Symbolic Music Similarity

Presentation

Ali Bektas Paul Kröger

February 3, 2020

Überblick

Darstellung von Noten

- Melodie : *"singbare, in sich geschlossene Folge von Tönen"* [1]
- Harmonie : *"wohltönender Zusammenklang mehrerer Töne oder Akkorde"* [2]
- Schlüssel : *"dient in der Musiknotation dazu, im Notensystem festzulegen, welche Tonhöhe die fünf Notenlinien repräsentieren."* [3]

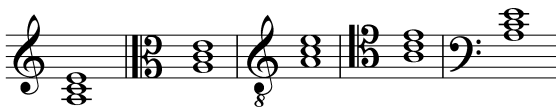
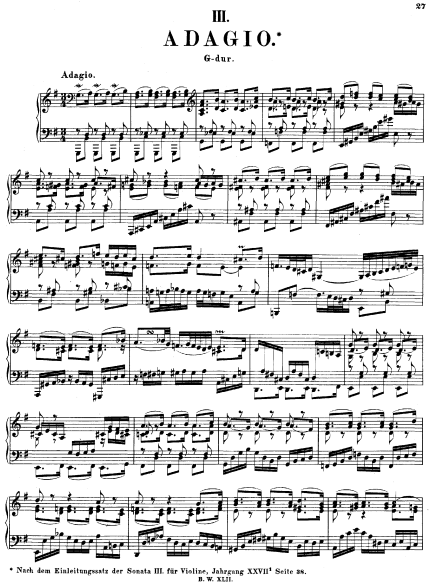


Figure: Source: [3]

Darstellung von Noten

Im Grunde genommen , ermöglicht die herkömmliche Methode von Notendarstellung , Informationen über Rhytmus , Tonlage , Gefühl beim Spielen , vortragsbetreffliche Elemente zu übermitteln.


III.

ADAGIO.*

G-dur.

Adagio.

* Nach dem Einleitungssatz der Sonata III. für Violine, Jahrgang XXVII[1] Seite 28.

B. W. XLII.

Figure: Source: IMLSP Archive

Darstellung von Noten

"Representing music as a weighted point set in a two-dimensional space has a tradition of many centuries. Since approximately the 10th century, one popular way of writing music has been to use a set of notes (points) in a two-dimensional space, with time and pitch as coordinates."[6]

Insert Title

"A Measure of Melodic Similarity Based on a Graph Representation of the Music Structure" [7]
von Nicola Orio und Antonio Rodá.

10th International Society for Music Information Retrieval Conference (ISMIR 2009)

A MEASURE OF MELODIC SIMILARITY BASED ON A GRAPH REPRESENTATION OF THE MUSIC STRUCTURE

Nicola Orio
Department of Information Engineering
University of Padova
nicola.orio@dei.unipd.it

Antonio Rodà
Lab. AVIRES
University of Udine
antonio.roda@uniud.it

ABSTRACT

Content-based music retrieval requires to define a similarity measure between music documents. In this paper, we propose a novel similarity measure between melodic content, as represented in symbolic notation, that takes into account musicological aspects on the structural function of the melodic elements. The approach is based on the representation of a collection of music scores with a graph structure, where terminal nodes directly describe the music content, internal nodes represent its incremental generalization, and arcs denote the relationships among them. The similarity between two melodies can be computed by analyzing the graph structure and finding the shortest path between the corresponding nodes inside the graph. Preliminary results in terms of music similarity are presented using a small test collection.

1. INTRODUCTION

One approach to content-based access to music documents is to provide users with tools to retrieve music documents that are *similar* to a set of one or more documents already known, which are the starting point of a query-by-example paradigm. The effectiveness of the results depends on the way a measure of music similarity is computed. This task is difficult to define, because the notion of music similarity is subjective and also because the role played by the different music dimensions – i.e., rhythm, melody, harmony, timbre, orchestration, tempo – is task dependent. For instance, the perceived similarity of two ballroom songs is mainly related to rhythm, while in jazz music it can depend on chord progressions.

