

# Versionsübergreifende Visualisierung harmonischer Verläufe: Eine Fallstudie zu Wagners Ring-Zyklus

Christof Weiß<sup>1</sup>, Frank Zalkow<sup>1</sup>, Meinard Müller<sup>1</sup>, Stephanie Klauk<sup>2</sup>, Rainer Kleinertz<sup>2</sup>

**Abstract:** Die computergestützte Erschließung und Analyse von Musikdaten ist eine zentrale Aufgabenstellung der Musikverarbeitung. Ein wichtiges Teilproblem ist die Erkennung von tonalen Strukturen wie Akkorden oder Skalen. In diesem Beitrag diskutieren wir Methoden für die Visualisierung harmonischer Verläufe von Musikwerken und wenden diese exemplarisch auf die Tetralogie *Der Ring des Nibelungen* von Richard Wagner an. Insbesondere stellen wir ein Verfahren vor, bei dem unterschiedliche Aufnahmen (Versionen) desselben Musikwerkes simultan analysiert werden. Als einen wesentlichen Beitrag führen wir ein geeignetes Farbschema ein, mit dem die Gemeinsamkeiten und Unterschiede der Analyseergebnisse über verschiedene Versionen hinweg geeignet dargestellt und visuell erfahrbar gemacht werden. Am Beispiel der automatisierten Analyse harmonischer Strukturen deuten wir an, welches Potenzial informatische Methoden für die musikwissenschaftliche Forschung bieten und wie Fragestellungen der Musikwissenschaft einen Ausgangspunkt für neue informatische Methoden darstellen können.

**Keywords:** Music Information Retrieval, Musikverarbeitung, Visualisierung, Harmonieanalyse

## 1 Einleitung

Im Forschungsbereich des *Music Information Retrieval* (MIR) geht es um die Entwicklung von computergestützten Methoden und Werkzeugen zur Erschließung und Analyse von Musikdaten, welche in unterschiedlichen Formaten und Ausprägungen gegeben sein können. Hierbei nimmt die harmonische Analyse von Musikaufnahmen in Form von Audiodaten eine zentrale Rolle ein [CB14, MD10, Mc14]. Dies ist nicht verwunderlich, stellt doch die Harmonik ein grundlegendes Prinzip westlicher Musik dar. Aufgrund ihrer fundamentalen Bedeutung werden harmoniebasierte Merkmale häufig als Mid-Level-Darstellungen für weiterführende Retrieval- und Strukturierungsaufgaben herangezogen [Mü15]. Abhängig von der Zeitskala, also dem betrachteten zeitlichen Kontext, korrespondiert die harmonische Analyse mit unterschiedlichen Aufgabenstellungen wie der Extraktion von Akkordfolgen [CB14, MD10], der lokalen Schätzung von Skalen [PP11, WH14] bis hin zur globalen Bestimmung der Grundtonart [NS06]. Die Analyse eines Musikstücks hinsichtlich auftretender tonaler Strukturen kann durchaus problematisch sein. Für komplexe Musik sind sich Musikwissenschaftler oft nicht einig in der Deutung harmonischer Verläufe.

---

<sup>1</sup> International Audio Laboratories Erlangen, [christof.weiss@audiolabs-erlangen.de](mailto:christof.weiss@audiolabs-erlangen.de)

<sup>2</sup> Universität des Saarlandes, Institut für Musikwissenschaft

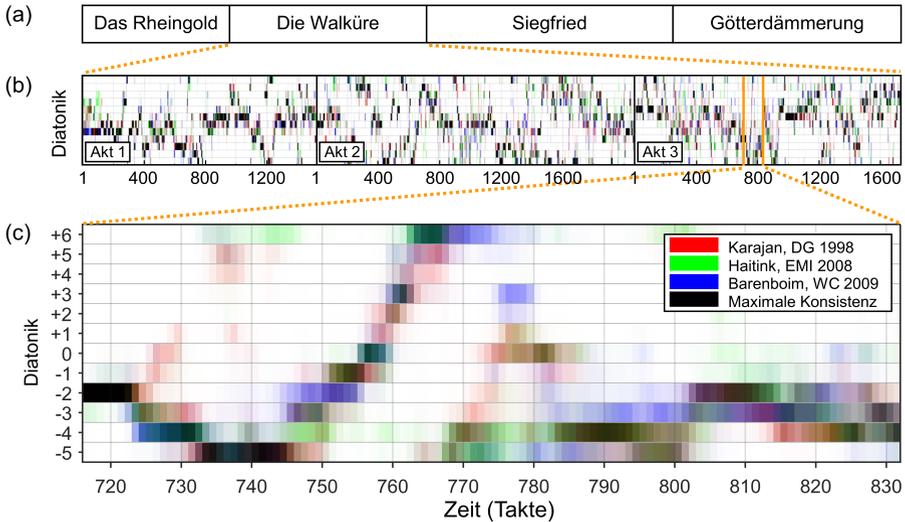


Abb. 1: Darstellung harmonischer Verläufe auf unterschiedlichen Zeitskalen. (a) Übersicht über den *Ring des Nibelungen*. (b) Gesamtüberblick über *Die Walküre*. (c) Detailansicht der Takte 716–832 aus dem dritten Akt der *Walküre*.

