

SOLVING THE (δ, α) -APPROXIMATE MATCHING PROBLEM UNDER TRANSPOSITION INVARIANCE IN MUSICAL SEQUENCES

Domenico Cantone Salvatore Cristofaro Simone Faro

Università di Catania, Dipartimento di Matematica e Informatica

Viale Andrea Doria 6, I-95125 Catania, Italy

cantone, cristofaro, faro@dmi.unict.it

ABSTRACT

The δ -approximate matching problem arises in many questions concerning musical information retrieval and musical analysis. In the case in which gaps are not allowed between consecutive pitches of the melody, transposition invariance is automatically taken care of, provided that the musical melodies are encoded using the pitch interval encoding. However, in the case in which nonnull gaps are allowed between consecutive pitches of the melodies, transposition invariance is not dealt with properly by the algorithms present in literature.

In this paper, we propose two slightly different variants of the approximate matching problem under transposition invariance and for each of them provide an algorithm, obtained by adapting an efficient algorithm for the δ -approximate matching problem with α -bounded gaps.

Keywords: approximate string matching, musical information retrieval, experimental algorithms.

1 INTRODUCTION

String matching is an extensively studied problem in computer science, mainly due to its direct applications to such diverse areas as text, image and signal processing, speech analysis and recognition, musical analysis, information retrieval, computational biology and chemistry, etc. Formally the string matching problem consists in finding *all* occurrences of a given pattern P in a text T , over some alphabet Σ .

The string matching problem has been generalized with various notions of approximate matching.

In this paper we focus on the δ -approximate string matching problem with α -bounded gaps, or (δ, α) -matching problem for short, which arises in many questions on musical information retrieval and musical analysis. This is especially true in the retrieval of given

melodies from complex musical scores in the context of monophonic music.

Musical sequences can be viewed schematically as sequences of integer numbers, representing either the notes in the chromatic or diatonic notation (*absolute pitch encoding*, abbreviated as *a.p.e.*), or the intervals between consecutive notes, expressed as number of semitones (*pitch interval encoding*, abbreviated as *p.i.e.*).

The latter representation is generally of greater interest in applications in tonal music, since absolute pitch encoding disregards tonal qualities of pitches. In particular, the pitch interval encoding is *transposition invariant*, i.e. any two melodies differing only by an interval transposition are encoded in p.i.e. by a same numerical sequence.

Note durations and note accents can also be encoded in numeric form, yielding richer alphabets whose symbols can really be regarded as sets of parameters. For this reason alphabets used for music representation are generally quite large.

In the δ -approximate matching problem two integer strings of the same length match if the corresponding integers differ by at most a fixed bound δ . Notice that if $\delta = 0$, the δ -approximate string matching problem reduces to the exact string matching problem.

A significant amount of research has been devoted to adapt solutions for exact string matching to δ -approximate matching (see, for instance, Cambouropoulos et al., 1999; Crochemore et al., 2001, 2002b; Cantone et al., 2004).

The δ -approximate matching problem has been generalized by Crochemore et al. (2002a) by admitting also bounded gaps. They proposed an algorithm based on the dynamic programming approach. More recently, faster solutions have been presented by Cantone et al. (2005a,b). At an intuitive level, we say that a melody has a δ -approximate occurrence with α -bounded gaps within a given musical score, if the melody has a δ -approximate matching with a subsequence of the musical score, in which it is allowed to skip up to a fixed number α of symbols (the *gap*) between any two consecutive positions.

For example, in classical music compositions, and in particular in compositions for *Piano Solo*, it is quite common to find musical pieces based on a sweet ground melody, whose notes are interspaced by rapidly executed arpeggios. In such cases, if one wants to retrieve all approximate occurrences of a given melody from a musical score, the value for α must be chosen large enough so that

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

©2005 Queen Mary, University of London

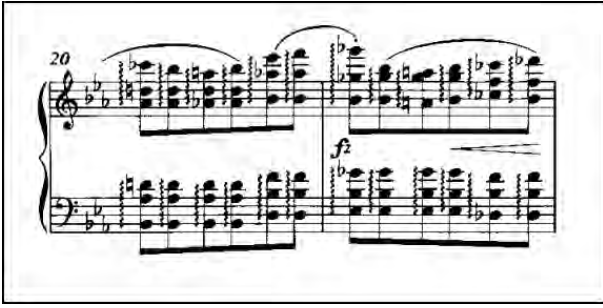


Figure 1: Bars Nr. 20 and Nr. 21 of the study *Opus 10 Nr. 11 for Piano Solo* by F. Chopin.

the secondary notes of each arpeggio can be skipped.

