

Symbolic Music Similarity

Presentation

Ali Bektas Paul Kröger

February 10, 2020

Überblick

1. Grundlegendes

2. Die Techniken

2.1 MIREX 2014

2.2 Urbano MelodyShape

3. MIREX : Algorithmen treten gegeneinander an

3.1 Ground Truth

3.2 Average Dynamic Recall

4. Bibliographie

Darstellung von Noten

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



Figure: Source: [?]

Symbolic Music Similarity

└ Grundlegendes

└ Darstellung von Noten

Darstellung von Noten

- Melodie : "singbare, in sich geschlossene Folge von Tönen" [7]
- Harmonie : "wohlklingender Zusammklang mehrerer Töne oder Akkorde" [7]
- Schlüssel : "dient in der Musiknotation dazu, im Notensystem festzulegen, welche Tonhöhe die fünf Notenslinien repräsentieren." [7]



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." [?]

Symbolic Music Similarity

└ Grundlegendes

└ 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."^[1]

1. In diesem Kontext kann "Gewicht" vieles sein: Die Position einer Note im Takt , die Länge einer Note im Takt , usw.

Darstellung von Noten

- Rhythmus
- Tonlage
- und vieles mehr

III.
ADAGIO:
4-8-8.

Adagio.



* Nach dem Bildbegriff der Ausgabe 81 der Partitur, Jahrgang 1853/54, Nr. 48.

Figure: Source: IMLSP Archive

- Rhythmus
- Tonhöhe
- und vieles mehr



Figure: Source: IMLSP Archive

1. Notendarstellung heißt nicht nur , zu welcher Zeit ein Ton gespielt wird , sondern auch , Informationen über , Gefühl beim Spielen , vortragsbetreffliche Elemente zu übermitteln.

Ein graphbasierter Ansatz



1

- Inhalt schrittweise vereinfacht.
- **Gewichte** der einzelnen Noten von Bedeutung.
 - unterliegende harmonische Funktion
 - metrische Position
 - Differenz der Tonlagen zwischen dem Ton und dem Grundton

¹"A Measure of Melodic Similarity Based on a Graph Representation of the Music Structure" [?] von Nicola Orio und Antonio Rodá.

Symbolic Music Similarity

└ Die Techniken

└ Ein graphbasierter Ansatz



- Inhalt schrittweise vereinfacht.
- Gewichte der einzelnen Noten von Bedeutung.
 - unterliegende harmonische Funktion
 - metrische Position
 - Differenz der Tonlagen zwischen dem Ton und dem Grundton

¹ "A Measure of Melodic Similarity Based on a Graph Representation of the Music Structure" [1] von Nicola Orio und Antonio Rodà.

1. Die Modelle die sich mit der Wahrnehmung von Musik beschäftigt , geht davon aus dass wir Melodien nicht so speichern, wie sie sind sondern vereinfachen wir sie , behalten nur Merkmale.

Ein graphbasierter Ansatz

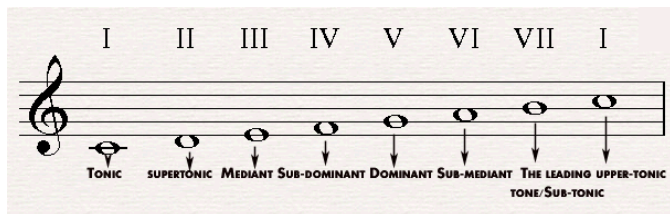


Figure: Funktionen der Noten im Tonleiter[?]



Figure: Funktionen der Noten im Tonleiter[7]

1. Tonic harmonisch relevanter als Dominant und Dominant als Sub-Dominant usw.

Ein graphbasierter Ansatz

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

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

1. Jeder Takt wird in jedem Schritt dadurch vereinfacht , dass einige Noten eliminiert sind , und die Bleibenden um die Länge der Eliminierten erweitert werden.
2. Diese Methode heißt Pseudo-Structural Representation (PSR)

Ein graphbasierter Ansatz

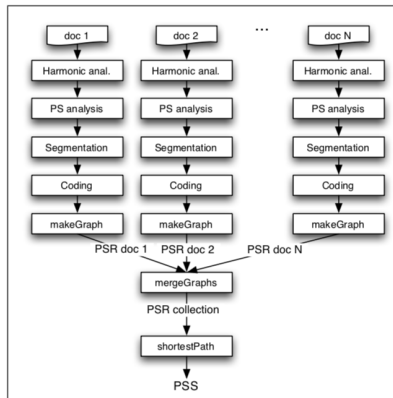


Figure: Ablauf des gesamten Verfahren [?]