In recent years, a major interest has been given to the retrieval of audio documents, typically in compressed formats such as MP3. This trend is explained by the increasing availability of large collections of audio files, and by the fact that users without a music training are usually not interested in accessing symbolic representations. For this reason, the notion of music similarity has been biased towards the music dimensions that, on the one hand, are more relevant for music listeners and, on the other hand, can be reliably extracted from audio files. A typical approach is to extract some timbre descriptors to address different tasks, such as genre and artist identification, automatic playlist generation, or music collection visualization [12].

Music documents can be represented also in symbolic forms, such as a notated digital score or a MIDI file. The access to these documents can be based on higher level features, such as melodic profile and harmonic progressions, which are not easily extracted from audio files. In particular, the melodic profile has been often used as the main dimension to compute the similarity between music documents [1, 7]. A typical task of melody-based retrieval is the automatic identification of a melody sung or hummed by the user. For this application, the similarity measure has to be robust to local mismatches due to imprecise recall from memory and to a lack of singing skills by the users, because it is assumed that the query and the relevant documents are representing the same information. Although the term *query-by-humming* was very popular in the early days of MIR research [3], recently it has been often replaced by the more general term query-by-example because it is assumed that users can easily record with a portable device an excerpt of the song they are interested to retrieve and are not willing to hum a melody in front of an user interface. Approaches of this kind may be based on approximate matching to identify the music work corresponding to the recorded performance, a task that is usually defined as cover [5] or music [8] identification.

The computation of melodic similarity can be useful also for applications other than an identification task. For instance, in musicological analysis the study of the melodic material used by different composers, or consistently used by a given composer, is of particular interest. Also ethnomusicological studies can take advantage from melodic similarity in order to track the evolution of a given song over the centuries and its diffusion in different geographical regions. Melodic similarity can be exploited to retrieve music that "sounds like" other well known songs, for instance to find a suitable soundtrack for a TV show.

We propose a novel similarity measure computed between music content in symbolic format, that takes into account the musical structure of the composition through the application of an analytic method. Our approach aims

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.
© 2009 International Society for Music Information Retrieval.

Insert Title

- Der Inhalt wird schrittweise vereinfacht.
- Dazu sind die **Gewichte** der einzelnen Noten von Bedeutung.
  - die unterliegende harmonische Funktion (harmonic weight)
  - die metrische Position (metric weight)
  - die Differenz der Tonlagen zwischen dem Ton und dem Grundton(melodic weight)

Insert Title

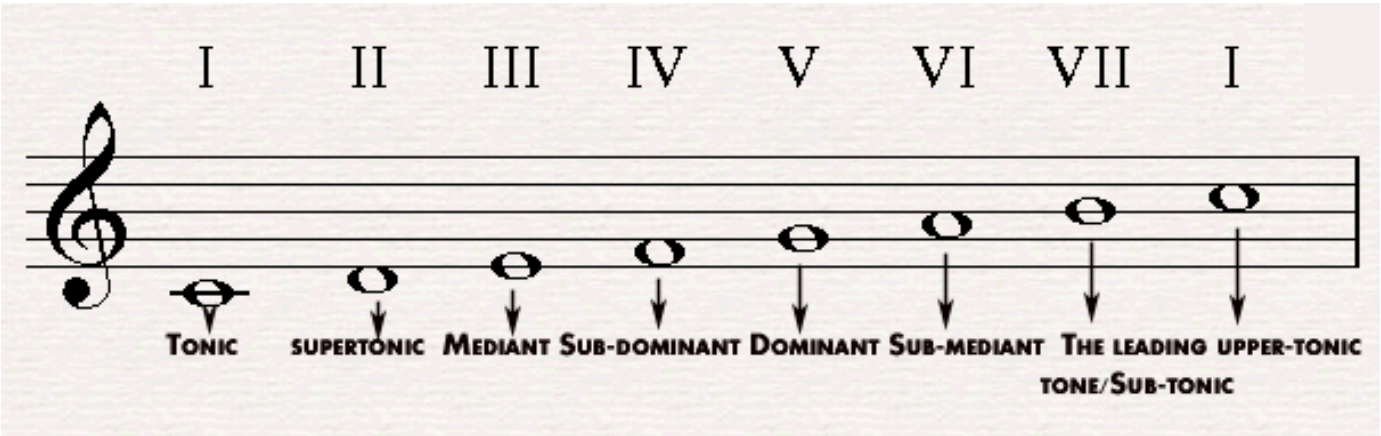


Figure: Funktionen der Noten im Skala [9]