Unterschiedliche musiktheoretische Konzepte (Tonarten-, Funktionsbezeichnungen oder reduktionistische Notationen im Stile Heinrich Schenkers) liefern mehr oder weniger komplexe Beschreibungen, welche eine aufwändige Analyse erfordern und oft auch dem Spezialisten keinen intuitiven Überblick ermöglichen. Weiterhin hängt die Analyse von der jeweiligen Fragestellung, der musikalischen Stilistik sowie der betrachteten Zeitskala ab. Automatisierte Methoden zur Harmonieanalyse stoßen hier schnell an ihre Grenzen.

Anstatt eine explizite harmonische Analyse durchzuführen, geht es in diesem Beitrag um Visualisierungskonzepte, die automatisch gemessene harmonische Strukturen grafisch darstellen. In [Sa05] wird eine hierarchische Darstellungsform eingeführt, die dem Benutzer Analyseergebnisse simultan auf unterschiedlichen Zeitskalen präsentiert. Im Projekt „PHENICX“ (Performances as Highly Enriched aNd Interactive Concert eXperiences) werden Visualisierungen harmonischer Strukturen eingesetzt, um das Konzerterlebnis von klassischen Konzertaufführungen zu bereichern [LGS15]. Grundlage für unseren Beitrag sind die in [WKM16] vorgestellten Methoden zur Visualisierung von Akkorden und Skalen. Hierbei werden in einem gewissen mathematischen Sinne Wahrscheinlichkeiten für das Auftreten lokaler harmonischer Strukturen über die Zeit hinweg dargestellt. Zum Beispiel zeigt Abb. 1c eine solche Visualisierung hinsichtlich diatonischer Skalen.

In diesem Artikel wollen wir die in [WKM16] vorgestellten Konzepte erweitern. Als einen wesentlichen Beitrag führen wir ein neuartiges Visualisierungskonzept ein, das eine simultane Darstellung von unterschiedlichen Analyseergebnissen ermöglicht. Hierbei ist

anzumerken, dass automatisierte Verfahren stets von konkreten Darstellungen des musikalischen Werks ausgehen, z. B. einem digitalisierten Notentext, einer MIDI-Darstellung oder einer CD-Einspielung. Das Analyseergebnis hängt damit nicht nur von werkimmanenten, sondern auch von darstellungsspezifischen Eigenschaften ab. In unserer Visualisierung führen wir ein geeignetes Farbschema ein, das sowohl die Gemeinsamkeiten als auch die Unterschiede der Analyseergebnisse über verschiedene Versionen hinweg darstellen kann. Eine solche Visualisierung ist in Abb. 1c dargestellt für den Fall von drei unterschiedlichen Einspielungen, dirigiert von Herbert von Karajan (Berliner Philharmoniker, aufgenommen 1966–1970, Deutsche Grammophon), Bernard Haitink (Symphonieorchester des Bayerischen Rundfunks 1988–1991, EMI) und Daniel Barenboim (Bayreuther Festspiele 1991–1992, Warner Classics). Unsere Hypothese ist, dass Konsistenzbereiche in solchen Darstellungen (in Abb. 1c durch die Farbe „Schwarz“ kodiert) auf werkimmanente Eigenschaften hinweisen, was in ähnlicher Weise auch in [KMK13] untersucht wurde.

Als einen weiteren Beitrag dieser Arbeit diskutieren wir, wie unsere Visualisierungskonzepte in Benutzerschnittstellen integriert werden können, die einem Musikwissenschaftler ermöglichen, die berechneten Analyseergebnisse intuitiv zu erfassen sowie in unterschiedlichen Auflösungen hierarchisch zu durchsuchen, zu interpretieren und interaktiv anzupassen. Anhand der Tetralogie *Der Ring des Nibelungen* von Richard Wagner deuten wir an, wie man auf diese Weise großformatige Übersichten harmonischer Verläufe für ganze Werkzyklen erzeugen kann. Die Verifizierung der musikwissenschaftlichen Relevanz solcher Konzepte erfolgt im Rahmen eines von der DFG geförderten Kooperationsprojektes.<sup>3</sup>

Der Artikel gliedert sich wie folgt: Zunächst gehen wir in Abschnitt 2 auf unser Fallbeispiel, *Der Ring des Nibelungen*, ein und erläutern die Relevanz des Beispiels für das Projekt. In Abschnitt 3 fassen wir unterschiedliche Analysemethoden zusammen, welche die Grundlage für diese Arbeit bilden. Daraufhin stellen wir in Abschnitt 4 unser versionsübergreifendes Visualisierungskonzept sowie das vorgeschlagene Farbschema vor. In Abschnitt 5 diskutieren wir abschließend die Ergebnisse und geben einen Ausblick auf weitere Forschungsarbeiten.

