descriptions of the piece. In other words, they are designed to explore the notion that music analysis is effectively just music compression.

2. USING POINT SETS TO REPRESENT MUSIC

In the algorithms described below, it is assumed that the piece of music to be analysed is represented in the form of a multi-dimensional point set called a dataset, as described by Meredith et al. [8]. Although these algorithms work with datasets of any dimensionality, it will be assumed here that each dataset is a set of two-dimensional points, \((t, p)\), where each point represents a single note or sequence of tied notes whose onset time is \(t\) in tatum times and whose morphetic pitch [6–8] is \(p\). If morphetic pitch information is not available (e.g., because the data is in MIDI format), then (at least for Western tonal music) it can be very reliably computed from chromatic pitch (i.e., MIDI note number) using an algorithm such as PS13s1 [6, 7].

3. MAXIMAL TRANSLATABLE PATTERNS

I shall use the term pattern to refer to any subset of a dataset. Suppose \(D\) is a dataset and \(P_1, P_2 \subseteq D\). The two patterns, \(P_1, P_2\), are said to be translationally equivalent, denoted by \(P_1 \equiv_T P_2\), if and only if there exists a vector \(v\), such that \(P_1\) translated by \(v\) is equal to \(P_2\). That is,

\[ P_1 \equiv_T P_2 \iff (\exists v \mid P_2 = P_1 + v). \]

Given a vector, \(v\), then the maximal translatable pattern (MTP) for \(v\) in the dataset, \(D\), is defined and denoted as follows:

\[ \text{MTP}(v, D) = \{ p \mid p \in D \land p + v \in D \} \]

where \(p + v\) is the point that results when one translates \(p\) by the vector \(v\). In other words, the MTP for a vector \(v\) in a dataset \(D\) is the set of points in \(D\) that can be translated by \(v\) to give other points that are also in \(D\).

The notion that COSIATEC and SIATECCOMPRESS can be used to discover the patterns in a piece of music that an analyst or a listener finds important, is founded upon the hypothesis that these patterns correspond in some way to MTPs in the pitch-time dataset representation of the piece. Meredith et al. [8] describe an algorithm called SIA for discovering all the MTPs in a dataset.

4. TRANSLATIONAL EQUIVALENCE CLASSES

When analysing a piece of music, we typically want to find all the occurrences of an interesting pattern, not just one occurrence. Given a pattern, \(P\), in a dataset, \(D\), the translational equivalence class (TEC) of \(P\) in \(D\) is defined and denoted as follows:

\[ \text{TEC}(P, D) = \{ Q \mid Q \equiv_T P \land Q \subseteq D \} . \]
We can also define the covered set of a TEC, $T$, denoted by $\text{COV}(T)$, to be the union of the occurrences in the TEC. That is,
\begin{equation}
\text{COV}(T) = \bigcup_{P \in T} P.
\end{equation}

Here we will be particularly concerned with \textit{MTP TECs}—that is, the translational equivalence classes of the maximal translatable patterns in a dataset. Meredith \textit{et al.} [8] describe an algorithm called SIATEC that uses TIA to find all the MTPs and then goes on to find the TEC of each of these MTPs (i.e., it finds all the (exact) occurrences of all the MTPs).