Insert Title

- Der Inhalt wird schrittweise vereinfacht.
- Dazu sind die **Gewichte** der einzelnen Noten von Bedeutung.
  - die unterliegende harmonische Funktion (harmonic weight)
  - die metrische Position (metric weight)
  - die Differenz der Tonlagen zwischen dem Ton und dem Grundton(melodic weight)

Insert Title

"Algorithms for Computing Geometric Measures of Melodic Similarity" [8]
von Greg Aloupis, Thomas Fevens, Stefan Langerman, Tomomi Matsui, Antonio Mesa, Yurai Nunez, David Rappaport, and Godfried Toussaint

Greg Aloupis,[*] Thomas Fevens,[†] Stefan Langerman,[‡] Tomomi Matsui,[§] Antonio Mesa,[¶] Yurai Nuñez,[¶] David Rappaport,[**] and Godfried Toussaint[†]

[*] School of Computer Science, McGill University
3480 University Street
Montreal, Quebec, Canada H3A 2A7
{athens,godfried}@cs.mcgill.ca

[†] Department of Computer Science and Software Engineering, Concordia University
1455 de Maisonneuve Boulevard West
Montreal, Quebec, Canada H3G 1M8
fevens@cs.concordia.ca

[‡] Chercheur Qualifié du FNRS, Département d'Informatique, Université Libre de Bruxelles
CP212 Boulevard du Triomphe, 1050 Bruxelles, Belgium
Stefan.Langerman@ulb.ac.be

[§] Department of Mathematical Informatics
Graduate School of Information Science and Technology, University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan
tomomi@misojiro.t.u-tokyo.ac.jp

[¶] Departamento Ciencias de la Computacion
Facultad de Matematica y Computacion,
Universidad de La Habana
San Lazaro y L, Vedado 10400
Ciudad de La Habana, Cuba
tonymesa@matcom.uh.cu, yurainn@yahoo.com

[**] School of Computing, Queen's University
Kingston, Ontario, Canada K7L 3N6
daver@cs.queensu.ca

Algorithms for Computing Geometric Measures of Melodic Similarity

We have all heard numerous melodies, whether they come from commercial jingles, jazz ballads, operatic arias, or any of a variety of different sources. How a human detects similarities in melodies has been studied extensively (Martinez 2001, Hofmann-Engl 2002, Müllensiefen and Frieler 2004). There has also been some effort in modeling melodies so that similarities can be detected algorithmically. Some results in this fascinating study of musical perception and computation can be found in a collection edited by Hewlett and Selfridge-Field (1998).

Similarity measures for melodies find application in content-based retrieval methods for large music databases such as query by humming (QBH) (Ghias et al. 1995; Mo, Han, and Kim 1999) but also in other diverse applications such as helping prove music copyright infringement (Cronin 1998). Previous formal mathematical approaches to rhythmic and melodic similarity, such as the one taken in this article, are based on methods like one-dimensional edit-distance computations (Toussaint 2004), approximate string-matching algorithms (Bainbridge et al. 1999, Lemström 2000), hierarchical correlation functions (Lu, You, and Zhang 2001), two-dimensional augmented suffix trees (Chen et al. 2000), transportation distances (Typke et al. 2003; Lubiw and Tanur 2004), and maximum segment overlap (Ukkonen, Lemström, and Mäkinen 2003).

Computer Music Journal, 30:3, pp. 67–76, Fall 2006
© 2006 G. Aloupis, T. Fevens, S. Langerman, T. Matsui, A. Mesa, Y. Nuñez, D. Rappaport, and G. Toussaint.

