

Graphenbasierte Musik-Klassifizierung

Studienarbeit

Abteilung Informatik

Herbstsemester 2013

Autoren: Christian Fässler, Jonas Furrer

Betreuer: Prof. Ruedi Stoop

Abgabedatum: 20. Dezember 2013

Inhaltsverzeichnis

Abstract	4
1. Einleitung	5
1.1. Vorbemerkungen	5
1.2. Aufgabenstellung	5
1.3. Einführung	5
1.4. Gliederung	6
2. Musik vergleichbar machen	7
2.1. Begriff des Musikstückes	7
2.2. Musikstücke auf Graphen abbilden	8
2.3. Melodie-Extraktion: Das Skyline-Verfahren	10
2.4. Einbettung	11
2.5. Zeit-Information der Knoten	12
2.6. Überblick der Verfahren	13
2.7. Reduktion der Information	13
2.8. Bewertung der Reduktion	13
2.9. Zufallspfade im Graphen	14
3. Charakterisierung der Graphen	15
3.1. Lokale Charakteristiken	15
3.1.1. Knotengrad	15
3.2. Orbits	15
3.3. Globale Charakteristiken	17
3.3.1. Anzahl der Knoten	17
3.3.2. Entropie	18
3.3.3. Eigenwerte	18
3.3.4. Small World Phänomen	19
3.4. Bereich der Features im Vergleich	20
3.5. Featureeigenschaften	22
4. Clustering der Graphen	23
4.1. Datenmaterial	23
4.2. Erwartungen an das Resultat des Clusterings	23
4.3. Bewertung der Resultate	24
4.4. HLC Clustering Parameter	25
4.5. Ergebnisse	26
4.6. Interpretation	27
5. Weiterführende Ideen	32

Anhang	34
A. Aufgabenstellung	35
B. Musikstücke	38
C. Clusteringergebnisse (Details)	40
C.1. Standard Features	40
C.2. Skylined Features	41
C.3. Timing Features	42

Abstract

Diese Studienarbeit beschreibt, wie man Musikstücke, welche im MIDI Format vorliegen, auf komplexe Netzwerke abbildet, um sie mit mathematischen Hilfsmitteln klassifizieren zu können. Die MIDI Dateien werden mit Hilfe eines Samplingverfahrens in Vektoren von gleichzeitig erklingenden Tönen unterteilt. Die Vektoren werden dedupliziert und zu einem gerichteten Graphen verbunden. Die gerichteten Kanten des Graphen geben die Reihenfolge vor, in der die Vektoren durchlaufen werden müssen.

Die so entstandenen Netzwerke werden dann mit Werkzeugen der Graphentheorie beschrieben. Dabei konnte festgestellt werden, dass es sich bei den Musik-Graphen um Small World Netzwerke handelt und dass die Graphen Eigenschaften von komplexen Netzwerken aufweisen. Die Charakteristiken der Graphen werden sodann pro Musikstück in Feature-Vektoren zusammengefasst, welche zusammen mit den gerichteten Kanten des Graphen das Musikstück charakterisieren.

Um festzustellen, ob die gewählten Features wirklich die Musikstücke charakterisieren, wurde ein Clustering der Feature-Vektoren durchgeführt, mit dem Ziel, in den entstehenden Clustern die Genres der Musikstücke wiederzufinden. Obwohl das Clustering gute Resultate liefert, stimmten die Musikgenres nicht vollumfänglich mit den Clustern überein. Wir vermuten, dass das Clustering eher eine Gruppierung nach der Komplexität der Stücke liefert, welche nicht mit der Einteilung in Genres übereinstimmen muss.

1. Einleitung

1.1. Vorbemerkungen

Bei diesem Dokument handelt es sich um die Studienarbeit von Christian Fässler und Jonas Furrer, erstellt an der Technischen Hochschule Rapperswil im Studiengang Informatik. Betreut und begleitet durch Prof. Ruedi Stoop, Institute of Neuroinformatics, Universität Zürich.

Sämtliche Aufgaben zu dieser Arbeit wurden von beiden Studenten zu gleichen Teilen erarbeitet. Für Berechnungen und Auswertungen wurde Wolfram Mathematica verwendet. Die Arbeit wurde innerhalb von 14 Wochen im Zeitraum vom 16. September 2013 bis zum 20. Dezember 2013 durchgeführt.

1.2. Aufgabenstellung

Es geht darum, Musikstücke als komplexe Harmonie-Netzwerke darzustellen. Diese Netzwerke sollen in der Sprache der Physik charakterisiert und beschrieben werden. Konkret heisst dies, dass MIDI-Files als Graphen dargestellt werden.

Dabei sollen die gleichzeitig vorkommenden Noten als Knoten eingesetzt werden und die Kanten die zeitliche Abfolge repräsentieren. Die so erhaltenen Graphen sollen mit moderner Graphentheorie analysiert werden, um Gemeinsamkeiten und Unterschiede herauszuarbeiten und festzuhalten.

1.3. Einführung

Musik existiert seit es Menschen gibt und sie hat sich zusammen mit dem Menschen verändert und entwickelt. Über Musik wird geredet, diskutiert und gestritten. Die Gefühle und Assoziationen, die Musik auslösen kann, sind individuell und subjektiv. Trotz unterschiedlicher Wahrnehmung wird immer wieder versucht, Musik zu kategorisieren und zu klassifizieren. Bei einigen Musikstücken kann eine Kategorisierung vorgenommen werden, die übereinstimmend als richtig angesehen wird. Bei anderen Stücken ist keine eindeutige Kategorisierung möglich, da die Meinungen der Hörer zu weit auseinander gehen.

Es stellt sich nun die Frage, ob man Musik derart mathematisch beschreiben kann, so dass eine deutliche Kategorisierung möglich ist. Obwohl in diesem Bereich geforscht wird, wie zum Beispiel in der Arbeit von J. Serra et al. 2012 [1], sind die Möglichkeiten der automatischen Analyse von Klangmaterial noch nicht so weit fortgeschritten wie beispielsweise die Analyse von Bildmaterial. Im Folgenden wird versucht, die Analyse und möglicherweise eine Kategorisierung von Musik aufgrund von komplexen Netzwerken vorzunehmen.

Kann Musik auf einen Graphen reduziert werden?

Die zentrale Frage unserer Arbeit ist, ob Musik als Graph dargestellt werden kann und ob das Modell eines Graphen dafür geeignet ist, Musik zu repräsentieren. Das Gebiet der Graphentheorie hat in den letzten 20 Jahren das Interesse der wissenschaftlichen Forschung angezogen und in vielen Gebieten zu neuen, fundamentalen Erkenntnissen geführt.

Ausgelöst wurden diese Fortschritte nicht zuletzt durch das erhöhte Interesse an der Social Network Analyse [11], welche heutzutage, in der Zeit von Social-Media Plattformen, ein grosses Anwendungsgebiet hat. Diese Erkenntnisse haben ihre Bedeutung auch in der Physik und in der Mathematik.

« Individuals are, as it were, tied to one another by invisible bonds which are knitted together into a criss-cross mesh of connections, much as a fishing net or a length of cloth is made from intertwined fabrics.» Scott, 1988 [4]

Entsprechend möchten wir untersuchen, ob - analog zu den Strukturen in Sozialen Netzwerken - auch bei der Musik in Form von Netzwerken solche Strukturen zu identifizieren sind und in wie fern sich das Konzept von komplexen Netzwerken auch auf die Musik anwenden lässt.

1.4. Gliederung

Diese Arbeit lässt sich grob in drei Teile gliedern. Im ersten Teil beschreiben wir, wie wir aus der vorliegenden Musikinformation, in Form von MIDI Dateien, Vektoren erstellen und diese anschliessend als Knoten in einem komplexen Netzwerk darstellen. Im zweiten Teil extrahieren wir, basierend auf den Charakteristiken der erstellten Netzwerke, Feature-Vektoren und versuchen so, die Stücke zu beschreiben. Im letzten Teil untersuchen wir mittels Clustering, wie gut die erstellten Feature-Vektoren die Musikstücke charakterisieren.

2. Musik vergleichbar machen

Um Musik analytisch oder numerisch vergleichbar zu machen, muss die vorliegende Klanginformation reduziert und in eine Repräsentation gebracht werden, welche mit mathematischen Mitteln greifbar wird.

Eine Modellierung mittels Graphen schafft diese Voraussetzungen und ist damit ein geeignetes Verfahren: Graphen bieten die Möglichkeit, Beziehungen auf einfache Art und Weise darstellen zu können, zum Beispiel die zeitliche Abfolge von Tönen in der Musik. Somit scheint uns die Wahl von Graphen als Mittel zur Abbildung geeignet.

Mit der Graphentheorie existiert zudem ein mächtiges Instrument zur Beschreibung, Analyse und Vergleich von Graphen. Als Quellformat der zu untersuchenden Musik eignet sich das MIDI-Format [3], da in MIDI Dateien die Abfolge der erklingenden Töne als Notenfolge bereits vorhanden ist. Es entfällt eine aufwändige Frequenz- und Spektrum-Analyse, wie sie bei Musik in Form von PCM-Rohdaten notwendig wäre. Im Unterschied zu einer textuellen Darstellung, zum Beispiel in Form von Notenblättern oder XML-Dateien, lässt sich eine MIDI Datei abspielen und erlaubt es, gefundene Phänomene durch Anhören besser zu verstehen.

2.1. Begriff des Musikstückes

Als erstes muss man sich auf ein Verständnis der zu betrachtenden Musik einigen.

Musik besteht nicht nur aus einer Abfolge von Tönen, sondern sollte mehrdimensional betrachtet werden. Neben der grundlegenden Tonabfolge spielt der Rhythmus eine tragende Rolle, auch lassen sich in den Tonabfolgen bestimmte Regeln oder Muster finden [2] und so zum Beispiel Tonleitern zuordnen. Weitere Betrachtungsweisen sind das Frequenzspektrum oder Amplitudenverhältnisse; Geschwindigkeit, Pausen und Takte verändern ebenfalls das Empfinden der Musik. Ein weiterer Punkt ist, dass ein Stück durch die Interpretation, die Wahl und Anzahl der Instrumente einen komplett anderen Klang erhalten kann und trotzdem erkennbar bleibt, solange gewisse Charakteristiken beibehalten werden. Daraus lässt sich schliessen, dass einige dieser Charakteristiken mehr zum Wesen des Stückes beitragen als andere.

Für unsere Analysen versuchen wir, die betrachteten Stücke auf Charakteristiken zu reduzieren, die uns wesentlich erscheinen.

Um diese Charakteristiken zu finden, haben wir versucht herauszufinden, warum man Musikstücke trotz unterschiedlicher Interpretation dennoch wiedererkennen kann. Unter Interpretation verstehen wir beispielsweise eine Transponierung in eine andere Tonart, die Verwendung von anderen Instrumenten, oder die Veränderung der Abspielgeschwindigkeit.

Wahrnehmungsmodell

Ein Gedankenspiel hilft uns dabei, ein Modell für die Identifizierung wichtiger Charakteristiken zu finden: Gegeben sei ein Spiel mit dem Ziel, Musikstücke möglichst schnell anhand von ge-

summierten Melodien zu erraten.

Der Summende versucht, das Erraten der Melodie möglichst einfach zu gestalten. Das geschieht sinnvollerweise durch Summen der markantesten Teile eines Stückes, den Teilen mit dem höchsten Wiedererkennungswert. Um nun herauszufinden, um welches Stück es sich handelt, wird der Ratende die gesummierten Intervalle mit ihm bekannten Melodien abgleichen.