A TEC is a set of patterns that are all translationally equivalent to each other. Suppose a TEC, $T$, contains $n$ occurrences of a pattern containing $m$ points. There are at least two ways in which one can specify $T$. First, one can list each of the $n$ occurrences in $T$ explicitly by listing all of the $m$ points in each occurrence. This requires one to write down $mn$ 2-dimensional points or $2mn$ integers. Alternatively, one can explicitly list the $m$ points in just one of the $n$ occurrences, $P$, and then give the $n-1$ vectors required to map $P$ onto the other occurrences. This requires one to write down $m$ 2-dimensional points and $n-1$, 2-dimensional vectors—that is, $2(m+n-1)$ integers. If $n$ and $m$ are both greater than one, then $2(m+n-1)$ is less than $2mn$, implying that the second method of specifying a TEC gives us a compressed encoding of the TEC (and therefore also of its covered set). Thus, in principle, if a dataset contains repeated (i.e., translationally equivalent) patterns, it may be possible to encode the dataset in a compact manner by representing it as the union of the covered sets of a set of TECs, where each TEC, $T$, is encoded as an ordered pair, $(P,V)$, where $P$ is one occurrence in $T$ and $V$ is the set of vectors that map $P$ onto the other occurrences in $T$. When a TEC, $T = (P,V)$, is represented in this way, we call $P$ the pattern and $V$ the translator set of the TEC.

5. THE COSIATEC ALGORITHM

COSIATEC [5, 9] (see Figure 1) is a greedy compression algorithm, based on SIATEC, that takes a dataset, $D$, as input and computes a compressed encoding of $D$ in the form of an ordered set of MTP TECs, $T$, such that
\begin{equation}
D = \bigcup_{T \in T} \text{COV}(T)
\end{equation}
and, for all $T_1, T_2 \in T$, $T_1 \neq T_2$,
\begin{equation}
\text{COV}(T_1) \cap \text{COV}(T_2) = \emptyset.
\end{equation}

In other words, COSIATEC partitions a dataset $D$ into the covered sets of a set of MTP TECs. If each of these MTP TECs is represented as a \{pattern, translator set\} pair, then this description of the dataset as a set of TECs is typically shorter than an \textit{in extenso} description in which the points in the dataset are simply listed explicitly.

COSIATEC begins by making a copy of the input dataset which it stores in the variable $P$ (line 1). Then, on each iteration of the \texttt{while} loop (lines 4–7), the algorithm finds the “best” MTP TEC in $P$, $T^*$, appends this TEC to $T$ and then removes the set of points covered by $T^*$ from $P$. When $P$ is empty, the algorithm terminates, returning the list of MTP TECs, $T$. The sum of the number of translators and the number of points in this output encoding is never more than the number of points in the input dataset and can be much less than this if there are many repeated patterns in the input dataset.

Given an input dataset, $D$, and what remains of a copy, $P$, of this dataset after the removal of zero or more MTP TEC covered sets, the COSIATEC algorithm finds the “best” MTP TEC in $P$ (line 5), using the \texttt{GetBestTEC} algorithm shown in Figure 2. In lines 1–2 of \texttt{GetBestTEC}, the SIATEC algorithm is used to find all the MTPs in the dataset. The first step in this process is to compute a so-called vector table, $V$, which is a two-dimensional array of ordered triples,
\begin{equation}
V[i][j] = (p_i - p_j, p_i, j)
\end{equation}
where $p_i - p_j$ is the vector from point $p_j$ to $p_i$ and $p_k = P[k]$, where $P$ is an ordered set that only contains every element in $P$, sorted into lexicographical order.

Having computed the vector table, $V$, the MTPs are found by sorting the triples in $V$, lexicographically by their vectors (i.e., their first elements), and then scanning this sorted list once: each MTP is then equal to the points associated with a run of consecutive triples with the same vector in this sorted list. This is accomplished in line 2 of \texttt{GetBestTEC} using the \texttt{ComputeMTPCISPairs} algorithm, which is shown in Figure 3.
The ` COMPUTEMtpCISPairs(V)` algorithm (Figure 3) first sorts the triples in the vector table, V, into increasing lexicographical order by their vectors. The resulting ordered set of triples is stored in the variable W (see line 1). In lines 2–19 of this algorithm, two lists are constructed, MTPs and CISs. MTPs contains all the MTPs in the dataset, each MTP being represented as an ordered set of points in lexicographical order. CISs contains, for each MTP, a list of the indices of the columns in the vector table corresponding to the points in the MTP. In lines 20–22 of ` COMPUTEMtpCISPairs`, a list of \( \langle mtp, cis \rangle \) pairs is constructed by combining corresponding elements in MTPs and CISs.