This content downloaded from 130.15.241.107 on Wed, 27 Jan 2016 17:47:18 UTC
All use subject to JSTOR Terms and Conditions

Insert Title

- Melodien werden als Polygonalketten dargestellt
- Tonlänge wird durch Länge der waagerechten Kanten modelliert
- Intervalle werden durch Länge der senkrechten Kanten modelliert

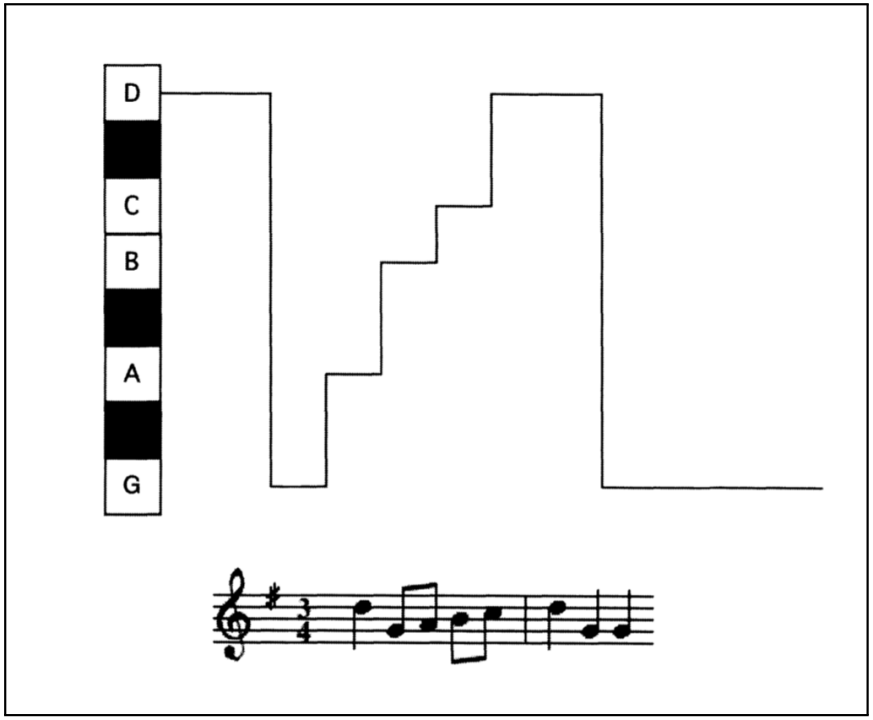
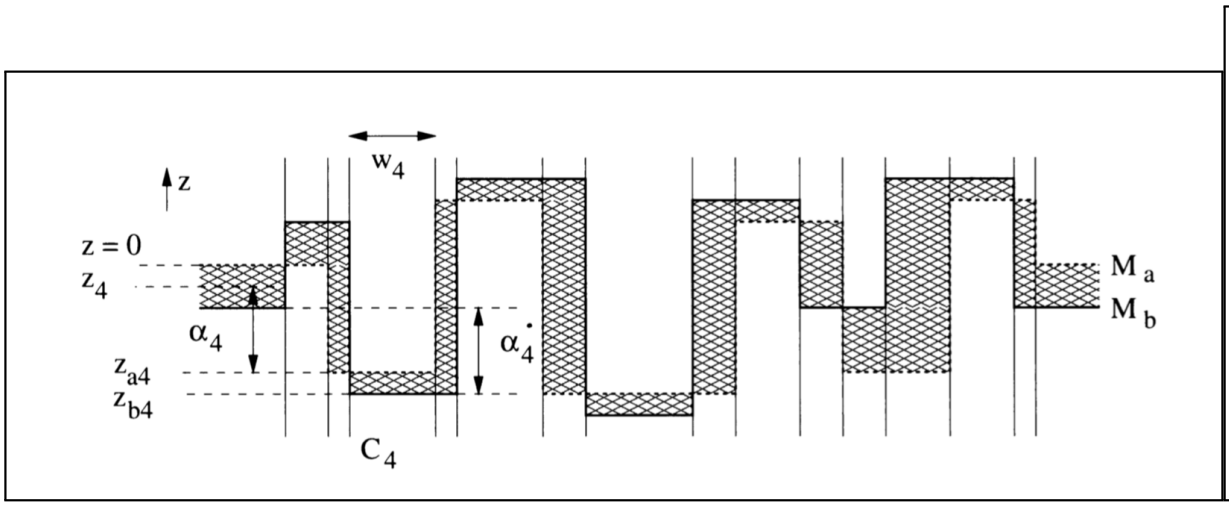
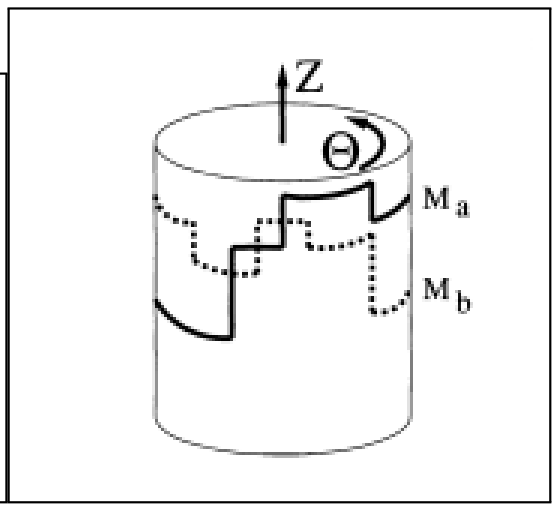