Folgendes lässt sich dabei festzustellen: Bei den gesummierten Teilen handelt es sich um vergleichsweise kurze Teile des Gesamtstückes, welche somit als besonders einprägsam oder charakteristisch bezeichnet werden können. Beim Summen der Teile scheint die absolute Tonlage (Tonart/Frequenz) eine untergeordnete Rolle zu spielen. Viel bedeutender sind die Intervalle und die jeweilige Tondauer.

Daraus lassen sich folgende Punkte ableiten:

- Ein Stück wird anhand der Melodie, der Abfolge der Intervalle, erkannt
- Die erkannte Melodie basiert auf einem kurzen, aber wichtigen Teil des Stückes.
- Takt und Rhythmus spielen eine untergeordnete Rolle

Das heisst, als Musikstück betrachten wir eine Musiksequenz mit Anfang und Ende, die von einem Menschen durch Anhören korrekt benannt werden kann. Von diesem Modell ausgehend versuchen wir, Musikstücke auf Netzwerke abzubilden.

2.2. Musikstücke auf Graphen abbilden

Um Musikstücke in die Form eines Graphen zu bringen, verwenden wir den folgenden Ansatz: In einem ersten Schritt werden die Perkussionsspuren entfernt, da diese keinen Beitrag zur Melodie leisten, beziehungsweise keine Tonhöheninformation aufweisen. Die Tonhöhen werden im nächsten Schritt auf ein einheitliches Spektrum normalisiert, das heisst, die Töne werden in die Oktave mit der Basis C4 gebracht, beispielsweise C#5 wird zu C#4 oder F6 wird zu F4.

Der Grund dafür ist, dass die in der westlichen Musik verwendete Tonleiter nur aus zwölf Halbtönen besteht. Die anderen Töne sind lediglich Oktaven der dieser zwölf Grundtöne. Durch diese Verdichtung lässt sich die Anzahl der gleichzeitig auftretenden Tonkombinationen wesentlich reduzieren.

Es sind so maximal noch

$$\sum_{i=1}^{12} \frac{12!}{(12-i)! \cdot i!} = 4095$$

Knoten möglich, wobei ohne diese Reduktion die Knotenmenge erheblich höher wäre:

$$\sum_{i=1}^{(x \cdot 12)} \frac{(x \cdot 12)!}{((x \cdot 12) - i)! \cdot i!} \quad \text{wobei } x \text{ die Anzahl Oktaven bedeutet}$$

Der Erwartungswert der Knotenzahl liegt im Bereich von

$$\sum_{i=1}^7 \frac{7!}{(7-i)! \cdot i!} = 127,$$

da die gängigen Skalen wie Dur oder Moll nur sieben Halbtöne besitzen.

Die Knoten des Graphen werden mit Hilfe eines Samplingverfahrens gefunden. Als Samplezeit dient die Dauer der kürzesten Note im analysierten Stück. Das Stück wird in Zeitschritten der Sampledauer durchlaufen, wobei die Kombination der gleichzeitig gespielten Noten einen Knoten ergibt.

Falls zwei identische Knoten direkt nacheinander auftreten, werden diese zu einem einzigen zusammengefügt. Demnach betrachten wir eine Abfolge von $\{\{C, F\}, \{A, E\}, \{A, E\}, \{A, E\}\}$ als einen einzigen Übergang vom Knoten $\{C, F\}$ zum Knoten $\{A, E\}$.

Sind aber zwischen den Knoten Pausen enthalten, wird jeweils ein Übergang, zu einem Pausenknoten vorgenommen. Ein Pausenknoten ist ein Knoten der aus einem leeren Vektor besteht. Falls innerhalb eines Knotens ein Note mehrfach auftritt, wird diese auf ein Vorkommen reduziert, beispielsweise $\{\{E, E, G\}, \{C, G\}, \dots\} \rightarrow \{\{E, G\}, \{C, G\}, \dots\}$.

Die so entstandene Liste von Knoten wird dann in eine Adjazenzmatrix umgewandelt. Die Matrix kann dann als Graph dargestellt und entsprechend analysiert werden.

Der nun vorliegende Graph ist gerichtet. Die Vektoren repräsentieren einen Zeitpunkt im Musikstück, die Komponenten dieser Vektoren sind die Noten, die zu diesem Zeitpunkt erklingen.

Die Kanten zeigen den Ablauf des Musikstücks, also die Übergänge von einer Tonkombination zur nächsten. Nicht zu sehen ist, wie oft ein Knoten besucht wird, sowie die Dauer der einzelnen Knoten.

In Abbildung 1 ist ein gemäss dem beschriebenen Vorgehen erhaltener Beispielgraph abgebildet.

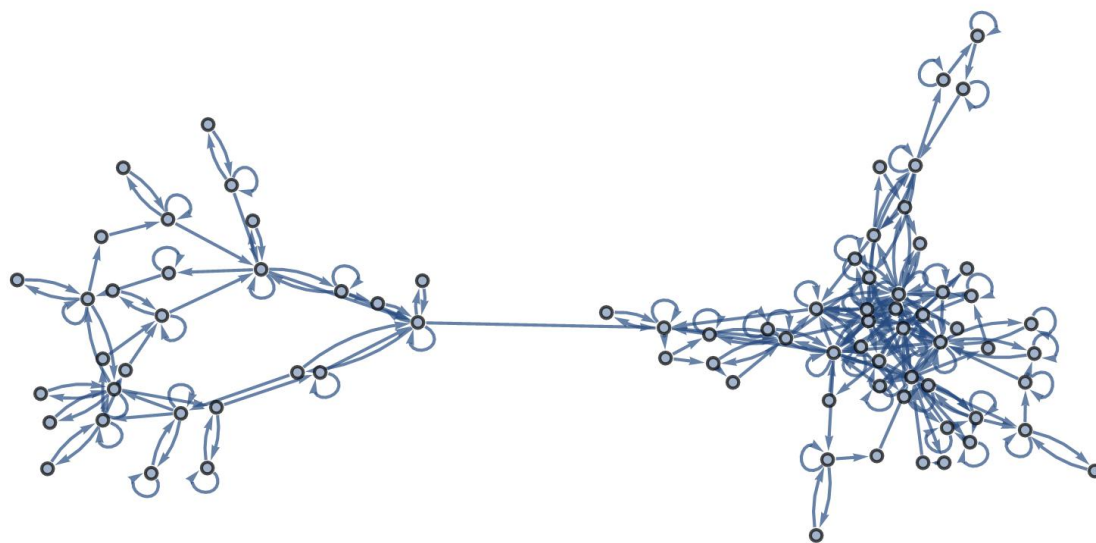


Abbildung 1: Abbildung des Stückes «Final Countdown» von *Europe* auf einen Graphen. Offenbar besteht dieser aus zwei Teilgraphen, welche durch eine einzige Verbindung aneinander gekettet sind. Der Übergang zwischen den Teilgraphen ist beim Hören nicht deutlich auszumachen.

2.3. Melodie-Extraktion: Das Skyline-Verfahren

Ein Problem beim Vergleich von Musikstücken stellt sich beim Vergleich von gleichen Stücken in unterschiedlicher Interpretation. Betrachten wir dasselbe Stück einmal als mehrspurige orchestrale Version und als einspurige Klavierversion, so ist für einen Menschen im Normalfall deutlich erkennbar, dass es sich um das gleiche Stück handelt. Die MIDI Dateien dieser beiden Versionen sind aber grundlegend verschieden. Die orchestrale Version besteht aufgrund ihrer Mehrspurigkeit tendenziell aus mehr gleichzeitig gespielten Tönen.

Um ein Stück unabhängig von seiner Interpretation untersuchen zu können, muss also das Kernmerkmal, nämlich die Melodie des Stückes, extrahiert werden.

Ein relativ einfaches Verfahren zu Melodie-Extraktion ist das Skyline-Verfahren [5].

Der Skyline-Algorithmus reduziert polyphone Musikstücke auf eine Spur. Es wird eine passende Samplezeit gewählt, mit der wir das Musikstück in Abständen der Samplezeit durchlaufen. Bei jedem Schritt wird aus den gleichzeitig erklingenden Noten diejenige der markantesten Ausprägung gewählt. Die Qualität des Ergebnisses variiert stark von Stück zu Stück und hängt von der Wahl des markanten Merkmales ab. Trotzdem ist dieses Verfahren eine einfache Variante, um die analysierten MIDI Dateien zu vereinheitlichen.

Nach dem ein Musikstück durch den Skyline-Algorithmus reduziert wurde, kann das bereits beschriebene Verfahren angewendet werden, um das Stück als Graphen darzustellen, wobei nun jeder Knoten aus genau einer Note besteht.

In unserer Implementation haben wir als Samplezeit die Dauer der kürzesten Note des Stückes

und als Merkmal für die markanteste Ausprägung die Tonhöhe gewählt; das heisst, es wird der höchste Ton zu jedem Samplezeitpunkt extrahiert.

Dies weil wir beobachtet haben, dass bei vielen Stücken die tragende Melodie durch die höchsten Töne charakterisiert wird und die Stücke häufig den so reduzierten Versionen wiedererkannt werden können.

In dieser Konfiguration geht der Skyline-Algorithmus davon aus, dass zu jedem Sample-Zeitpunkt ein Ton der Hauptmelodie gespielt wird und die Hauptmelodie stets die markanteste Ausprägung besitzt. Sind diese Voraussetzungen erfüllt, liefert das Verfahren sehr gute Ergebnisse.

Es gibt allerdings durchaus Stücke, deren Wiedererkennungswert stark durch die Hauptmelodie gegeben ist, die Hauptmelodie im Verhältnis zu anderen Spuren aber viele Pausen beinhaltet. Dadurch bleiben nach dem Anwenden des Skyline-Algorithmus diese Nebentöne dominant und die Hauptmelodie verliert an Gewicht, sodass in einigen Fällen das Stück kaum wieder zu erkennen ist.

2.4. Einbettung

Wie bereits erwähnt, gehen wir davon aus, dass gewisse Sequenzen eines Stückes mehrfach vorkommen und eine stärkere Relevanz für den Wiedererkennungswert eines Stückes aufweisen. Diese Teile sind oftmals im Vergleich zur Gesamtdauer relativ kurz.

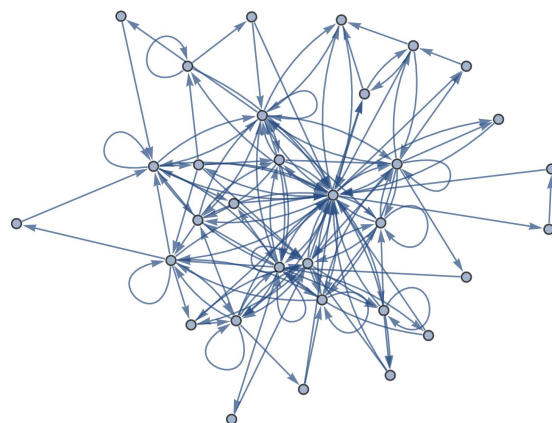
Das Ziel ist es, mit Hilfe einer Einbettung das Stück auf einer grösseren Zeitskala zu betrachten, um so Basisstrukturen hervorzuheben.

Um nicht nur eine Momentaufnahme – also jeden Knoten für sich – zu betrachten, werden mehrere aufeinander folgende Knoten zu einem neuen Knoten zusammengefasst. Die so entstehenden Knoten sind Vektoren, welche aus einer festen Anzahl aufeinanderfolgender ursprünglicher Knoten aufgebaut sind. Diese Anzahl definieren wir als Einbettungsdimension, welche für einen Einbettungsvorgang konstant bleibt.