In lines 5–14 of ` GETBESTTEC`, the `for` loop iterates over this ordered set of \( \langle mtp, cis \rangle \) pairs computed by ` COMPUTEMtpCISPairs`. For each pair, the TEC of the MTP is computed in line 7 using the technique employed in the SIATEC algorithm [8]. Then, in line 8, the conjugate TEC [1] is computed for each MTP TEC found in line 7.

The concept of a conjugate TEC is illustrated in Figure 4. Given a TEC, \( T = \langle P, V \rangle \), the conjugate of \( T \) is denoted and defined as follows:

\[
\text{CONJ}(T) = \langle P', V' \rangle
\]

and defined as follows:

\[
\text{CONJ}(T) = \langle P', V' \rangle
\]

where, if \( p_0 \) is the lexicographically first point in \( P \),

\[
P' = \{ p_0 \} \cup \{ p_0 + v \mid v \in V \},
\]

and

\[
V' = \{ p - p_0 \mid p \in P \} \setminus \{ (0,0) \}.
\]

Given a pair of conjugate TECs, one may be “better” than the other (e.g., because its pattern might be more compact).

In lines 9 and 10 of ` GETBESTTEC`, redundant translators are removed from both the TEC, \( T \), and its conjugate using the ` REMREDTRAN` algorithm. A translator is defined to be `redundant` if it can be removed from the translator set of a TEC without changing the covered set of the TEC. Ideally, in order to get the most compact description of the covered set of a TEC, one would want to remove as many redundant translators as possible. However, in general, finding the smallest subset of the translator set of a TEC that is sufficient to generate the TEC’s covered set is an NP-hard problem. In the implementation of COSIATEC submitted to the MIREX 2013 competition, a greedy approximation algorithm is used to remove as many redundant translators as possible from a TEC within a reasonable running time.

Finally, in lines 11–14 of ` GETBESTTEC`, each MTP TEC and its conjugate are compared with the “best” TEC so far and replace it if they are deemed superior to it by the ` ISBESTTEC` function, defined in Figure 5. This function takes two TECs as its arguments and returns true if the first is “better than” the second. In lines 1–2 of ` ISBESTTEC`, the compression ratio of the two TECs are compared. If \( P(T) \) and \( V(T) \) are defined to return the pattern and translator set of a TEC, \( T \), respectively, then the compression ratio of a TEC is defined as follows:

\[
\text{COMPRATIO}(T) = \frac{|\text{COV}(T)|}{|P(T)| + |V(T)| - 1}.
\]

If the two TECs to be compared have the same compression ratio, then they are compared for bounding-box compactness (lines 3–4 of `ISBESTTEC`) [8]. The bounding-box compactness of a TEC is the number of points in the TEC’s pattern divided by the number of dataset points in the bounding box of this pattern. If the two TECs have the same compression ratio and compactness, the TEC with the largest covered set is considered superior (lines 5–6). If the two covered sets are also the same size, then the TEC with the larger pattern is considered superior (lines 7–8). If the patterns are also the same size, then the TEC with the pattern that has the shorter temporal duration is considered superior (lines 9–10). Finally, if the two TECs also have the same temporal duration, then the TEC with the pattern whose bounding box has the smaller area is considered superior (lines 11–12).

6. THE SIATEC COMPRESS ALGORITHM

COSIATEC runs SIATEC on each iteration of its `while` loop. Since SIATEC has worst case running time \( O(n^3) \)
where $n$ is the number of points in the input dataset, running COSIATEC on large datasets can be time-consuming (see Table 1 for some example running times). On the other hand, because COSIATEC strictly partitions the dataset into non-overlapping MTP TEC covered sets, it tends to achieve high compression ratios for many point-set representations of musical pieces (typically between 2 and 4 for a piece of classical or baroque music).