Figure: Source : [8]

Insert Title





Source : [8]

MIREX

- Ein Wettbewerb und Plattform für Interessierte
- Es gibt verschiedene Kategorien
  - Real-time Audio to Score Alignment (a.k.a Score Following)
  - Discovery of Repeated Themes and Sections
  - Audio Melody Extraction
  - **Symbolic Melodic Similarity**
  - ...
- Gegeben ein Ziel , treten verschiedene Algorithmen gegeneinander zum Wettkampf an. Derjenige, der die besten Ergebnisse hat , gewinnt.
- **Nun eine Frage:**Wie kann man Algorithmen miteinander vergleichen?
- Es kommt nicht auf die Laufzeit oder Speicherbedarf an , sondern auf die Qualität der Ergebnisse.
- Welche Messmethoden gibt es , um die Qualität von solcen Ergebnissen zu beurteilen?

MIREX

Overall	AP1	AP2	AU1	AU2	AU3	GAR1	GAR2	FHAR
ADR	0.031	0.024	0.666	0.698	0.706	0.712	0.739	0.730
NRGB	0.028	0.027	0.601	0.590	0.616	0.617	0.683	0.666
AP	0.017	0.023	0.525	0.477	0.500	0.508	0.545	0.545
PND	0.044	0.056	0.527	0.495	0.515	0.494	0.588	0.557
Fine	0.292	0.281	0.532	0.528	0.532	0.586	0.581	0.540
Psum	0.234	0.190	0.522	0.524	0.527	0.589	0.580	0.517
WCsum	0.179	0.146	0.470	0.480	0.486	0.537	0.526	0.470
SDsum	0.152	0.123	0.444	0.458	0.465	0.511	0.498	0.447
Greater0	0.397	0.323	0.677	0.653	0.650	0.743	0.743	0.657
Greater1	0.070	0.057	0.367	0.393	0.403	0.433	0.417	0.377

Figure: Source: [5]

MIREX

Overall	AP1	AP2	AU1	AU2	AU3	GAR1	GAR2	FHAR
ADR	0.031	0.024	0.666	0.698	0.706	0.712	0.739	0.730
NRGB	0.028	0.027	0.601	0.590	0.616	0.617	0.683	0.666
AP	0.017	0.023	0.525	0.477	0.500	0.508	0.545	0.545
PND	0.044	0.056	0.527	0.495	0.515	0.494	0.588	0.557
Fine	0.292	0.281	0.532	0.528	0.532	0.586	0.581	0.540
Psum	0.234	0.190	0.522	0.524	0.527	0.589	0.580	0.517
WCsum	0.179	0.146	0.470	0.480	0.486	0.537	0.526	0.470
SDsum	0.152	0.123	0.444	0.458	0.465	0.511	0.498	0.447
Greater0	0.397	0.323	0.677	0.653	0.650	0.743	0.743	0.657
Greater1	0.070	0.057	0.367	0.393	0.403	0.433	0.417	0.377