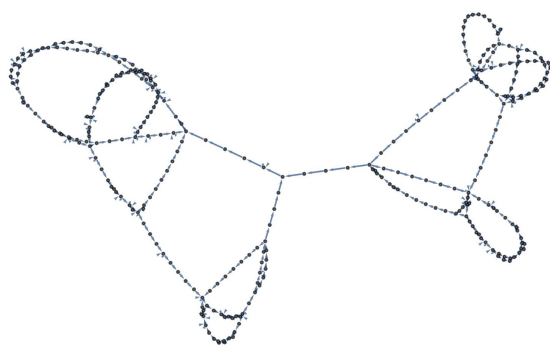
Die Zusammenfassung geschieht fortlaufend, so dass diese jeweils überlappend ist (auch bekannt als sliding window). Die Überlappung hat jeweils die Länge *Dimension* – 1, das heisst, dass beispielsweise bei einer Einbettung zur Dimension drei, drei Knoten zu einem zusammengefasst werden und sich zwei Knoten mit dem nächsten überlappen, zum Beispiel:

$$\{\{E, G\}, \{C, G\}, \{A, F\}, \{A, E\}, \{E, F\}\} \rightarrow \{\{\{E, G\}, \{C, G\}, \{A, F\}\}, \{\{C, G\}, \{A, F\}, \{A, E\}\}, \{\{A, F\}, \{A, E\}, \{E, F\}\}\}.$$

Ab einer gewissen Einbettungsdimension werden sich wiederholende Teile in Form von Orbits in den Graphen sichtbar. Beim Vergleich eines Stückes in verschiedenen Einbettungsdimensionen ist zu beobachten, dass sich die Zahl der Orbits anfangs bei jeder Dimension stark verändert, aber ab einer gewissen Dimension sich zu stabilisieren beginnt. Mehr zur Analyse der eingebetteten Stücke ist in Abschnitt 3 zu finden.



(a) Ohne Einbettung



(b) Einbettung Dimension 9

Abbildung 2: Vergleich von Graphen von «Auseer Schottisch», beim eingebetteten Graphen sind grössere Strukturen und Orbits besser zu sehen als beim Graphen ohne Einbettung.

2.5. Zeit-Information der Knoten

Weder das erste beschriebene Verfahren zur Erstellung von Graphen (mit oder ohne Einbettung), noch das Skylining-Ergebnis, enthält Informationen zur Dauer der einzelnen Knoten, beziehungsweise Tonkombinationen.

Das Hinzufügen der Dauer der einzelnen Knoten hat einige interessante Folgen. Einerseits wird grundsätzlich die Anzahl der Knoten erhöht, da gleiche Töne, die unterschiedlich lange erklingen, nun separate Knoten erzeugen. Es werden auch unterschiedlich lange Pausen, also Knoten ohne Noten, sichtbar, zuvor war nur eine Pause beliebiger Dauer vorhanden.

Sehr interessant ist, dass es durch diese Zeit-Information möglich ist, aus dem Graphen wieder eine MIDI Datei zu erstellen. Diese Möglichkeit erlaubt uns, durch Anhören zu prüfen, ob ein Stück, welches zu einem Graphen reduziert wurde und somit viel Information verloren hat, noch erkennbar ist. Diese Hörprobe ist zwar subjektiv, trotzdem aber ein relevantes Mass, da Musik letztendlich vom Menschen beurteilt wird.

Um die Dauer den einzelnen Knoten hinzuzufügen, haben wir das verwendete Verfahren (Unterabschnitt 2.2) zur Grapherstellung geringfügig ergänzt: Im Schritt, in dem aufeinanderfolgende,

identische Knoten zusammengefügt werden, wird jeweils den bestehenden Noten noch die Dauer beigelegt. Jeder Knoten enthält so nebst den bestehenden Noten, auch noch die Dauer als zusätzliche Komponente, zum Beispiel: $\{A, E, G, 1.3\}$. Der weitere Verlauf der Grapherstellung bleibt unverändert.

2.6. Überblick der Verfahren

Bisher haben wir verschiedene von uns verwendete Verfahren beschrieben. Für diese Verfahren werden wir folgende Begriffe verwenden:

Standard Erstes Verfahren, Sampling, Knoten enthalten gleichzeitig erklingende Töne

Skylined Stücke mittels Skyline-Verfahren auf eine Spur reduziert

Embedded Einbettung eines Graphen

Timing Dauer in den Knoten ergänzt

Die genannten Verfahren basieren alle auf dem «Standard-Verfahren» und können teilweise auch miteinander kombiniert werden. Skylining wird vorgängig auf den MIDI Dateien durchgeführt und kann zusätzlich noch eingebettet und mit Zeitinformationen ergänzt werden. Graphen mit Zeitinformation können auch zusätzlich noch eingebettet werden.

2.7. Reduktion der Information

Die Überführung eines Musikstücks in ein Netzwerk, gemäss beschriebenem Verfahren, basiert auf dem Wahrnehmungsmodell in Abschnitt 2.1 und vernachlässigt dabei gewisse Informationen gänzlich oder reduziert diese teilweise:

- Die Tonhöhen werden auf eine Oktave reduziert, ein Teil der Tonhöhen-Information geht verloren
- Perkussion, falls vorhanden, wird entfernt
- Das Stück wird mit einem Samplingverfahren analysiert, sodass die Gefahr besteht, dass einzelne Noten oder Pausen verloren gehen
- Der Graph hat keine Gewichtung, es ist nicht klar, wie oft ein Knoten durchlaufen wird
- Die Information über die gespielten Instrumente werden ignoriert

2.8. Bewertung der Reduktion

Es stellt sich die Frage, ob das Netzwerk nach der Reduktion das Musikstück noch genügend gut repräsentiert, oder letztlich, wie gut unser Wahrnehmungsmodell die Musikwahrnehmung

beschreibt.

Um einen Hinweis zu erhalten, ob trotz der Informationsreduktion noch genügend Information enthalten ist, haben wir die Liste der reduzierten Knoten mit Zeit-Information wieder in eine MIDI Datei zurück gewandelt und durch Anhören versucht, das jeweilige Stück dem korrekten Titel zuzuordnen.

Tests mit rund 30 Stücken haben gezeigt, dass es möglich ist, alle reduzierten Stücke problemlos zu identifizieren. Dieses Ergebnis erlaubt es uns, zu sagen, dass die von uns durchgeführte Reduktion der Stücke in einem sinnvollen Mass geschieht. Dies erlaubt den Schluss, dass die Graphen das Musikstück mit genügender Genauigkeit darstellen können.

Wichtig zu erwähnen ist, dass für die Tests die korrekte Traversierung der Graphen bekannt war, diese aber aus dem Graph alleine nicht ersichtlich ist.

2.9. Zufallspfade im Graphen

Wenn man den Graphen eines Stückes betrachtet, sieht man alle Tonkombinationen, welche im Stück vorkommen. Das Stück entsteht, indem der Graph in einer ganz bestimmten Reihenfolge durchlaufen wird. Alle anderen Pfade auf dem Graphen ergeben eine andere Tonabfolge, die zwar aus demselben Tonmaterial besteht, aber nicht mehr der Reihenfolge des ursprünglichen Musikstücks entspricht.

Nun stellt sich die Frage: Was sind die Eigenschaften eines «zufälligen» Stückes, ist es eventuell möglich, künstliche Stücke zu erschaffen, die für einen Hörer einen guten Klang besitzen?

Aus Zeitgründen war es nicht möglich, uns mit dieser Fragestellung genauer zu beschäftigen, allerdings haben wir trotzdem versucht, zufällige Stücke zu erstellen.

Dazu haben wir einen Random-Walk auf dem Graphen simuliert, indem wir nach folgendem Algorithmus Knoten aus der Knotenliste mit Zeit-Information gewählt haben:

Als Startpunkt wird zufällig ein Knoten aus der Liste gewählt, danach werden alle Vorkommnisse dieses Knotens in der Liste gesucht und deren Nachfolger gesammelt. Aus den Nachfolgern wird wieder zufällig ein Knoten ausgesucht, die ausgesuchten Knoten werden einer neuen Liste hinzugefügt, welche dann das zufällige Stück repräsentiert. Danach wird der Schritt mit der zufälligen Wahl eines Nachfolgers wiederholt, bis die gewünschte Länge erreicht ist.

Dieser Algorithmus ist sehr einfach und hat ein hohes Verbesserungspotenzial. Es ist uns zwar gelungen, mit diesem Vorgehen einige interessant klingende Stücke zu generieren, allerdings ist zu bemerken, dass die Samplerate einen starken Einfluss auf das Ergebnis hat.

Trotz allen Mängeln sind die zufällig generierten Stücke eine interessante Grundlage für weitere Analysen.

3. Charakterisierung der Graphen

Die vorliegenden Graphen können nun mit Hilfe der Mathematik, im Speziellen der Graphentheorie, untersucht und allfällige Strukturen und Gemeinsamkeiten erkannt werden. Wir gehen davon aus, dass es sich bei den erstellten Graphen nicht um Zufallsgraphen, sondern um komplexe Netzwerke, wie von Steven Strogatz in *Exploring complex networks* [6] beschrieben, handelt. Solche komplexe Netzwerke weisen diverse interessante Charakteristiken auf, welche sich grob in Lokal- und Global-Charakteristiken einteilen lassen.

Im Rahmen dieser Arbeit haben wir aus den Netzwerken von 42 Stücken Charakteristiken extrahiert und betrachtet.

Aus diesen Charakteristiken liess sich für jedes Stück ein Feature-Vektor erstellen, welcher im letzten Teil der Arbeit dazu dient, das Stück zu beschreiben und zu klassifizieren. Eine Liste der verwendeten Stücke ist im Anhang zu finden.

3.1. Lokale Charakteristiken

3.1.1. Knotengrad

Der Grad eines Knotens im Netzwerk sagt etwas über dessen Zentralität aus. Je höher der Grad, desto besser ist seine Verbindung zum Gesamtnetz. Im Zusammenhang mit dem Grad betrachten wir folgende Werte:

- Durchschnittlicher Knotengrad (eingehende und ausgehende Verbindungen)
- Standardabweichung des Knotengrades. Dieses Mass weist auf die Verteilung der Zentralitäten hin. Es zeigt das Verhältnis von zentralen Knoten zu dezentralen Knoten.

Knotengrad Verteilung

Die Verteilung der Knotengrade ist nicht linear, sondern folgt einem Potenzgesetz - es gibt wenig Knoten mit hohem Grad (hohe Zentralität) und viele Knoten mit niedrigem Grad. Diese Verteilung ist ein typisches Merkmal von Small-World Netzwerken (siehe Unterunterabschnitt 3.3.4).

3.2. Orbits

Da Orbits weder rein lokalen noch rein globalen Charakter haben, sondern dazwischen einzuordnen sind, ist dieses Feature nicht wie die anderen Features in global oder lokal einteilbar.

Als Orbits betrachten wir geschlossene Wege auf den Graphen, welche sich nicht in kleinere geschlossene Wege zerlegen lassen; diese Wege müssen kantendisjunkt sein und die Pfadlänge muss grösser als zwei sein.

Die Beobachtung der Orbits ist besonders bei den eingebetteten Graphen interessant, denn bei der kontinuierlichen Einbettung in höhere Dimensionen werden gewisse sich wiederholende Strukturen in der Form von Orbits sichtbar, siehe zum Beispiel Abbildung 2 und Abbildung 3.

Es stellt sich nun die Frage, bei welcher Einbettungsdimension das Musikstück am besten repräsentiert wird und die Anzahl der Orbits die höchste Aussagekraft hat.

Je höher die Einbettungsdimension, desto eher fallen die Orbits zusammen. Im Extremfall ist die Einbettungsdimension so hoch wie die Anzahl Knoten des Graphen und der resultierende Graph besteht lediglich aus einem einzigen Knoten.

In Abbildung 3 sieht man, dass sich im Bereich gewisser Einbettungsdimensionen die Struktur des Graphen nicht mehr wesentlich zu ändern scheint: Die Struktur stabilisiert sich, bis der Graph zu einer Kette degeneriert.