Figure 1 shows two (partial) bars of the study *Opus 10 Nr. 11 for Piano Solo* by F. Chopin illustrating such a point. The notes of the melody are the higher notes of each arpeggiated chord. In this case a value $\alpha = 5$ would suffice to retrieve the melody.

The above musical technicality is not by any means the only one for which approximate matching with bounded gaps turns out to be very useful. Other examples are given by the use of *musical ornaments, pedal-notes*, etc.

2 MATCHING UNDER TRANSPOSITION INVARIANCE

In the problem of finding musical similarities between given melodies, it is quite natural, among other things, to regard as similar any two melodies when they differ by an interval transposition. However, the definition of (δ, α) -matching used in previous works does not take into account the requirement of transposition invariance.

Since transposition invariance is a property which depends only on the intervals between the pitches, rather than on the pitches themselves, when gaps are not allowed the approximation problem under transposition invariance can be solved by any algorithms for the δ -approximate matching problem, provided that the melodies are represented using the pitch interval encoding.

For instance, let us consider the melodies S_1 and S_2 in Figure 2. If S_1 and S_2 are represented using the absolute pitch encoding, then an approximation bound of $\delta \geq 7$ is needed to retrieve a match, whereas a value of $\delta = 1$ would suffice if both melodies were represented using the pitch interval encoding.

However, the use of the pitch interval encoding works well only in the case in which gaps between pitches are not allowed. Referring again to Figure 2, although the melody S_3 has plainly a transposed δ -approximate occurrence with α -bounded gaps in melody S_2 (with $\alpha = 1$), such an occurrence would not be retrieved by the existing (δ, α) -approximate matching, unless a quite large value for δ were used. The main reason behind this fact is that when gaps are allowed, one has to devise a way to keep relevant tonal information contained in the pitches of the gap which otherwise would be lost.



Figure 2: Two excerpts from *Toccata and Fuga in D minor* by J.S. Bach.

In this section we propose two slightly different definitions of the δ -approximate matching problem with α -bounded gaps, which take into account transposition invariance.

2.1 (δ, α) -Matching Under Transposition Invariance

Let Σ be an alphabet of integer numbers and let δ , α , and τ be nonnegative integers. Two symbols a and b of Σ are said to be δ -approximate if $|a - b| \leq \delta$. In this case we say that a and b have a δ -match and we write $a =_{\delta} b$.

Given a text T of length n and a pattern P of length m , a τ -transposed (δ, α) -occurrence of P in T at position i is an increasing sequence of indices $(i_0, i_1, \dots, i_{m-1})$ such that

1. $0 \leq i_0$ and $i_{m-1} = i \leq n - 1$;
2. $i_h - i_{h-1} \leq \alpha + 1$, for $1 \leq h \leq m - 1$;
3. $|P[0] - T[i_0]| = \tau$;
4. $P[j] - P[j - 1] =_{\delta} T[i_j] - T[i_{j-1}]$, for $1 \leq j \leq m - 1$.

The above definition of δ -approximate occurrence with α -bounded gaps is slightly different from the one given in Cantone et al. (2005a) in that it takes into account δ -matching of intervals between (immediately consecutive) pitches, rather than δ -matching of the pitches themselves.

Given a text T and a pattern P , we write $P \triangleleft_{\delta, \alpha, \tau}^i T$ to mean that P has a τ -transposed (δ, α) -occurrence in T at position i .

Thus we give the following definition

Definition 1 Given a text T , a pattern P , and two non-negative integers δ and α , the δ -APPROXIMATE MATCHING PROBLEM WITH α -BOUNDED GAPS UNDER TRANSPOSITION INVARIANCE consists in finding all positions i in T such that $P \triangleleft_{\delta, \alpha, \tau}^i T$, for some $\tau \geq 0$.

According to the above definition, the melody S_3 of Figure 2 has a transposed (δ, α) -match within S_1 , provided that both melodies are represented in a.p.e. and that $\delta, \alpha \geq 1$.