Inhaltsübersicht

MIREX

Overall	AP1	AP2	AU1	AU2	AU3	GAR1	GAR2	FHAR
ADR	0.031	0.024	0.666	0.698	0.706	0.712	0.739	0.730
NRGB	0.028	0.027	0.601	0.590	0.616	0.617	0.683	0.666
AP	0.017	0.023	0.525	0.477	0.500	0.508	0.545	0.545
PND	0.044	0.056	0.527	0.495	0.515	0.494	0.588	0.557
Fine	0.292	0.281	0.532	0.528	0.532	0.586	0.581	0.540
Psum	0.234	0.190	0.522	0.524	0.527	0.589	0.580	0.517
WCsum	0.179	0.146	0.470	0.480	0.486	0.537	0.526	0.470
SDsum	0.152	0.123	0.444	0.458	0.465	0.511	0.498	0.447
Greater0	0.397	0.323	0.677	0.653	0.650	0.743	0.743	0.657
Greater1	0.070	0.057	0.367	0.393	0.403	0.433	0.417	0.377

Average Dynamic Recall - ADR

Inhaltsübersicht

Insert Title

"MelodyShape at MIREX 2014 Symbolic Melodic Similarity" [10]
von Julian Urbano

MelodyShape at MIREX 2014 Symbolic Melodic Similarity

Julián Urbano
Universitat Pompeu Fabra
Barcelona, Spain
julian.urbano@upf.edu

ABSTRACT

This short paper describes our three submissions to the 2014 edition of the MIREX Symbolic Melodic Similarity task. All three submissions rely on a geometric model that represents melodies as spline curves in the pitch-time plane. The similarity between two melodies is then computed with a sequence alignment algorithm between sequences of spline spans: the more similar the shape of the curves, the more similar the melodies they represent. As in the previous MIREX 2010, 2011, 2012 and 2013 editions, our systems ranked first for all effectiveness measures. The main difference with last year is that we submitted a re-implementation of all algorithms, contained in the new open source library MelodyShape.

1. INTRODUCTION

For the 2014 edition of the MIREX Symbolic Melodic Similarity task we submitted the same three algorithms as last year. JU1-ShapeH implements the same algorithm that has consistently obtained the best or second-best results in MIREX 2010–2013. The second submission is called JU2-ShapteTime, and it contains the same algorithm as in MIREX 2013 and 2012. It works like ShapeH, except that the top-1 retrieved results are further re-ranked using the third system, called JU3-Time (also submitted in MIREX 2013 and 2012). This system was shown to be especially good at ranking results, so it is used to complement ShapeH for rank-aware measures.

We submitted these algorithms again to evaluate them with a different set of queries and assessors, and to serve as strong and cross-year baselines to measure possible improvements in other submissions. In addition, we recently published MelodyShape [1], a Java tool as library implementing all our algorithms since 2010 from their original code in C#. The three submissions to MIREX 2014 are the exact implementations found in MelodyShape v1.3.

In MIREX 2010, 2011, 2012 and 2013 all our systems ranked first [2–5]. In this MIREX 2014 edition the three systems again ranked at the very top [6].

[1] https://github.com/julian-urbano/MelodyShape

This document is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.
http://creativecommons.org/licenses/by-nc-sa/3.0/
© 2014 The Authors.

Figure 1. Melody as a curve in the pitch-time plane.

2. GEOMETRIC MELODY REPRESENTATION

Melodies are represented as curves in the pitch-time plane, arranging notes according to their pitch height and onset time. For the pitch dimension we use a directed interval representation, while for the time dimension we use the onset ratio between successive notes. We then calculate the interpolating curve passing through the notes (see Figure 1). From that point on, only the curves are used to compute the similarity between melodies [10].

We use Uniform B-Splines to interpolate through the notes [1], which give us a parametric polynomial piecewise function for the spline: one function for the pitch dimension and another one for the time dimension. Their first derivatives measure how much the melodies change at any point. This representation is transposition invariant because two transposed melodies have the same first derivative (see Figure 2). It is also time-scale invariant because we use duration ratios within spline spans instead of actual durations.