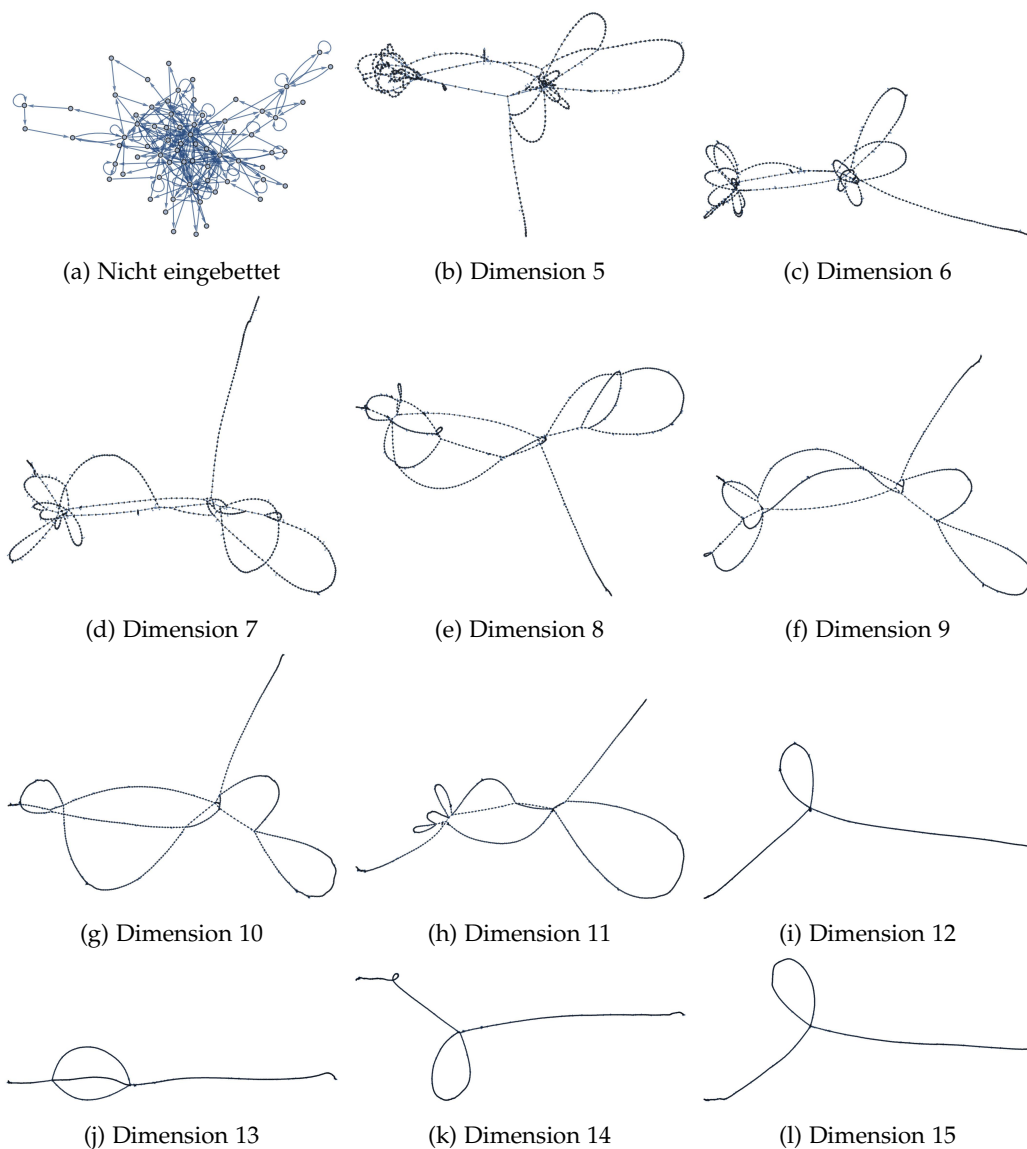


Abbildung 3: Kontinuierliche Einbettung von «ACDC Highway to Hell».

Diese auftretenden Stabilitäten sind ein interessantes Phänomen und es scheint möglich zu sein, mit Hilfe der Anzahl Orbits den Bereich der Einbettungsdimension zu finden, ab welchem diese Stabilität eintritt.

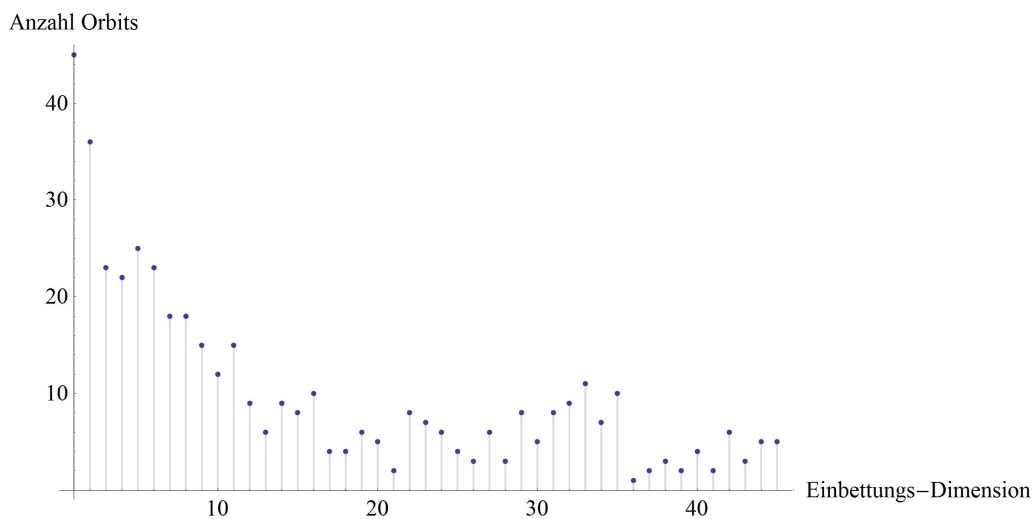


Abbildung 4: Kontinuierliche Einbettung bis zur Dimension 45 von «Highway to Hell» von ACDC. Man sieht, dass die Anzahl Orbits schnell sinkt und dann in einem ähnlichen Bereich verbleibt.

Bei der Betrachtung der Anzahl Orbits über aufsteigende Einbettungsdimensionen, wie in Abbildung 4 zu sehen, sind die stabilen Bereiche sichtbar. Das Diagramm zeigt die Anzahl der Orbits (Y-Achse) bei der jeweiligen Einbettungsdimension (X-Achsenwert). Der Abstieg in den geringen Einbettungsdimensionen ist sehr hoch, mit zunehmender Dimension flacht er allerdings ab, bis er sich Null annähert, da bei genügend grosser Einbettung der Graph zu einer Kette degeneriert, welche keine Orbits mehr beinhalten kann.

Aufgefallen ist uns auch, dass die Anzahl der Orbits bei diversen Stücken nach einem ersten Minimum noch einmal stark ansteigt. Wir konnten den Grund dafür nicht feststellen, die Vermutung ist, dass solche Anstiege bei Vielfachen von Dimensionen auftreten, welche zu einer optimalen Anzahl von Orbits führen. Diese Vermutung konnten wir allerdings, aus Zeitgründen, nicht mehr bestätigen.

3.3. Globale Charakteristiken

3.3.1. Anzahl der Knoten

Eine einfache Charakteristik ist die Anzahl der Knoten im Netzwerk. Hier ist zu beachten, dass einerseits eine Abhängigkeit zur Anzahl der Spuren und andererseits zur Gesamtdauer des Stückes besteht. Die Anzahl gleichzeitiger Töne (bei mehreren Spuren) beeinflusst die Knotenzahl direkt, siehe dazu auch die Berechnungen in Unterabschnitt 2.2. Die Dauer eines Musikstücks

beeinflusst die Knotenzahl nur indirekt, da die Knoten mehrfach besucht werden können. Die Knotenzahl zeigt daher die Diversität der verwendeten Tonkombinationen. Die durchschnittliche Knotenzahl über alle 42 Stücke beträgt 121.7 und liegt damit nahe den erwarteten 127 (siehe Unterabschnitt 2.2).

Knotenzahl und Skylining

Beim Analysieren von Stücken, die mit dem Skyline-Verfahren reduziert wurden, gibt die Anzahl der Knoten einen Hinweis darauf, wie gross die verwendete tonale Skala ist, beziehungsweise, ob vielleicht mehrere Skalen im selben Stück verwendet wurden, oder ob Tonartwechsel vorgenommen werden.

Würde bei einem Stück in beispielsweise C-Dur (sieben Töne) noch nach C-Moll gewechselt, so betrüge die maximale Knotenzahl $7 + 1$, da die Tonleitern sechs gemeinsame Töne haben. Wechsel zu Paralleltonarten werden hierbei nicht berücksichtigt, da diese dasselbe Tonmaterial besitzen. Zusätzlich können einzelne zusätzliche Töne, wie zum Beispiel «Blue Notes», welche einem Stück «Blues-Charakter» verleihen, sich in der Anzahl der Knoten bemerkbar machen.

3.3.2. Entropie

Als Mass der «Informationsdichte» des Stückes dient die Informations-Entropie (wir benutzen die Basis e), über die Abfolge der Knoten. Die unterschiedlichen Knoten bilden das Alphabet für die Bestimmung der Entropie.

Bei der Anwendung des Entropiebegriffes in dieser Form ist die resultierende Entropie nur bedingt abhängig von der Länge des Stückes, da man davon ausgehen kann, dass das Stück aus sich wiederholenden Teilen besteht. Um ein normiertes Mass für den Vergleich zu erhalten, drücken wir die Entropie als Verhältnis zur maximal möglichen Entropie aus:

$$\text{Entropie Verhältnis} = \frac{H}{H_{max}}$$

3.3.3. Eigenwerte

Aus den Adjazenzmatrizen, welche die Graphen beschreiben, lassen sich Eigenwerte berechnen. Da die Eigenwerte für eine Matrix charakteristisch sind, sollten sie auch für den Graphen eine beschreibende Eigenschaft besitzen. Für die Feature-Vektoren verwenden wir aus der Menge der nicht-komplexen Eigenwerte folgende Werte:

- Anzahl Eigenwerte
- Spektrum der Eigenwerte (wobei nur reelle Eigenwerte, ungleich 0, berücksichtigt werden):

$$\text{Spektrum} = \text{Max}(\text{Eigenwerte}) - \text{Min}(\text{Eigenwerte})$$

Auffallend ist, dass die Anzahl der Eigenwerte eine grosse Spannweite hat, die geringste Anzahl liegt bei 3 und die höchste bei 73. Was man auch sagen kann, ist, dass die Graphen mit mehr Knoten und somit einer Adjazenzmatrix von grösserer Dimension, auch mehr Eigenwerte haben. Der Bereich der Eigenwerte ist aber nicht abhängig von der Knotenzahl.

3.3.4. Small World Phänomen

Bei Small World Netzwerken sind unter anderem die zwei Phänomene geringer Durchmesser und hohe Transitivität festzustellen [9]. Bisher konnte man diese Phänomene vor allem in sozialen Netzwerken beobachten [10]. Man geht aber davon aus, dass in vielen natürlichen Netzwerken diese Small World Phänomene vorkommen.

Bei den Graphen der Musikstücke scheint es sich ebenfalls um Small World Netzwerke zu handeln. Alle im Folgenden beschriebenen Features weisen Werte auf, die für Small World Netzwerke charakteristisch sind.

Charakteristische Pfadlänge

Zwischen allen Knotenpaaren im vorliegenden Graph werden jeweils die kürzesten Pfade (nach Dijkstra) berechnet und davon der Mittelwert verwendet.