Ein graphbasierter Ansatz

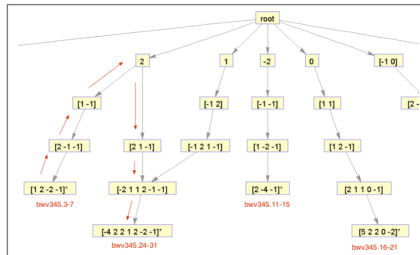
level 4

level 3

level 2

level 1

surface



Ein graphbasierter Ansatz

$$s(c_i, q) = \left(1 + \frac{d(c_i, q)}{\sum_{j=1}^N \frac{d(c_i, c_j)}{N-1}} \right)^{-1}$$

└ Die Techniken

└ Ein graphbasierter Ansatz

$$s(c_i, q) = \left(1 + \frac{d(c_i, q)}{\sum_{j=1}^N \frac{d(c_i, c_j)}{N-1}} \right)^{-1}$$

1. $d(c_i, c_j)$: Wir gucken , was die Distanzen zwischen Segmenten von den beiden Dokumenten sind.
2. Nehmen dann den Median und der Medianwert bildet dann die Distanz. Dieser Wert wird dann normalisiert , indem er durch die durchschnittliche Distanz des Dokuments c_i zu allen anderen Dokumenten in der Sammlung geteilt wird. Die Werte zur Normalisierung können im Voraus berechnet werden.

Evaluierung

- RISM-Sammlung
- Basiswissen von Experten als Maßstab

LBDM

- Local Boundary Detection Model
- Change Rule (CR) : Je größer die Differenz , desto wahrscheinlicher wird die Nichtzusammengehörigkeit.
- Proximity Rule (PR) : Change Rule angewandt auf Intervalle.

Ähnliche Anwendung der Gestaltstheorie



2

- Implication/Realization Model.
- Dies besagt, dass man nach seinen Erfahrungen (sowohl kulturellen, als auch angeborenen) Erwartungen hat, wie ein Musikstück weitergeht.
- Wir beschäftigen uns hier mit den angeborenen Aspekten.

¹"Melody Retrieval using the Implication/Realization Model" [?] Maarten Grachten, Josep Lluís Arcos and Ramon López de Mantaras

Ähnliche Anwendungen der Gestaltstheorie

- I/R Modell besagt : Wir sind dazu geneigt , Elemente nach Konzepten zu gruppieren. Diese Konzepten sind denen der Gestalttheorie ähnlich
 - Proximity : Werden zwei Elemente gleich wahrgenommen?
 - Similarity : Haben zwei Elemente Ähnlichkeiten?
- PRD : kleines Intervall in eine Richtung impliziert noch ein Intervall in dieselbe Richtung
- PID : kleines Intervall impliziert ein kleines Intervall.
- Nach diesen Prinzipien ist ein Alphabet von Strukturen definiert.
- Mithilfe von Edit Distance wird die Ähnlichkeit festgestellt.

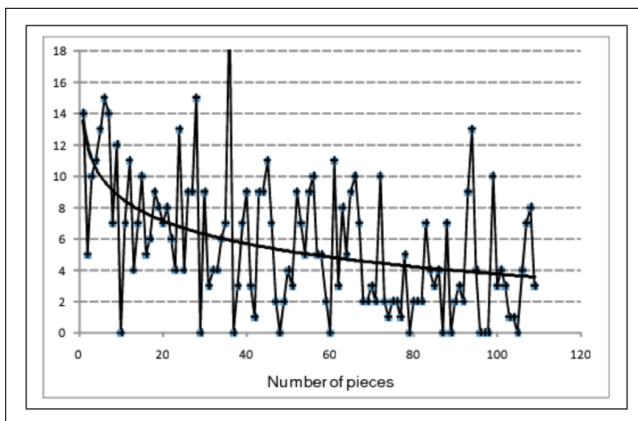
Evaluierung

# symbols	ADR	AP	R-P
3	0.65	0.60	0.54
5	0.66	0.60	0.52
7	0.65	0.59	0.51
no quantization	0.67	0.64	0.56

segmentation	ADR	AP	R-P
manual	0.67	0.64	0.56
gestalt	0.69	0.64	0.55
probabilistic	0.67	0.61	0.53
LBDM	0.61	0.53	0.50