A melody is thus represented as a sequence of spline spans, each of which can be considered the same as an n-gram. Given two arbitrary melodies, we compare them with a sequence alignment algorithm, which computes the differences between two spans based on their geometry.

3. SYSTEM DESCRIPTIONS

3.1 ShapeH

In this system we completely ignore the time dimension and use spans 3-notes long, which result in splines defined by polynomials of degree 2. These are then differentiated, so we actually use polynomials of degree 1 to represent melodies. In addition, we implemented a heuristic very similar to the classical *idf* (Inverse Document Frequency) in Text Information Retrieval: the more frequent a spline

Urbano MelodyShape

- Töne werden als Punkt auf Pitch-Time plane dargestellt.
- Darstellung als Funktion durch Interpolation mithile von Splines.

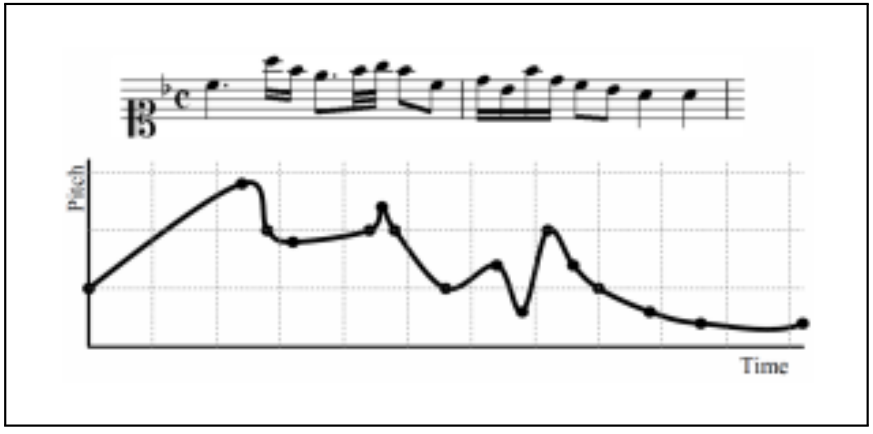


Figure: Source : [10]

Needlemann - Wunsch Algorithmus

$$D = \begin{pmatrix} & - & A & G & T & C \\ - & 0 & -1 & -2 & -3 & -4 \\ A & -1 & 0 & 0 & 0 & 0 \\ C & -2 & 0 & 0 & 0 & 0 \\ G & -3 & 0 & 0 & 0 & 0 \\ T & -4 & 0 & 0 & 0 & 0 \\ C & -5 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 & -1 & -2 & -3 & -4 \\ -1 & 1 & 0 & -1 & -2 \\ -2 & 0 & 0 & -1 & 0 \\ -3 & -1 & 1 & 0 & -1 \\ -4 & -2 & 0 & 2 & 1 \\ -5 & -3 & -1 & 1 & 3 \end{pmatrix}$$

Source : [11]

ShapeH

- Insertion :
  $s(-, n) = -(1 - f(n))$
- Deletion:
  $s(n, -) = -(1 - f(n))$
- Match:
  $s(n, n) = 1 - f(n)$

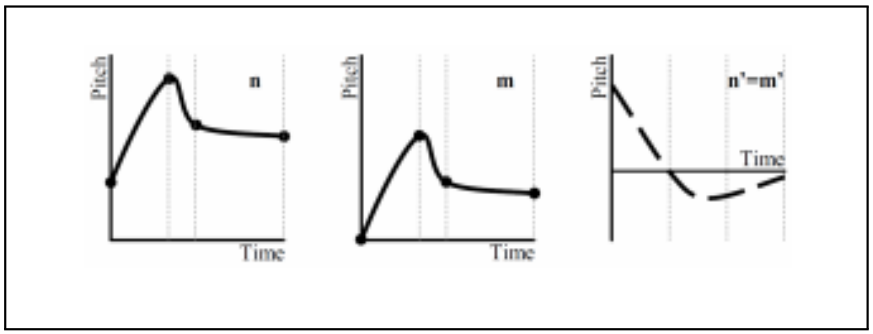


Figure: Source : [10]