Like COSIATEC, the SIATECCOMPRESS algorithm shown in Figure 6 is a greedy compression algorithm based on SIATEC that computes an encoding of a dataset in the form of a union of TECs. SIATECCOMPRESS closely resembles the algorithm described by Forth [3,4], but is simpler and non-parametric. Like Forth’s algorithm, but unlike COSIATEC, SIATECCOMPRESS runs SIATEC only once to get a list of TECs in decreasing order of quality (as defined by the ISBETTERTEC function in Figure 5). It then works its way down this list, selecting TECs to include in the encoding, until the input dataset is covered. SIATECCOMPRESS does not generally produce as compact an encoding as COSIATEC, since the TECs in its output may share points. However, it is faster than COSIATEC and can therefore be used practically on much larger datasets.

The first steps in SIATECCOMPRESS are to compute a vector table and compute MTPs using the SIATCOMPRESS function (Figure 7), which computes a compact encoding of the input dataset.

### 7. RESULTS

Three versions of each of the two algorithms described above were run on the JKU Patterns Development Database (JKU PDD) [2]. The results are shown in Table 1. The values in the table were computed using Tom Collins’ MATLAB implementation of the metrics defined in [2], bundled with the JKU PDD.

Each row in Table 1 gives the results of running one version of an algorithm on one of the five pieces in the JKU PDD. The first column gives the name of the algorithm. Each name either has no suffix (e.g., “COSIATEC”, “SIATECCOMPRESS”) or one of the two suffixes, “BB” or “Segment”. A name with no suffix indicates that the row shows the results of running the plain algorithm as described above, with the discovered patterns equal to the MTP TECs in the output encoding. A name with the suffix “BB” indicates that each occurrence within a TEC in the output of the algorithm is replaced with the set of dataset points in the bounding-box of the occurrence. A name with the suffix “Segment” indicates that each occurrence within a TEC in the output of the algorithm is replaced with the set of dataset points in the time segment spanned by the occurrence.

The following preliminary observations can be made from studying these results:

1. The algorithms generally score better on establishment recall than establishment precision; whereas each algorithm’s occurrence recall and occurrence...
<table>
<thead>
<tr>
<th>Algorithms</th>
<th>Phase</th>
<th>Mono/Policy/Appendix</th>
<th>T</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>COSIATEC</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>SIATECCOMPRESS</td>
<td>bash/wg2/Duo</td>
<td>Mono</td>
<td>721</td>
<td>50.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>
| Table 1. Results of running COSIATEC and SIATECCOMPRESS on the JKU Patterns Development Database.
precision scores tend to be more similar to each other.

2. The algorithms score higher on occurrence measures than establishment measures.

3. On “three-layer” measures (P\textsubscript{3}, R\textsubscript{3} and TLF\textsubscript{1}), the algorithms generally score better on recall than precision.

4. SIATECCOMPRESS is 5–10 times faster than COSIATEC and clearly has a lower order of growth with respect to input size. A more detailed analysis of runtime will be given in a later paper.

5. The highest establishment \( F_1 \) score of 0.78 was obtained using SIATECCompressSegment on the Beethoven Sonata movement.

6. The highest occurrence \( F_1 \) score with \( c = 0.75 \) of 0.94 was obtained using COSIATEC on the monophonic Mozart Sonata movement. On this movement, the algorithm also achieved occurrence precision and occurrence recall of 0.94.

7. The highest values of the “three-layer” \( F_1 \) score (0.62–0.65) were obtained using SIATECCompressSegment on the Beethoven Sonata movement (similar values were obtained for both the monophonic and polyphonic versions).

8. The highest values of the occurrence \( F_1 \) score with \( c = 0.5 \) were 0.85–0.87 obtained using COSIATEC and COSIATECBB on the monophonic version of the Mozart Sonata movement.

9. REFERENCES