Diskussion

- Sublineares Wachstum des Baumes
- Manuelle Annotierung der Akkorde



Ein mathematischer Ansatz

"Algorithms for Computing Geometric Measures of Melodic Similarity" [?]

von Greg Aloupis, Thomas Fevens, Stefan Langerman, Tomomi Matsui, Antonio Mesa, Yurai Nunez, David Rappaport, and Godfried Toussaint



Ein mathematischer Ansatz

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

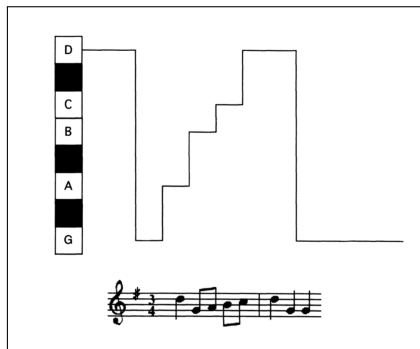


Figure: Source: [?]

Ein mathematischer Ansatz

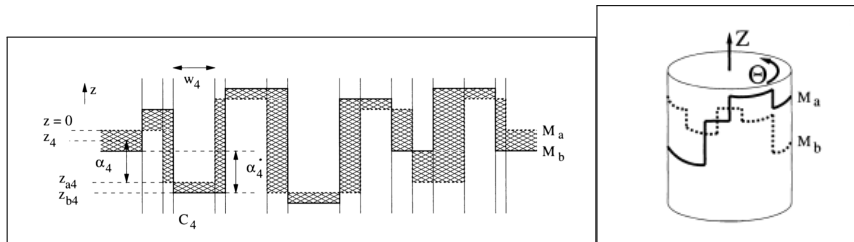


Figure: Source: [?]

Inhaltsübersicht

1. Grundlegendes

2. Die Techniken

2.1 MIREX 2014

2.2 Urbano MelodyShape

3. MIREX : Algorithmen treten gegeneinander an

3.1 Ground Truth

3.2 Average Dynamic Recall

4. Bibliographie

Ähnlichkeitssuche durch Pattern Mining



1

- Nur Note-On Events
- Dauer einer Note spielt keine Rolle.
- Grundton spielt keine Rolle. Es werden die Differenzen zwischen Tonlagen in Betracht gezogen.
- Jede Melodie wird durch primitive 'Items' dargestellt.

¹"MIREX 2014 Symbolic Melodic Similarity : Extracting Similar Melodies Based on Top-N Colossal Pattern Mining" [14] von Shihō Sugimoto, Yuto Nakashima, Masayuki Takeda.

Symbolic Music Similarity

Die Techniken

MIREX 2014

Ähnlichkeitssuche durch Pattern Mining



- Nur Note-On Events
- Dauer einer Note spielt keine Rolle.
- Grundton spielt keine Rolle: Es werden die Differenzen zwischen Tonlagen in Betracht gezogen.
- Jede Melodie wird durch primitive 'Items' dargestellt.

¹ "MIREX 2014 Symbolic Melodic Similarity: Extracting Similar Melodies Based on Top-N Colonial Pattern Mining" [14] von Shiko Sugimoto, Yuto Nakashima, Masayoshi Taketaka

1. Obwohl es nachgewiesen ist , dass die Tonlage wichtigsten Aspekt der Wahrnehmung der Musik bildet , ist es immer noch keine gute Idee die restlichen Aspekte komplett zu vernachlässigen.

Ähnlichkeitssuche durch Pattern Mining

- Einer Melodie werden alle N-Gramme entnommen.
- $TDB_M = \{(ID(x), trans(x)) | x \in M\}$
- $TDB_M[Q]$
- $X = \{X_1, \dots, X_m\}$ die Menge der Matches in der Datenbank , also $X_i \subset Q$
- Hole $M(X_i) \{x | x \in M \wedge X_i \subset trans(x)\}$

└ Die Techniken

└ MIREX 2014

└ Ähnlichkeitssuche durch Pattern Mining

- Einer Melodie werden alle N-Gramme entnommen.
- $TDB_M = \{(ID(x), trans(x)) | x \in M\}$
- $TDB_M[Q]$
- $X = \{X_1, \dots, X_n\}$ die Menge der Matches in der Datenbank , also $X_i \subset Q$
- Hole $M(X) = \{x | x \in M \wedge X_i \subset trans(x)\}$