Time

- Insertion : $s(-, n) = -diff_p(n, \Theta(n)) - \lambda k_t * diff_t(n, \Theta(n))$
- Deletion: $s(n, -) = -diff_p(n, \Theta(n)) - \lambda k_t * diff_t(n, \Theta(n))$
- Match: $2\mu_p + 2\lambda k_t \mu_t = 2\mu_p(1 + k_t)$
- Substitution $s(n, m) = -diff_p(n, m) - \lambda k_t * diff_t(n, m)$

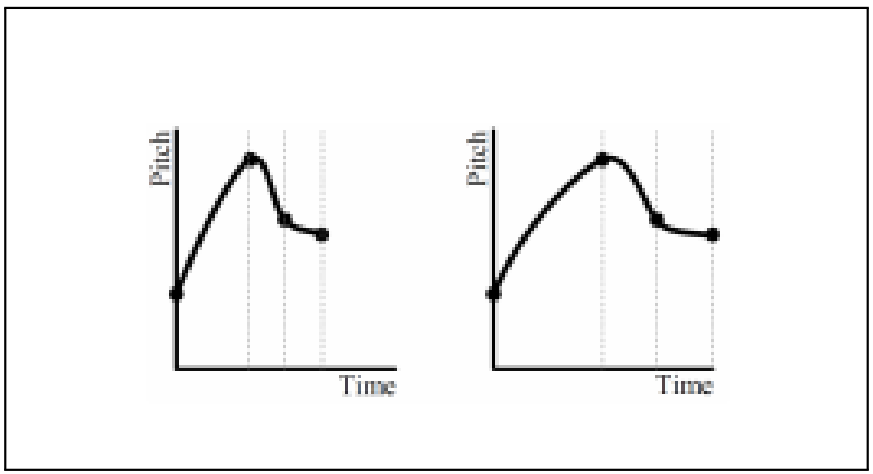


Figure: Source : [10]

Bibliographie I

[1] Duden: Melodie: Rechtschreibung, Bedeutung, Definition, Herkunft https://www.duden.de/rechtschreibung/Melodie.

[2] Duden: Harmonie : Rechtschreibung, Bedeutung, Definition, Herkunft https://www.duden.de/rechtschreibung/Harmonie.

[3] "Notenschlüssel." Wikipedia, Wikimedia Foundation, 11 Dec. 2019, de.wikipedia.org/wiki/Notenschlüssel.

[4] MIREX,Symbolic Melodic Similarity 2005,https://www.music-ir.org/mirex/wiki/2005:Symbolic_Melodic.

[5] MIREX,Symbolic Melodic Similarity Results 2007, https://www.music-ir.org/mirex/wiki/2007:Symbolic_Melodic_Similarity_Results.

Bibliographie II

[6] Typke, Rainer. (2007). Music Retrieval based on Melodic Similarity.

[7] Orio, N., and A. Rodá. 2009. "A Measure of Melodic Similarity Based on a Graph Representation of the Music Structure." In Proceedings of the International Conference for Music Information Retrieval, pp. 543– 548.

[8] Greg Aloupis, Thomas Fevens, Stefan Langerman, Tomomi Matsui, Antonio Mesa, Yurai Nunez, David Rappaport, and Godfried Toussaint, "Algorithms for Computing Geometric Measures of Melodic Similarity" Computer Music Journal, Vol.30, No. 3 (Autumn, 2006), pp. 67-76

Bibliographie III

[9] Tonal Degrees [Online]. [Accessed 30 Jan 2020]. Available from : http://www.piano-play-it.com/musical-scales.html

[10] J. Urbano. MelodyShape at MIREX 2014 Symbolic Melodic Similarity. Technical report, Music Information Retrieval Evaluation eXchange, 2014

[11] Wikipedia, Needlemann-Wunsch-Algorithmus, https://de.wikipedia.org/wiki/Needleman-Wunsch-Algorithmus, aberufen am 02.02.20