Notice that when $\alpha = 0$, the δ -approximate matching problem with α -bounded gaps under transposition invariance reduces to the δ -approximate matching problem relative to melodies represented in p.i.e.

2.2 (δ, α) -Matching Under Ranged-Transposition Invariance

The notion of τ -transposed (δ, α) -occurrence given before takes into consideration only the intervals between immediately consecutive pitches in the pattern melody P , namely the intervals $P[j] - P[j - 1]$, for $j = 1, \dots, m - 1$.

This definition allows to bound the *local* tonal variations between pitches, but may lead to large *global* tonal variations along the whole sequence. More precisely, a difference δ between any two consecutive pitches in a sequence P of length m can lead to an overall difference of $\delta \cdot (m - 1)$ between the first and the last symbols of the sequences. For instance, consider Figure 3, which shows the **C** major scale (first score) and the **A** whole tone scale (second score). If we set $\delta = 2$ and $\alpha = 1$, then we can retrieve a transposed (δ, α) -occurrence of the whole tone scale as part of the major scale (in fact, we retrieve the **Dm13** chord in arpeggiated form) indicated in the first score by the squared notes. Observe, however, that the overall difference between the first and the last symbols of the sequences is 9 semitones.

However, to take care of transposition invariance, we do not need to consider necessarily intervals formed between immediately consecutive pitches in the pattern melody. To bound the global tonal variations of a matched sequence it may be more convenient to fix a particular pitch of the melody —the pivot— and then consider the intervals relative to such a pitch.

Choosing the first pitch of a melody as pivot leads to a new definition of (δ, α) -occurrence of a pattern P in a text T , which we call τ -ranged-transposed (δ, α) -occurrence of P in T (at a given position i). The new definition differs from the previous one in that clause 4 is replaced by the following:

- 4'. $P[j] - P[0] =_{\delta} T[i_j] - T[i_0]$, for $1 \leq j \leq m - 1$.

Given a text T and a pattern P , we write $P \triangleleft_{\delta, \alpha, \tau}^i T$ to mean that P has a τ -ranged-transposed (δ, α) -occurrence in T at position i . Thus we end up with the following definition:

Definition 2 Given a text T , a pattern P , and two nonnegative integers δ and α , the δ -APPROXIMATE MATCHING PROBLEM WITH α -BOUNDED GAPS UNDER RANGED-TRANSPPOSITION INVARIANCE consists in finding all positions i in T such that $P \triangleleft_{\delta, \alpha, \tau}^i T$, for some $\tau \geq 0$.

In general, approximate matching under ranged-transposition invariance leads to more accurate matching results than approximate matching under transposition invariance defined previously in Section 2.1. For instance, in Figure 3, to retrieve a ranged-transposed (δ, α) -occurrence, we need a value of $\delta \geq 9$ for the approximation bound.

3 TWO SEQUENTIAL-SAMPLING ALGORITHMS

In this section we present two algorithms which solve the approximate matching problems introduced in the previ-

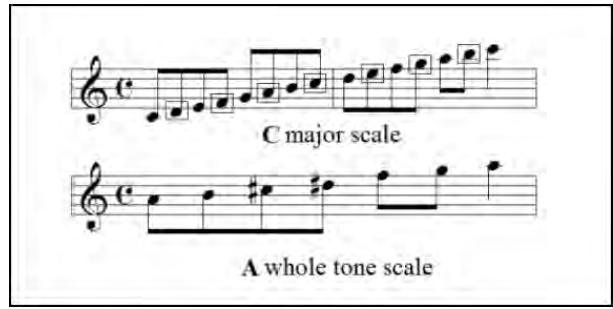


Figure 3: The **C** major scale and the **A** whole tone scale. The squared notes in the first score represent an approximate occurrence under transposition invariance of the whole tone scale.

ous section. Our algorithms, respectively named (δ, α) -T-Sequential-Sampling and (δ, α) -RT-Sequential-Sampling, are modified versions of the (δ, α) -Sequential-Sampling algorithm presented by the authors in Cantone et al. (2005a).

We begin by considering first the (δ, α) -matching problem under transposition invariance (cf. Definition 1), and then we turn our attention to the (δ, α) -matching problem under ranged-transposition invariance (cf. Definition 2).