Bei Small World Netzwerken bewegt sich diese Länge im Bereich von 5 - 7. Der Durchschnitt bei den von uns berechneten Stücken liegt bei 3.71, das heisst der Wert liegt sogar leicht darunter, was durch die teilweise tiefe Knotenzahl erklärt werden kann.

Durchmesser

Als Durchmesser des Graphen verwenden wir die grösste Distanz zwischen zwei Knoten. Der Durchmesser der Musik-Graphen bewegt sich zwischen 4 und 20. Dies sind tiefe Werte, wenn man beachtet, dass sich unter den untersuchten Graphen solche mit beinahe 500 Knoten befinden.

Globaler Clustering Koeffizient

Der Clustering Koeffizient beschreibt die Ausprägung der Cliquenbildung. Sind zwei Knoten mit einem dritten verbunden, so ist es wahrscheinlich, dass diese auch untereinander verbunden sind.

$$\text{Clustering Koeffizient} = \frac{3 \cdot \text{Anzahl Dreiecke (geschlossene Pfade der Länge 3)}}{\text{Anzahl Tripel (Mögliche Pfade der Länge 2)}}$$

Der Faktor 3 kompensiert das 3-fache Zählen eines Knotens in den Tripeln.

Small-World'ness

Wie in Tabelle 1 zu sehen ist, scheint der Graph des Stückes «Autumn Leaves» durchaus Small-World Tendenzen aufzuweisen. Bei einer relativ hohen Anzahl von Knoten bleibt der Durchmesser und die Charakteristische Pfadlänge relativ klein.

Eigenschaft	Wert
Anzahl Knoten	286
Globaler Clustering Koeffizient	0.11
Durchmesser	14
Charakteristische Pfadlänge	4.53

Tabelle 1: Grapheigenschaften vom Stück «Autumn Leaves».

Es ist allerdings naheliegend, dass Netzwerke mit einer Knotenzahl von 20 und weniger eine charakteristische Pfadlänge von 6 und weniger aufweisen. Aus diesem Grund verwenden wir das Verhältnis der Knotenzahl zur charakteristischen Pfadlänge als normalisierte Kenngrösse. Dieses Verhältnis zeigt auf, ob es sich lediglich um einen kleinen Graphen handelt, oder ob die Pfadlänge wirklich ein interessantes Mass aufweist.

3.4. Bereich der Features im Vergleich

Um einen Überblick über die insgesamt fast 600 Featurewerte der 42 Stücke zu bekommen, haben wir die Bereiche der Features, aufgeteilt nach Genres, visualisiert. Man sieht, dass die Wertebereiche meist eine kleine Spannweite aufweisen und dass sich die Bereiche zwischen den Genres meistens überlappen.

Die Diagramme sind wie folgt zu lesen: Jeder Balken steht für den Wertebereich von den jeweils sieben Stücken eines Genres, die X-Achse ist der Wertebereich, die Y-Achse sind die Genres, abgekürzt nach folgender Tabelle:

Abkürzung	Genre
E	Elektronische Musik
F	Folklore/«Ländler»
J	Jazz
K	Klassik
P	Pop
R	Rock

Tabelle 2: Genre Abkürzungen.

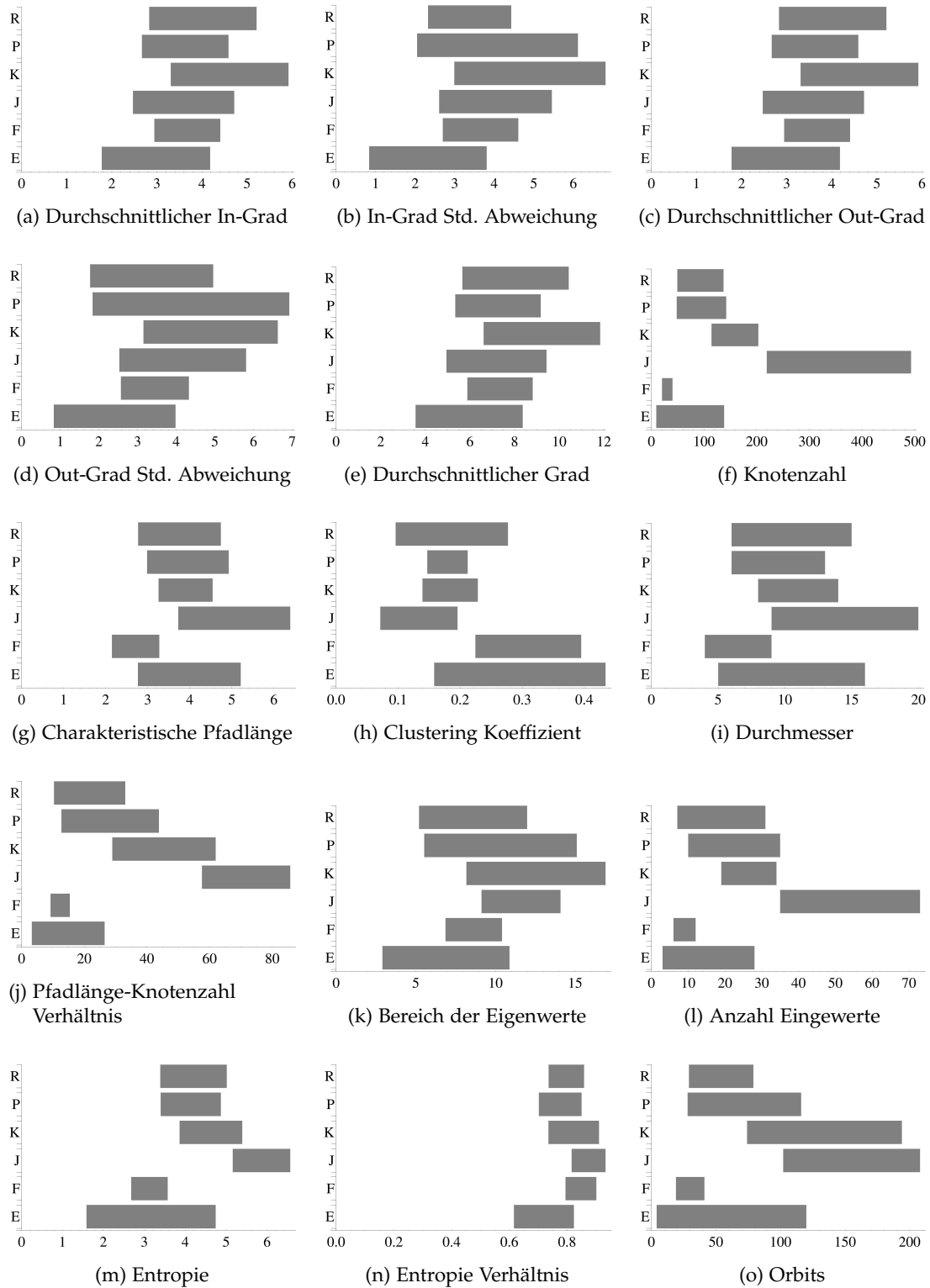


Abbildung 5: Vergleich der Feature-Bereiche per Genre. Man sieht, dass die Wertebereiche meist eine kleine Spannweite aufweisen und dass sich die Bereiche zwischen den Genres teilweise überlappen.

3.5. Featureeigenschaften

Die Features der Musik-Graphen weisen durchwegs interessante Werte auf. Es lässt sich sagen, dass die Charakteristiken Eigenschaften von komplexen Netzwerken und Small World Netzwerken aufweisen. Ausserdem liegen die Werte in einem ähnlichen Bereich und die Ergebnisse sind nicht von Stück zu Stück komplett verschieden. Das zeigt, dass bei den Musik-Graphen über alle Stücke aus allen Genres ähnliche Muster und Verhalten vorliegen, die nicht zufällig sind.

Da die Feature-Vektoren der Stücke scheinbar gute und interessante Werte aufweisen, stellt sich nun die Frage, ob diese Features nur den Graphen oder auch das dargestellte Musikstück gut charakterisieren.

Um diese Frage zu klären, haben wir uns für ein Clustering dieser Feature-Vektoren entschieden. Das Vorgehen und die Ergebnisse werden wir in den nächsten Abschnitten beschreiben.

4. Clustering der Graphen

Mittels Clustering möchten wir Gemeinsamkeiten zwischen den generierten Graphen anhand der untersuchten Charakteristiken finden. Als Clusteringverfahren verwenden wir den HLC Algorithmus (Hebbian Learning Clustering, Landis, Ott, Stoop 2010 [8]). Dieser Algorithmus bietet den Vorteil gegenüber bekannten Algorithmen wie K-Means, dass er eine beliebige Anzahl von Clustern finden kann, ohne dass diese Anzahl im voraus bekannt sein muss. Ausserdem ist es nicht notwendig, dass die Datenpunkte der jeweiligen Cluster durch eine Hyperebene separierbar sind. Es können so auch konvexe Formen gefunden werden.

Das Clustering kann unter verschiedenen Voraussetzungen durchgeführt werden. Zum einen lässt sich der Clustering Algorithmus parametrisieren und zum anderen ist die Wahl der Komponenten des Feature-Vektors entscheidend für das Ergebnis. Aus diesem Grund wird das Clustering mehrfach mit verschiedenen Eingangsvoraussetzungen durchgeführt. Anschliessend werden die Ergebnisse verglichen, mit dem Ziel, eine optimale Konfiguration zu finden.

4.1. Datenmaterial

Als Datenbasis für die Analyse dienen uns 42 Musikstücke, welche wir in die folgenden sechs Genres, zu jeweils sieben Stücken, aufgeteilt haben:

- Elektronische Musik
- Folklore
- Jazz
- Klassische Musik
- Pop
- Rock

Eine detaillierte Auflistung der Stücke ist im Anhang zu finden. Bei der Auswahl der Stücke wurde darauf geachtet, möglichst genretypische Vertreter zu wählen. Die Genres werden im weiteren Verlauf des Textes oft mit ihrem Anfangsbuchstaben abgekürzt.

4.2. Erwartungen an das Resultat des Clusterings

Die Unterteilung in Genres ist ein gängiger Weg, Musik zu klassifizieren, ist aber gleichzeitig schwer zu verifizieren, insbesondere im mathematischen Sinn. Die Klassifizierung nach Genres, wie sie der Menschen vornimmt, ist subjektiv, aber dennoch in vielen Fällen erstaunlicherweise eindeutig. Die Stücke, die einem Genre zugeteilt sind, müssen demnach Gemeinsamkeiten aufweisen.

Die untersuchten Charakteristiken der Graphen beschreiben im Wesentlichen die strukturelle Komplexität. Die Frage ist nun, ob in dieser strukturellen Komplexität Gemeinsamkeiten zu finden sind, welche eine Klassifizierung nach Genres zulassen.

Die Erwartung an das Clustering-Resultat ist also, eine Gruppierung der Stücke in sechs Gruppen zu je sieben Stücken, wie sie durch die Genres vorgegeben ist, zu finden. Wie in Abbildung 5 bereits zu sehen war, gibt es überlappende Bereiche der Feature-Werte zwischen den Genres. Diese Überlappungen sind auch in den Clusteringergebnissen zu erwarten.

4.3. Bewertung der Resultate

Um verschiedene Clustering-Ergebnisse rasch bewerten zu können, haben wir einen einfachen Bewertungsalgorithmus entwickelt.

Der Algorithmus funktioniert folgendermassen:

1. Die gleichen Genres jedes Clusters werden zusammengezählt, das Genre mit dem höchsten Aufkommen ist der «Besitzer» des Clusters.
2. Die Anzahl der Stücke mit gleichem Genre wird quadriert, um eine nichtlineare Bewertung der Anzahl zu erhalten.
3. Die quadrierte Anzahl der Stücke eines Genres, die nicht Besitzer des Clusters sind, werden von der quadrierten Anzahl des Besitzers abgezogen. Das Ergebnis ist die Bewertung des Clusters.
4. Alle Bewertungen der einzelnen Clusters werden addiert und ergeben so die Bewertung des gesamten Clusterings.