1. Die entnommenen bezeichnen wir als Transaktionen und dann erstellen wir durch diese Transaktionen eine Datenbank
2. Wir extrahieren eine Menge von Items Q aus der Anfrage und dann suchen nach diesen Items in der Datenbank
3. Die Menge der geholten Melodien aus der Datenbank sind nun die Kandidaten , aus denen wir eine kleiner Menge wählen

$$\text{sim}(q, x) = \frac{1}{|Q \cap \text{trans}(x)|} \sum_{f \in Q \cap \text{trans}(x)} \text{weight}(f)$$

wobei für $f = (d_1, \dots, d_l)$

$$\text{weight}(f) = \sum_{i=1}^l |d_i|$$

$$\text{sim}(q, x) = \frac{1}{|Q \cap \text{trans}(x)|} \sum_{f \in Q \cap \text{trans}(x)} \text{weight}(f)$$

wobei für $f = \{d_1, \dots, d_l\}$

$$\text{weight}(f) = \sum_{i=1}^l |d_i|$$

1. / bezeichnet die Länge der Sequenz der Differenzen der Tonlagen

Diskussion

- Kann man rhythmische Werte vernachlässigen ?
- Was ist das richtige n für das N-Gramm?

Inhaltsübersicht

1. Grundlegendes

2. Die Techniken

2.1 MIREX 2014

2.2 Urbano MelodyShape

3. MIREX : Algorithmen treten gegeneinander an

3.1 Ground Truth

3.2 Average Dynamic Recall

4. Bibliographie

Urbano MelodyShap

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



Urbano MelodyShape

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

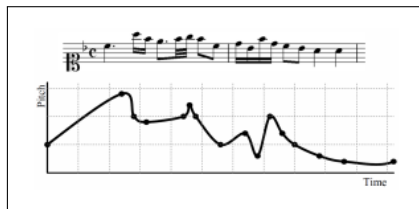


Figure: Source: [?]

Needlemann - Wunsch Algorithmus

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

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

Figure: Source: [13]

ShapeH

- Insertion :

$$s(-, n) = -(1 - f(n))$$

- Deletion:

$$s(n, -) = -(1 - f(n))$$

- Match:

$$s(n, n) = 1 - f(n)$$

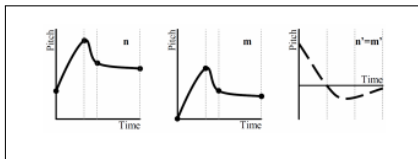


Figure: Source: [?]

Symbolic Music Similarity

└ Die Techniken

└ Urbano MelodyShape

└ ShapeH

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

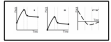


Figure: Source: [9]

1. Hybrid Alignment Approach, heißt lokales maximum wird genommen nicht globales.
2. Lokales maximum dann similarity score
3. Bei mismatch folgende Malusregeln :
4. 1.Ableitung betrachtet, vorzeichen der beiden Melodien am Anfang und Ende identisch kaum malus
5. Anfang oder Ende unterschiedlich, anderes gleich größere Malus
6. Beides unterschiedlich größter Malus

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: $s(n, m) = -diff_p(n, m) - \lambda k_t * diff_t(n, m)$

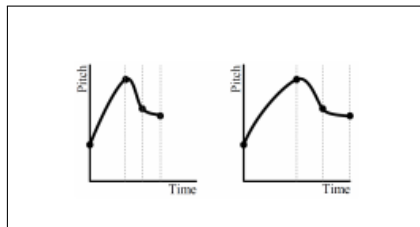


Figure: Source: [?]

MIREX

- Ein Wettbewerb und Plattform für Interessierte
- 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**
 - ...
- Welche Messmethoden gibt es , um den Erfolg eines Algorithmus festzustellen?

MIREX

SCORE	JU1	JU2	JU3	YO1
ADR	0.7089	0.7962	0.7997	0.6912
NRGB	0.6786	0.7493	0.7602	0.6378
AP	0.7344	0.7534	0.7992	0.5535
PND	0.7361	0.7444	0.7611	0.5611
Fine	53.7767	54.5967	51.1933	36.9633
PSum	1.1167	1.13	1.1267	0.69
WCSum	1.5033	1.5133	1.5433	0.93667
SDSum	1.89	1.8967	1.96	1.1833
Greater0	0.73	0.74667	0.71	0.44333
Greater1	0.38667	0.38333	0.41667	0.24667

Figure: Source: [?]