Given a text T of length n , a pattern P of length m , and two integers $\delta, \alpha \geq 0$, for $i = 0, 1, \dots, n$ we put

$$\mathcal{S}_i = \{(j, k) : 0 \leq j < i, \\ 0 \leq k < m, \text{ and} \\ P_k \triangleleft_{\delta, \alpha, \tau}^i T \text{ for some } \tau \geq 0\},$$

where P_k denotes the prefix of P of length $k + 1$.

Notice that $\mathcal{S}_0 = \emptyset$. If we put $\mathcal{S} = \mathcal{S}_n$, the problem of finding all positions i in T such that $P_k \triangleleft_{\delta, \alpha, \tau}^i T$ for some $\tau \geq 0$ translates into the problem of finding all values i such that $(i, m - 1) \in \mathcal{S}$. To begin with, note that for $i = 0, 1, \dots, n - 1$ we have

$$\mathcal{S}_{i+1} = \mathcal{S}_i \cup \{(i, 0)\} \\ \cup \{(i, k) : k > 0, (i - h, k - 1) \in \mathcal{S}_i \text{ and} \\ T[i] - T[i - h] =_{\delta} P[k] - P[k - 1] \\ \text{for some } 1 \leq h \leq \alpha + 1\}.$$

The above recursive relation, coupled with the initial condition $\mathcal{S}_0 = \emptyset$, allows one to compute the set \mathcal{S} in an iterative fashion, by computing in turn the sets $\mathcal{S}_0, \mathcal{S}_1, \dots, \mathcal{S}_n$, starting from \mathcal{S}_0 . Such computation can be accomplished by means of the (δ, α) -T-Sequential-Sampling algorithm shown in Figure 4, whose time and space complexities are easily seen to be respectively $\mathcal{O}(mn\alpha)$ and $\mathcal{O}(mn)$.

Since during each iteration of the **for**-loop at line 4 of the (δ, α) -T-Sequential-Sampling algorithm at most $m \cdot (\alpha + 1)$ pairs of \mathcal{S} need to be considered, it would be enough to maintain only $m \cdot (\alpha + 1)$ pairs of \mathcal{S} at each step of the computation. With this modification, the space complexity of the (δ, α) -T-Sequential-Sampling algorithm can be reduced to $\mathcal{O}(m\alpha)$.

To solve the (δ, α) -matching problem under ranged-transposition invariance, for each pair (i, k) with $0 \leq i < n$ and $0 \leq k < m$, we define a set $\Gamma(i, k)$ which col-

```

 $(\delta, \alpha)$ -T-Sequential-Sampling( $T, P, \delta, \alpha$ )
1.  $n = \text{length}(T)$ 
2.  $m = \text{length}(P)$ 
3.  $S = \emptyset$ 
4. for  $i = 0$  to  $n - 1$  do
5.   for  $k = m - 2$  downto  $0$  do
6.     if  $\exists h \in \{1, \dots, \alpha + 1\} : (i - h, k) \in S$  and
7.        $P[k + 1] - P[k] =_{\delta} T[i] - T[i - h]$  then
8.          $S = S \cup \{(i, k + 1)\}$ 
9.    $S = S \cup \{(i, 0)\}$ 
10. for  $i = 0$  to  $n - 1$  do
11.   if  $(i, m - 1) \in S$  then output( $i$ )

```

Figure 4: The algorithm (δ, α) -T-Sequential-Sampling for the δ -approximate matching problem with α -bounded gaps under transposition invariance.

```

 $(\delta, \alpha)$ -RT-Sequential-Sampling( $T, P, \delta, \alpha$ )
1.  $n = \text{length}(T)$ 
2.  $m = \text{length}(P)$ 
3. for  $i = 0$  to  $n - 1$  do
4.   for  $k = m - 2$  downto  $0$  do
5.      $\Gamma(i, k + 1) = \emptyset$ 
6.     for  $h = 1$  to  $\alpha + 1$  do
7.       for each  $j \in \Gamma(\max(0, i - h), k)$  do
8.         if  $P[k + 1] - P[0] =_{\delta} T[i] - T[j]$  then
9.            $\Gamma(i, k + 1) = \Gamma(i, k + 1) \cup \{j\}$ 
10.    $\Gamma(i, 0) = \{i\}$ 
11. for  $i = 0$  to  $n - 1$  do
12.   if  $\Gamma(i, m - 1) \neq \emptyset$  then output( $i$ )