Daraus ergibt sich eine maximale Punktzahl von 294 Punkten bei einem perfekten Clustering, sowie eine minimale Punktzahl von -196 bei einem einzigen Cluster.

Verhalten des Bewertungsverfahrens

Das Bewertungsverfahren wurde über 100000 zufällig generierte Sets von Stücken, die der Struktur von Clusteringergebnissen entsprechen, durchgeführt. Erreicht wurde eine durchschnittliche Punktzahl von zirka -10. Die 50 besten Werte bewegen sich im Bereich von 45 bis 75, die 50 tiefsten Werte sind alle -196. Wir gehen davon aus, dass dieses Bewertungsverfahren gute Ergebnisse liefert, da bei zufälligen Sets die durchschnittlich erreichte Punktzahl klein ist und die maximale Punktzahl nicht erreicht wurde.

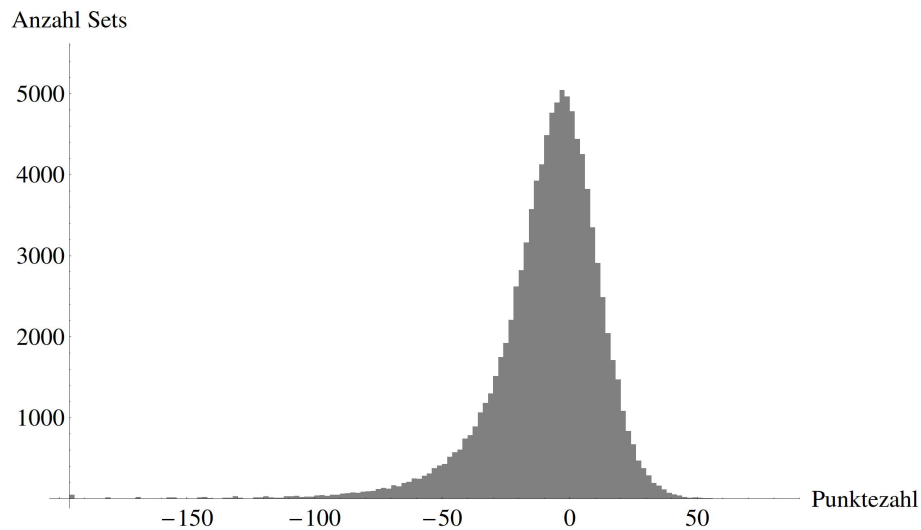


Abbildung 6: Histogramm der Punkteverteilung bei zufällig generierten Sets von Musikstücken. Die Maximalpunktzahl wird nie erreicht. Im Durchschnitt erhalten Zufallsergebnisse -10 Punkte.

4.4. HLC Clustering Parameter

Bei unseren Clustering-Versuchen konnte mit den Parametern wie sie in Tabelle 3 zu sehen sind, das beste Resultat erzielt werden. Bei anderen Konfigurationen wurden entweder alle Proben einem einzigen Cluster zugeordnet, oder es wurde gar kein Cluster gefunden. Wenn die Konfiguration nur geringfügig geändert wurde, konnten keine Unterschiede im Clustering festgestellt werden.

Daraus schliessen wir, dass diese Konfiguration ein stabiles Ergebnis liefert und somit für unsere Daten gut gewählt ist.

Parameter	Wert
neighbourhoodNeurons	1
d0Constant	0.5
tauConstant	0.5
voltageIR	25
theta	16

Tabelle 3: HLC Clustering Parameter

4.5. Ergebnisse

Das Clustering wurde über drei verschiedene Feature-Arten gemacht, nämlich über die Standard-, Skylined- und Timing-Features.

Bei allen Arten wurde ein Feature-Vektor mit elf Dimensionen verwendet. Dabei waren folgende Features enthalten:

- Durchschnittlicher Grad
- Standardabweichung der Grade
- Spektrum der Eigenwerte
- Anzahl der Eigenwerte
- Entropie
- Anzahl der Orbits
- Clustering Koeffizient
- Charakteristische Pfadlänge
- Durchmesser
- Verhältnis zwischen charakteristischer Pfadlänge und Knotenzahl
- Verhältnis der Entropie zur maximalen Entropie

Bei keinem der Feature-Sets konnte ein perfektes Ergebnis erzielt werden. Es wurden tendenziell zu viele Cluster gefunden, also mehr als die erwarteten sechs. Es gibt auch diverse Cluster, die man nicht eindeutig einem Genre zuordnen kann. Wenn man also eine unbenannte Probe in so einen Cluster legen würde, könnte man nicht eindeutig bestimmen zu welchem Genre diese Probe wohl gehören würde.

Wenn man den Bewertungsalgorithmus auf die Ergebnisse anwendet, sieht man, dass sich die Punktzahlen deutlich vom Durchschnitt der Zufallssets abheben. Die erreichte Punktzahl ist trotzdem weit vom erwünschten Maximum entfernt.

Feature Art	Punktzahl	Anzahl Cluster
Standard	68	10
Skylined	38	10
Timing	42	12

Tabelle 4: Clustering Ergebnisse, Anzahl Cluster und Punktzahl

Die Clustering Resultate sind in Form von Tabellen auf den folgenden Seiten visualisiert, Tabelle 6 zeigt das Clustering der Standard-Features, Tabelle 7 zeigt das Clustering der Skylined-Features und Tabelle 8 zeigt das Clustering der Timing-Features.

Die Tabellen zeigen pro Zeile einen Cluster. Die Spalten repräsentieren die sieben Stücke des jeweiligen Genres.

Abkürzung	Genre
E	Elektronische Musik
F	Folklore/«Ländler»
J	Jazz
K	Klassik
P	Pop
R	Rock

Tabelle 5: Genre-Abkürzungen.

Welche Stücke verwendet wurden und wie das detaillierte Clustering mit Stücknamen aussieht, ist im Anhang zu finden.

4.6. Interpretation

Es scheint, dass das Clustering-Ergebnis durchgehend nicht zufriedenstellend ist. Eine genauere Betrachtung der Cluster, legt aber den Schluss nahe, dass das Clustering zwar gut ist, aber dass nicht Genres erkannt wurden, sondern andere Gemeinsamkeiten. Abbildung 7, Abbildung 8 und Abbildung 9 zeigen, dass die Clustering-Ergebnisse deutlich besser sind, als eine zufällige Gruppierung der Stücke.

Für die Diagramme wurden die Features der einzelnen Clusters verglichen und die Standardabweichung errechnet. Auf der Y-Achse sieht man den Wert der durchschnittlichen Standardabweichung aller Features, die X-Achse steht für die Einzelnen Cluster. Die Gerade zeigt die durchschnittliche Standardabweichung der Features von 10000 zufälligen Sets mit den Eigenschaften von Clusters.

Durchschnittl. Std. Abweichung

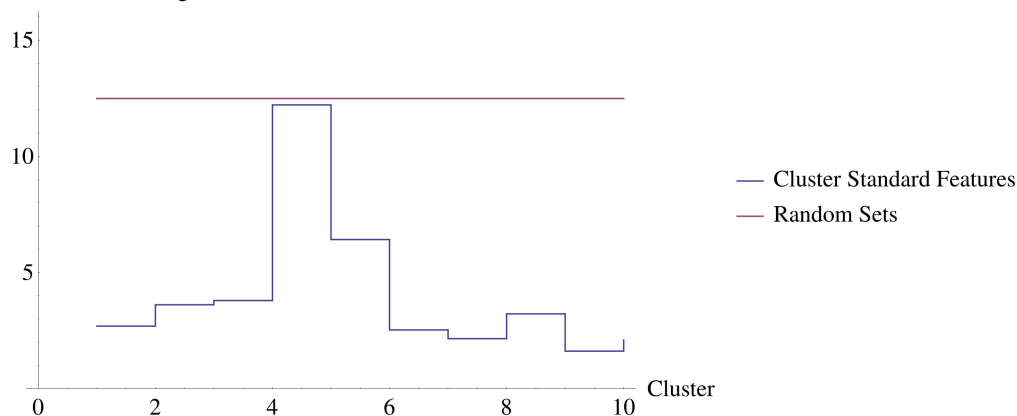


Abbildung 7: Vergleich der durchschnittlichen Standardabweichung der Standard-Features innerhalb der Clusters, mit der durchschnittlichen Standardabweichung von 10000 zufälligen Sets von Standard-Features.

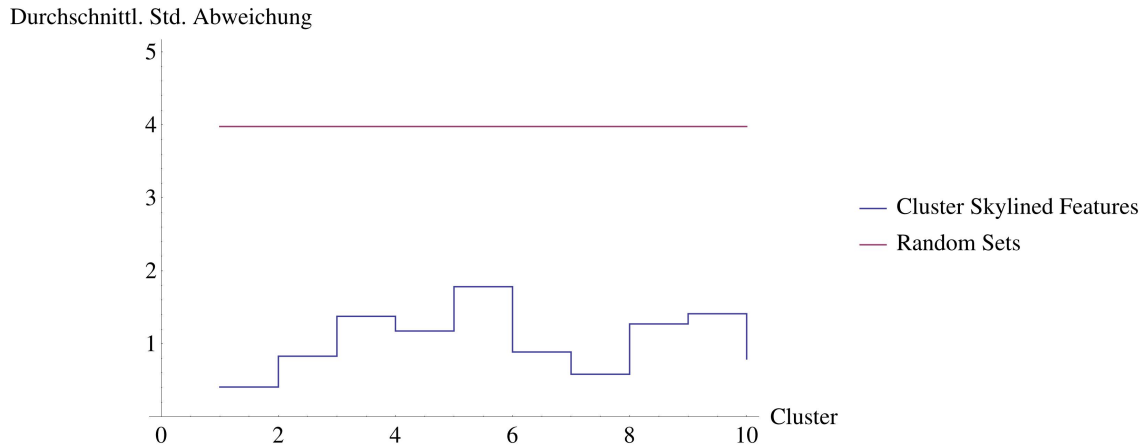


Abbildung 8: Vergleich der durchschnittlichen Standardabweichung der Skylined-Features innerhalb der Clusters, mit der durchschnittlichen Standardabweichung von 10000 zufälligen Sets von Skylined-Features.

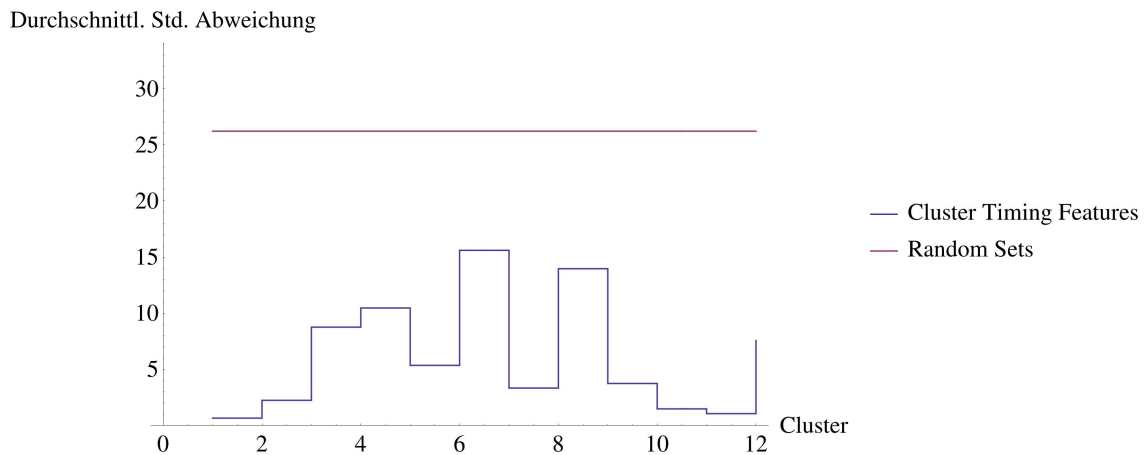


Abbildung 9: Vergleich der durchschnittliche Standardabweichung der Timing-Features innerhalb der Clusters, mit der durchschnittlichen Standardabweichung von 10000 zufälligen Sets von Timing-Features.