Inhaltsübersicht

1. Grundlegendes

2. Die Techniken

2.1 MIREX 2014

2.2 Urbano MelodyShape

3. MIREX : Algorithmen treten gegeneinander an

3.1 Ground Truth

3.2 Average Dynamic Recall

4. Bibliographie

Ground Truth

- Experten werden befragt , Stücke aus der RISM A/II Sammlung nach deren Ähnlichkeiten zu einer Anfrage zu beurteilen.
- Die Sammlungen sind groß deswegen sind einige Techniken zur Eliminierung unrelevanter Elementen vorzunehmen , wie z.B.
 - Nach der Differenz zwischen dem tiefsten und höchsten Ton.
 - Nach dem Verhältnis der kürzesten Note zu der längsten.
 - usw.
- Nicht für alle Stücke werden dieselben Eliminierungsverfahren vorgenommen. Die Aspekte , durch die sich ein Stück auszeichnet sind beizubehalten. Das ist wiederum für die Experten zu entscheiden.

Ground Truth I



Query: Anonymus: Roslin Castle, RISM A/II signature: 800.000.193

1



Anonymus: Roslin Castle. 000.109.446

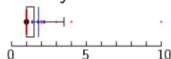


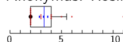
Figure: Abbildung: Ergebnisse der Befragung [?]

Ground Truth II

2



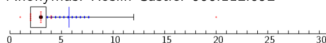
Anonymus: Roslin Castle. 000.111.779



3



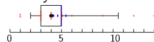
Anonymus: Roslin Castle. 000.112.692



4

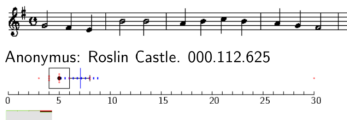


Anonymus: Roslin Castle. 000.132.330

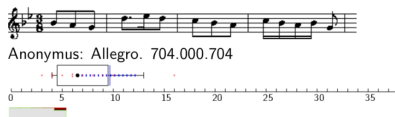


Ground Truth III

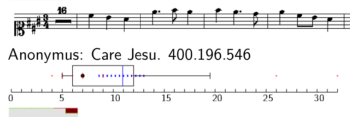
5



6.5



7



MIREX

SCORE	JU1	JU2	JU3	YO1
ADR	0.7089	0.7962	0.7997	0.6912
NRGB	0.6786	0.7493	0.7602	0.6378
AP	0.7344	0.7534	0.7992	0.5535
PND	0.7361	0.7444	0.7611	0.5611
Fine	53.7767	54.5967	51.1933	36.9633
PSum	1.1167	1.13	1.1267	0.69
WCSum	1.5033	1.5133	1.5433	0.93667
SDSum	1.89	1.8967	1.96	1.1833
Greater0	0.73	0.74667	0.71	0.44333
Greater1	0.38667	0.38333	0.41667	0.24667

Figure: Source: [?]

Inhaltsübersicht

1. Grundlegendes

2. Die Techniken

2.1 MIREX 2014

2.2 Urbano MelodyShape

3. MIREX : Algorithmen treten gegeneinander an

3.1 Ground Truth

3.2 Average Dynamic Recall

4. Bibliographie

Beispiel : Average Dynamic Recall - ADR

Betrachte die Gruppierungen $\langle (1, 2), (3, 4, 5) \rangle$ und die Ergebnisse $(2, 3, 1, 5, 7, 8, 9, 4)$

Pos.	encountered	relevant	#found	recall
1	2	1, 2	1	1
2	2, 3	1, 2	1	0.5
3	2, 3, 1	1, 2, 3, 4, 5	3	1
4	2, 3, 1, 5	1, 2, 3, 4, 5	4	1
5	2, 3, 1, 5, 7	1, 2, 3, 4, 5	4	0.8

Figure: Abbildung: ADR Berechnung [?]

five_point_three J. Urbano, *MIREX2013 Symbolic Melodic Similarity* A Geometric Approach, Cuadrado, *MIREX2012 Symbolic Melodic Similarity*, *Music Information Retrieval*

Wikipedia, Needleman-Wunsch-Algorithmus,
<https://de.wikipedia.org/wiki/Needleman-Wunsch-Algorithmus>,
abgerufen am 02.02.20

Okubo Yoshiaki , Haraguchi Makoyo , " MIREX 2014 Symbolic Melodic Similarity : Extracting Similar Melodies Based on Top-N Colossal Pattern Mining". Technical report, Music Information Retrieval Evaluation eXchange, 2014