```

Figure 5: The (δ, α) -RT-Sequential-Sampling algorithm for the δ -approximate matching problem with α -bounded gaps under ranged-transposition invariance.

lects the starting positions j of all τ -ranged-transposed (δ, α) -occurrences of P_k in T ending at position i . Then, the problem of finding all positions i in T such that $P \triangleleft_{\delta, \alpha, \tau}^i T$, for some $\tau \geq 0$, translates into the problem of finding all values i such that $\Gamma(i, m - 1) \neq \emptyset$. We observe that $\Gamma(i, 0) = \{i\}$, for $0 \leq i < n$. In addition, we have:

$$\Gamma(i, k) = \{j : P[k] - P[0] =_{\delta} T[i] - T[j] \text{ and } j \in \Gamma(\max(0, i - h), k - 1) \text{ for some } h \in \{1, \dots, \alpha + 1\}\},$$

for $0 \leq i < n$ and $0 < k < m$.

Thus the computation of the sets $\Gamma(i, k)$ can be done iteratively, by the (δ, α) -RT-Sequential-Sampling algorithm shown in Figure 5, whose time and space complexities are easily seen to be respectively $\mathcal{O}(m^2 n \alpha)$ and $\mathcal{O}(mn)$.

4 CONCLUSIONS

Existing solutions to the approximate matching problem with bounded gaps do not cover well the case in which transposition invariance is enforced, regardless of the encoding used (either a.p.e. or p.i.e.).

In this paper, we have proposed two slightly different variants of the approximate matching problem under transposition invariance and for both of them we have shown how to adapt the (δ, α) -Sequential-Sampling algorithm to take care of transposition invariance.

We plan to further optimize the algorithms presented. We also intend to address other variants of the approximation matching problem, relevant to applications in musical analysis and musical information retrieval.

REFERENCES

- E. Cambouropoulos, M. Crochemore, C. S. Iliopoulos, L. Mouchard, and Y. J. Pinzon. Algorithms for computing approximate repetitions in musical sequences. In R. Raman and J. Simpson, editors, *Proceedings of the 10th Australasian Workshop on Combinatorial Algorithms*, pages 129–144, Perth, WA, Australia, 1999.
- D. Cantone, S. Cristofaro, and S. Faro. Efficient algorithms for the δ -approximate string matching problem in musical sequences. In M. Šimánek and J. Holub, editors, *Proceedings of the Prague Stringology Conference '04*, pages 69–82, Czech Technical University, Prague, Czech Republic, 2004.
- D. Cantone, S. Cristofaro, and S. Faro. An efficient algorithm for δ -approximate matching with α -bounded gaps in musical sequences. In S.E. Nikolettseas, editor, *Proceedings of the 4th International Workshop on Experimental and Efficient Algorithms (WEA 2005)*, number 3503 in Lecture Notes in Computer Science, pages 428–439. Springer-Verlag, Berlin, 2005a.
- D. Cantone, S. Cristofaro, and S. Faro. On tuning the (δ, α) -sequential-sampling algorithm for δ -approximate matching with α -bounded gaps in musical sequences. Submitted to ISMIR 2005. 2005b. Available at the URL <http://www.dmi.unict.it/~cantone/CCF-ISMIR05.pdf>.
- M. Crochemore, C. S. Iliopoulos, T. Lecroq, and Y. J. Pinzon. Approximate string matching in musical sequences. In M. Balík and M. Šimánek, editors, *Proceedings of the Prague Stringology Conference '01*, pages 26–36, Prague, Czech Republic, 2001. Annual Report DC–2001–06.
- M. Crochemore, C. Iliopoulos, C. Makris, W. Rytter, A. Tsakalidis, and K. Tsihlias. Approximate string matching with gaps. *Nordic Journal of Computing*, 9 (1):54–65, 2002a.
- M. Crochemore, C. S. Iliopoulos, T. Lecroq, W. Plandowski, and W. Rytter. Three heuristics for δ -matching: δ -BM algorithms. In A. Apostolico and M. Takeda, editors, *Proceedings of the 13th Annual Symposium on Combinatorial Pattern Matching*, number 2373 in Lecture Notes in Computer Science, pages 178–189. Springer-Verlag, Berlin, 2002b.