Diese Tatsache lässt drauf schliessen, dass nicht das Clustering das Problem darstellt, sondern, dass die verwendeten Features nicht genretypisch sind. Die Folge ist, dass entweder andere Features gesucht, oder genauer analysiert werden muss, was die Features über die Stücke aussagen.

Scheinbar sind Stücke innerhalb eines Genres verschiedener als erwartet. Jedoch scheint es trotzdem Gemeinsamkeiten unter den Stücken zu geben.

[illegible]

Tabelle 6: Clustering der Standard-Features.

[illegible]

Tabelle 7: Clustering der Skylined-Features.

	E	E	E	E	E	E	E	F	F	F	F	F	F	F	F	J	J	J	J	J	J	J	J	K	K	K	K	K	K	K	P	P	P	P	P	P	P	R	R	R	R	R	R	R					
Cluster 1	x	x																																															
Cluster 2			x																					x																									
Cluster 3				x						x	x	x																																					
Cluster 4																																																	
Cluster 5																																																	
Cluster 6																																																	
Cluster 7																																																	
Cluster 8																																																	
Cluster 9																																																	
Cluster 10																																																	
Cluster 11																																																	
Cluster 12																																																	

Tabelle 8: Clustering der Standard-Features.

5. Weiterführende Ideen

In diesem Kapitel möchten wir Ideen festhalten, für deren Umsetzung die Zeit fehlte und welche in Folgearbeiten genauer betrachtet werden können.

Extraktion der Hauptmelodie

Ein bereits beschriebenes Problem stellen die diversen Spuren der MIDI Dateien dar. Mit dem Ziel, die wichtigsten Töne besser zu extrahieren, schlagen wir folgende Ansätze vor:

- die Spur mit den hochfrequentesten Änderungen genauer zu betrachten
- die Spur mit der grössten Entropie genauer zu betrachten
- Skylining Ansatz mit anderen Ausprägungen, wie: tiefste Note, längste Note

Einbettung

Es wäre interessant zu sehen, ob anhand einer Einbettung ohne Überlappung, Wiederholungen besser erkennbar sind, als anhand einer Einbettung mit Überlappung.

Knotenbesuche

Bis jetzt ist aus dem Graphen nicht ersichtlich, wie oft ein Knoten durchlaufen wird. Wäre diese Information im Graphen enthalten, liessen sich zusätzliche Zentralitätseffekte identifizieren. Eine Möglichkeit wäre, im Graphen Übergänge derart zu gewichten, dass pro Transition das Kantengewicht hochgezählt wird. Eine andere Möglichkeit wäre, die Zahl der Knotenbesuche im Knoten selber abzulegen.

Features

Folgende Features könnten noch betrachtet werden

Zentralität Durchschnittliche Erdős Zahl zum zentralsten Knoten (Knoten mit höchstem Knoten-grad).

Chromatische Nummer Siehe: <http://mathworld.wolfram.com/EdgeChromaticNumber.html>.

Zufallspfade und künstliche Musik

Man könnte Random-Walks mit Hilfe von Markov-Ketten durchführen, in Verbindung mit Kantenengewichten als mögliche Beschreibung der Übergangswahrscheinlichkeiten.

Literatur

- [1] Joan Serrà et al., *Measuring the Evolution of Contemporary Western Popular Music* Scientific Reports 2, Artikel 521, 2012
- [2] Fred Lerdahl, Ray Jackendoff *A generative theory of tonal music* MIT Press, Cambridge, USA, 1983
- [3] MIDI Manufacturers Association, *The Complete MIDI 1.0 Detailed Specification: Incorporating All Recommended Practices ; Document Version 96.1* MIDI Manufacturers Association Incorporated, 2006
- [4] John Scott, *Trend Report Social Network Analysis*. Sociology Vol. 22, Seiten 109-127, 1967
- [5] Giyasettin Ozcan, Cihan Isikhan, Adil Alpkocak, *Melody extraction on MIDI music files* Seventh IEEE International Symposium on Multimedia, Seite 8, 2005
- [6] Steven H. Strogaz, *Exploring complex networks* Nature Vol. 410, Seiten 268-276, 2001
- [7] Stefano Boccaletti, Vito Latora, Yamir Moreno, Martin Chavez, D-U Hwang, *Complex networks: Structure and dynamics* Physics reports Vol. 424, Seiten 175-308, 2006
- [8] Florian Landis, Thomas Ott, Ruedi Stoop, *Hebbian Self-Organizing Integrate-and-Fire Networks for Data Clustering*. Neural Computation 22, Seiten 273-288, 2010
- [9] Duncan J. Watts, *Small Worlds: the dynamics of networks between order and randomness*. Princeton University Press 1971
- [10] Stanley Milgram, *The Small-World Problem*. Psychology Today Vol. 1, Seiten 61-67, 1967
- [11] Peter J. Carrington, John Scott, Stanley Wasserman, *Models and Methods in Social Network Analysis* Cambridge university press, 2005
- [12] Abbe Mowshowitz, Valia Mitsou *Entropy, Orbits, and Spectra of Graphs* Wiley-VCH Weinheim, Analysis of Complex Networks, 2009
- [13] Aniruddh D. Patel *Music, language, and the brain* Oxford University Press, USA, 2010.

A. Aufgabenstellung

Harmonie-Netzwerke in der Pop/Rock-Musik: Vorschlag für eine Bachelor-Arbeit

R. Stoop, Uni Zürich / ETH Zürich

23. September 2013

Zusammenfassung

Es geht darum, moderne Lieder harmonisch als komplexe Netzwerke darzustellen. Diese Netzwerke sollen in der Sprache der Physik charakterisiert und verglichen werden.

1 Einführung

Musik ist in unserem Leben allgegenwärtig und wird von vielen Menschen als zentraler Punkt in ihrem Leben empfunden. Gerade Europa hat durch seine lange und reiche Musiktradition in diesem Sinne einen besonderen Status. Obwohl z.B. die Musik von Johann Sebastian Bach oft als genial und überaus „komplex“ empfunden werden, sind nur wenige Methoden angewandt worden, um diese Komplexität zu untersuchen. Die in letzter Zeit wissenschaftlich sehr populäre Methode der Analyse von komplexen Netzwerken erlaubt hier neue Einblicke. Es soll nun darum gehen, die Harmonien von modernerer Musik (Pop/Rock) als komplexe Netzwerke zu erfassen und zu untersuchen.

2 Methoden

Grundlage der Arbeit bilden Files im MIDI-Format. Dieses Format enthält detaillierte Angaben über die einzelnen Noten, nämlich über das Instrument, die Länge der Note (Anfangs- und Endpunkt) etc. Diese MIDI-Files sind im Internet oft einfach und gratis erhältlich. Es geht nun darum, diese MIDI-Files einzulesen und zu schauen, welche Noten zusammen erklingen (d.h. zur selben Zeit). So ein Akkord (z.B. C-E-G) ist nun ein Knoten des Netzwerks. Geht die Musik weiter, so wird ein neuer Knoten erreicht und ein „Link“ zwischen diesen beiden Knoten erzeugt (z.B. wäre der nächste Knoten C-E-A, wenn die Noten C und E bleiben, aber aus einem G ein A wird). Die Knoten sollen nur im Sinne der Harmonie charakterisiert werden: So ist z.B. C-E-G, E-G-C oder C-E-G-C alles dasselbe, d.h. mehrfach vorhandene Töne und die absolute Tonhöhe werden ignoriert. Insgesamt gibt es so:

- 1 möglicher Knoten für „Pause“
- 12 mögliche einstimmige Knoten (C,Cis,D,Dis,...,H)
- 66 mögliche zweistimmige Knoten (C-Cis,C-D,C-Dis,...,Cis-D,Cis-Dis,...,D-Dis,D-E.....)
- 220 mögliche dreistimmige Knoten

- 495 mögliche vierstimmige Knoten
- 792 mögliche fünfstimmige Knoten
- etc.

Sind von einem Lied nicht die kompletten Noten, sondern nur Melodie und Akkorde angegeben, so müssen die Akkordtöne miteinbezogen werden (Ein Melodieton „A“ + G-Dur entspricht dann dem Knoten D-G-A-H). Weiter: Cis und Des, Dis und Es etc. sollen als gleiche Note betrachtet werden (wie wir uns das heute mit der wohltemperierten Stimmung gewohnt sind). Ein typisches Lied besucht natürlich nur einen kleinen Teil der möglichen Knoten. Indem analysiert wird, welcher Knoten auf welchen Knoten folgt, entsteht nun ein kleines/mittelgroßes komplexes Netzwerk, welches untersucht werden soll. Einfachheitshalber kann man das Netzwerk als „undirected“ betrachten. Anschauen kann man sich nun:

- Degree Distribution
- Clustering Coefficient der einzelnen Nodes und Average Clustering Coefficient
- Centrality der einzelnen Nodes und Average Centrality

Weitere Möglichkeiten:

- Die Netzwerk-Links gewichten, indem geschaut wird, wie oft ein Übergang im Lied vorhanden ist.
- (Controlled) Random Walks auf dem (gewichteten) Netzwerk erzeugen und so *künstliche Musik* kreieren.
- Ein vollautomatisiertes Programm, das das MIDI-File einliest und eine Animation des Netzwerks inklusive Musik ausspuckt (siehe Beispiel-Video).

3 Beispiel

Figur 1 zeigt das harmonische Netzwerk des C-Dur-Präludiums aus dem 1. Band von J.S. Bach's Wohltemperiertem Klavier. Siehe auch das Video mit Musik, wie dieses Netzwerk durch das Stück gebildet wird oder auch das MIDI-File als Beispiel für künstliche Musik (random walk auf dem gewichteten Netzwerk).

4 Hilfreiches

- Ref. [1] macht etwas ähnliches, aber „beat-based“ (siehe Fig. 1 im Paper). Dies macht vom harmonischen Standpunkt nur bedingt Sinn: Als Beispiel der zweite Knoten in Figur 1 (G-Dur, Melodie macht A-H). Bei ihnen ist dies dann ein Knoten D-G-A-H, im Sinne der Arbeit hier wären das zwei aufeinanderfolgende Knoten D-G-A-H und D-G-H. Hilfreich könnten die Referenzen im Paper sein (z.B. „Million Song Dataset“).
- Ref. [2] gibt einen Überblick über Komplexe Netzwerke.

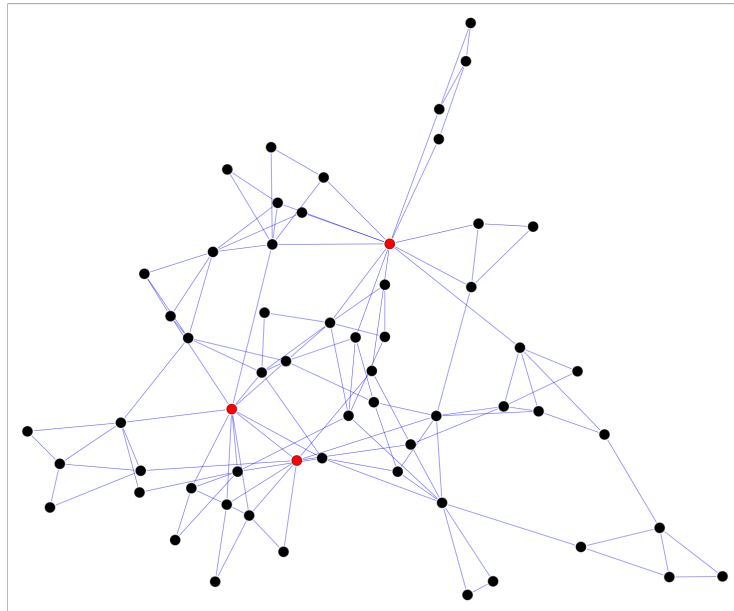


Abbildung 1: Harmonisches Netzwerk des C-Dur Präludiums (Bach, WTK I). Die drei Knoten mit dem höchsten Degree (C, G, C-E-G) sind rot angefärbt.

- Für einen Vergleich mit klassischer Musik kann F. Gomez kontaktiert werden (fgomez@ini.phys.ethz.ch).
- Hilfreiche Mathematica-Befehle: „Import“ für das Importieren von MIDI-Files, „AdjacencyGraph“ für das Plotten von Netzwerken

Literatur

- [1] J. Serra et al., Measuring the Evolution of Contemporary Western Popular Music, *Scientific Reports* 2:52, 2012.
- [2] S. Boccaletti et al., Complex networks: Structure and dynamics, *Physics Reports* 424:4, 2006.

Ruedi Stoop
 Ann. Kr.

B. Musikstücke

```
-- electronic
|   |-- E-Chemical-Brothers_Loops-Of-Fury.mid
|   |-- E-Daft-Punk_One-More-Time.mid
|   |-- E-DJ-Taucher_Infinity.mid
|   |-- E-Kernkraft400-Zombie_Nation.mid
|   |-- E-Members_of_Mayday-Save_the_robots.mid
|   |-- E-Rollergirl_dear-jessie.mid
|   '-- E-The-Prodigy_Breathe.mid
-- jazz
|   |-- J-A_string_of_Pearls_jc.mid
|   |-- J-Autumn_leaves.mid
|   |-- J-C_Brown-Joy_Spring.mid
|   |-- J-Coltrain-Blue_Train.mid
|   |-- J-Ill_Walk_Alone.mid
|   |-- J-Miles_Davis_Four.mid
|   '-- J-Stompin_At_The_Savoy.mid
-- klassik
|   |-- K-Bach-Praelidium_C_Dur.mid
|   |-- K-Beethoven-Fuer_Elise.mid
|   |-- K-Mozart-Kleine_Nachtmusik.mid
|   |-- K-Saint-Saens_Lions.mid
|   |-- K-Strauss-An-Der-Blauen-Donau.mid
|   |-- K-Vivaldi-Four_Seasons_Spring.mid
|   '-- K-Wagner-Ride-Of-The-Valkries.mid
-- folklore
|   |-- L-44Huehner.mid
|   |-- L-AuseerSchottisch.mid
|   |-- L-Gretlpolka.mid
|   |-- L-HanslHupf.mid
|   |-- L-LeckerfasslPolka.mid
|   |-- L-Muntipolka.mid
|   '-- L-SchottischOllersdorf.mid
-- pop
|   |-- P-Alphaville-Big_In_Japan.mid
|   |-- P-Cher-Believe.mid
|   |-- P-Europe-Final_Countdown.mid
|   |-- P-Katy_Perry-ET.mid
|   |-- P-Michael_Jackson-Billie-Jean.mid
|   |-- P-Pet_Shop_Boys-Its_A_Sin.mid
|   '-- P-Prince-Kiss.mid
```

```
'-- rock
  |-- R-ACDC-Highway_To_Hell.mid
  |-- R-Bon_Jovi-Bounce.mid
  |-- R-Deep_Purple-Smoke_On_The_Water.mid
  |-- R-Green_Day-Holiday.mid
  |-- R-Lynyrd_Skynyrd-Sweet_Home_Alabama.mid
  |-- R-Rolling_Stones-Satisfaction.mid
  '-- R-Scorpions-Rock_You_Like_A_Hurricane.mid
```

C. Clusteringergebnisse (Details)

C.1. Standard Features

```
{E-Chemical-Brothers_Loops-Of-Fury-raw-features,  
E-The-Prodigy_Breathe-raw-features,  
E-Kernkraft400-Zombie_Nation-raw-features,  
L-SchottischOllersdorf-raw-features,  
L-HanslHupf-raw-features,  
E-Daft-Punk_One-More-Time-raw-features,  
L-Gretlpolka-raw-features,  
L-LeckerfasslPolka-raw-features,  
L-AuseerSchottisch-raw-features,  
P-Alphaville-Big_In_Japan-raw-features,  
L-44Huehner-raw-features,  
L-Muntipolka-raw-features}
```

```
{E-DJ Taucher - Infinity-raw-features,  
P-Michael_Jackson-Billie-Jean-raw-features,  
R-ACDC-Highway_To_Hell-raw-features,  
K-Beethoven-Fuer_Elise-raw-features,  
E-Members_of_Mayday-Save_the_robots-raw-features}
```

```
{E-Rollergirl_dear-jessie-raw-features,  
K-Saint-Saens_Lions-raw-features,  
K-Bach-Praelidium_C_Dur-raw-features,  
R-Lynyrd_Skynyrd-Sweet_Home_Alabama-raw-features}
```

```
{J-A_string_of_Pearls_jc-raw-features,  
J-Autumn_leaves-raw-features,  
J-C_Brown-Joy_Spring-raw-features,  
J-Ill_Walk_Alone-raw-features}
```

```
{J-Coltrain-Blue_Train-raw-features,  
K-Vivaldi-Four_Seasons_Spring-raw-features,  
K-Wagner-Ride-Of-The-Valkries-raw-features,  
K-Mozart-Kleine_Nachtmusik-raw-features,  
P-Prince-Kiss-raw-features}
```

```
{J-Miles_Davis_Four-raw-features,  
J-Stompin_At_The_Savoy-raw-features}
```

```
{K-Strauss-An-Der-Blauen-Donau-raw-features,  
P-Cher-Believe-raw-features,  
R-Rolling_Stones-Satisfaction-raw-features}  
  
{P-Europe-Final_Countdown-raw-features,  
P-Pet_Shop_Boys-Its_A_Sin-raw-features,  
R-Bon_Jovi-Bounce-raw-features}  
  
{P-Katy_Perry-ET-raw-features,  
R-Green_Day-Holiday-raw-features}  
  
{R-Deep_Purple-Smoke_On_The_Water-raw-features,  
R-Scorpions-Rock_You_Like_A_Hurricane-raw-features}
```

C.2. Skylined Features

```
{E-Chemical-Brothers_Loops-Of-Fury-skylined-raw-features,  
E-Kernkraft400-Zombie_Nation-skylined-raw-features,  
E-The-Prodigy_Breathe-skylined-raw-features}  
  
{E-Daft-Punk_One-More-Time-skylined-raw-features,  
P-Pet_Shop_Boys-Its_A_Sin-skylined-raw-features,  
R-Scorpions-Rock_You_Like_A_Hurricane-skylined-raw-features,  
E-Rollergirl_dear-jessie-skylined-raw-features,  
P-Cher-Believe-skylined-raw-features,  
E-Members_of_Mayday-Save_the_robots-skylined-raw-features,  
R-Deep_Purple-Smoke_On_The_Water-skylined-raw-features}  
  
{E-DJ Taucher - Infinity-skylined-raw-features,  
K-Bach-Praelidium_C_Dur-skylined-raw-features,  
K-Beethoven-Fuer_Elise-skylined-raw-features,  
P-Prince-Kiss-skylined-raw-features,  
P-Michael_Jackson-Billie-Jean-skylined-raw-features,  
R-ACDC-Highway_To_Hell-skylined-raw-features,  
J-Coltrain-Blue_Train-skylined-raw-features,  
R-Green_Day-Holiday-skylined-raw-features}  
  
{J-A_string_of_Pearls_jc-skylined-raw-features,  
J-C_Brown-Joy_Spring-skylined-raw-features,  
J-Stompin_At_The_Savoy-skylined-raw-features}  
  
{J-Autumn_leaves-skylined-raw-features,
```

```
J-Ill_Walk_Alone-skylined-raw-features,  
J-Miles_Davis_Four-skylined-raw-features,  
K-Saint-Saens_Lions-skylined-raw-features,  
K-Wagner-Ride-Of-The-Valkries-skylined-raw-features}
```

```
{K-Mozart-Kleine_Nachtmusik-skylined-raw-features,  
K-Strauss-An-Der-Blauen-Donau-skylined-raw-features,  
K-Vivaldi-Four_Seasons_Spring-skylined-raw-features}
```

```
{L-44Huehner-skylined-raw-features,  
L-LeckerfasslPolka-skylined-raw-features,  
L-Gretlpolka-skylined-raw-features,  
P-Alphaville-Big_In_Japan-skylined-raw-features}
```

```
{L-AuseerSchottisch-skylined-raw-features,  
R-Rolling_Stones-Satisfaction-skylined-raw-features}
```

```
{L-HanslHupf-skylined-raw-features,  
L-SchottischOllersdorf-skylined-raw-features}
```

```
{L-Muntipolka-skylined-raw-features,  
P-Europe-Final_Countdown-skylined-raw-features,  
P-Katy_Perry-ET-skylined-raw-features,  
R-Bon_Jovi-Bounce-skylined-raw-features}
```

C.3. Timing Features

```
{E-Chemical-Brothers_Loops-Of-Fury-raw-features,  
E-The-Prodigy_Breathe-raw-features}
```

```
{E-Daft-Punk_One-More-Time-raw-features,  
K-Strauss-An-Der-Blauen-Donau-raw-features,  
R-Green_Day-Holiday-raw-features}
```

```
{E-DJ Taucher - Infinity-raw-features,  
R-Scorpions-Rock_You_Like_A_Hurricane-raw-features,  
R-ACDC-Highway_To_Hell-raw-features,  
L-Gretlpolka-raw-features,  
L-AuseerSchottisch-raw-features,  
L-HanslHupf-raw-features,  
L-SchottischOllersdorf-raw-features}
```

```
{E-Rollergirl_dear-jessie-raw-features,  
R-Rolling_Stones-Satisfaction-raw-features,  
P-Cher-Believe-raw-features,  
P-Europe-Final_Countdown-raw-features}  
  
{J-A_string_of_Pearls_jc-raw-features,  
K-Saint-Saens_Lions-raw-features,  
R-Lynyrd_Skynyrd-Sweet_Home_Alabama-raw-features,  
P-Michael_Jackson-Billie-Jean-raw-features}  
  
{J-Autumn_leaves-raw-features,  
J-Ill_Walk_Alone-raw-features,  
J-Miles_Davis_Four-raw-features,  
J-Stompin_At_The_Savoy-raw-features}  
  
{J-C_Brown-Joy_Spring-raw-features,  
K-Bach-Praelidium_C_Dur-raw-features}  
  
{J-Coltrain-Blue_Train-raw-features,  
K-Wagner-Ride-Of-The-Valkries-raw-features,  
K-Vivaldi-Four_Seasons_Spring-raw-features,  
P-Prince-Kiss-raw-features,  
R-Deep_Purple-Smoke_On_The_Water-raw-features}  
  
{K-Beethoven-Fuer_Elise-raw-features,  
K-Mozart-Kleine_Nachtmusik-raw-features}  
  
{L-44Huehner-raw-features,  
L-LeckerfasslPolka-raw-features}  
  
{L-Muntipolka-raw-features,  
P-Alphaville-Big_In_Japan-raw-features}  
  
{P-Katy_Perry-ET-raw-features,  
R-Bon_Jovi-Bounce-raw-features,  
P-Pet_Shop_Boys-Its_A_Sin-raw-features}
```