## 2 Fallbeispiel: Wagners „Ring“

In diesem Artikel stellen wir aufbauend auf [WKM16] ein Visualisierungskonzept vor, welches wir dann exemplarisch auf Richard Wagners Tetralogie *Der Ring des Nibelungen* anwenden. Beim *Ring* handelt es sich um einen Zyklus von vier musikdramatischen Werken (siehe Abb. 1), welche in Aufzüge (mit Ausnahme des *Rheingolds*) und Szenen untergliedert sind. Eine Aufführung des gesamten Zyklus nimmt bis zu 15 Stunden in Anspruch. Bei einem Werk solcher Länge kommen die Vorteile computergestützter Methoden in besonderer Weise zur Geltung. Werkzeuge zur Suche, Navigation und Visualisierung erleichtern den Zugang und ermöglichen einen Überblick über solche umfangreichen Musikdatenbestände. Zum Beispiel deutet Abb. 1b eine Visualisierung des harmonischen Verlaufs für den

<sup>3</sup> Computergestützte Analyse harmonischer Strukturen (DFG MU 2686/7-1, DFG KL 864/4-1)

zweiten Teil, *Die Walküre*, an. Nicht nur aufgrund seines großen Umfangs stellt der *Ring* eine Herausforderung dar. Bei der harmonischen Analyse des *Rings* wird auch musikwissenschaftlich weitgehend Neuland betreten. Hierzu sind großformatige Analysen tonaler Strukturen und deren Bezüge erforderlich. Eine der wenigen Arbeiten in dieser Richtung sind die Analysen von Alfred Lorenz [Lo24] aus den 1920er Jahren, welche in der Musikwissenschaft höchst umstritten sind [K114]. Die Validität der Analysen von Lorenz in ihrem methodischen Ansatz und ihrer praktischen Umsetzung wurde in den 1960er Jahren von Carl Dahlhaus hinterfragt [Da65, Br02]. Um Beschreibungen der harmonischen Anlage des *Rings* zu entwickeln, welche sich näher an den tatsächlich beobachteten Phänomenen orientieren, sind objektivere Analysen nötig. Hierfür kann die automatisierte Messung und Visualisierung tonaler Strukturen ein geeignetes Werkzeug darstellen.

### 3 Computergestützte Harmonieanalyse

Ein wesentlicher Bestandteil musikalischer Analyse ist die Erkennung von tonalen Strukturen wie Akkorden, Skalen oder Tonarten im Verlauf eines Werkes. Im MIR-Bereich werden solche Analysen mittels automatisierter Verfahren umgesetzt. Die Grundlage für solche Analysen bilden meist sogenannte *Chromagramm-Darstellungen*, welche die Energieverteilung des Musiksignals über die zwölf chromatischen Tonhöhenklassen im Verlauf der Zeit erfassen. Dazu wird das Signal zunächst in lokale Zeitfenster der Breite  $w$  (window size) unterteilt. Innerhalb eines Fensters werden nun die Anteile verschiedener Frequenzen berechnet, was unter anderem durch die Fouriertransformation realisiert werden kann [Mü15, Kapitel 2]. Dieses Zeitfenster wird nun über das Signal geschoben (shift), wobei der Vorschub durch einen Parameter  $h$  (hop size) festgelegt wird. Die Chroma-Darstellungen können grafisch visualisiert werden, wie in Abb. 2a gezeigt wird. In diesem Beispiel sind die Zeitfenster bereits auf einer musikalischen Zeitachse spezifiziert ( $w = 1$  Takt, ebenso  $h = 1$  Takt), was eine vergleichende Betrachtung von Notentext und Visualisierung ermöglicht. Zu diesem Zweck muss vorher die Position der Taktgrenzen manuell bestimmt oder durch geeignete Algorithmen ermittelt werden [KMK13, Za17]. Für die drei Aufnahmen, die in dieser Arbeit herangezogen werden, wurden alle Taktgrenzen durch Studenten anhand des Klavierauszugs händisch markiert. Jede Einspielung wurde dabei von mindestens zwei Annotatoren bearbeitet und größere Abweichungen anschließend nachkorrigiert. Einzelheiten zu diesem Vorgehen sowie eine Analyse der Konsistenz zwischen den Annotatoren finden sich in [We16].

Für jeden Zeitpunkt  $m$  zeigt die Visualisierung in Abb. 2a die lokale Energieverteilung über die zwölf Tonhöhenklassen in dem jeweiligen Zeitfenster (zentriert um  $m$ ).<sup>4</sup> Hohe Energie wird durch einen dunklen (schwarzen) Farbwert dargestellt, niedrige Energie durch einen hellen (weißen) Farbwert. Grauwerte deuten auf mittlere Energie hin. In einstimmigen Passagen lassen sich die Visualisierungen der Chroma-Darstellungen besonders intuitiv

---

<sup>4</sup> Die enharmonische Unterscheidung von Tonhöhen wie z. B. *cis* und *des* ist hierbei nicht möglich, was einen wesentlichen Unterschied zur Analyse symbolischer Musikdaten wie Partituren darstellt.

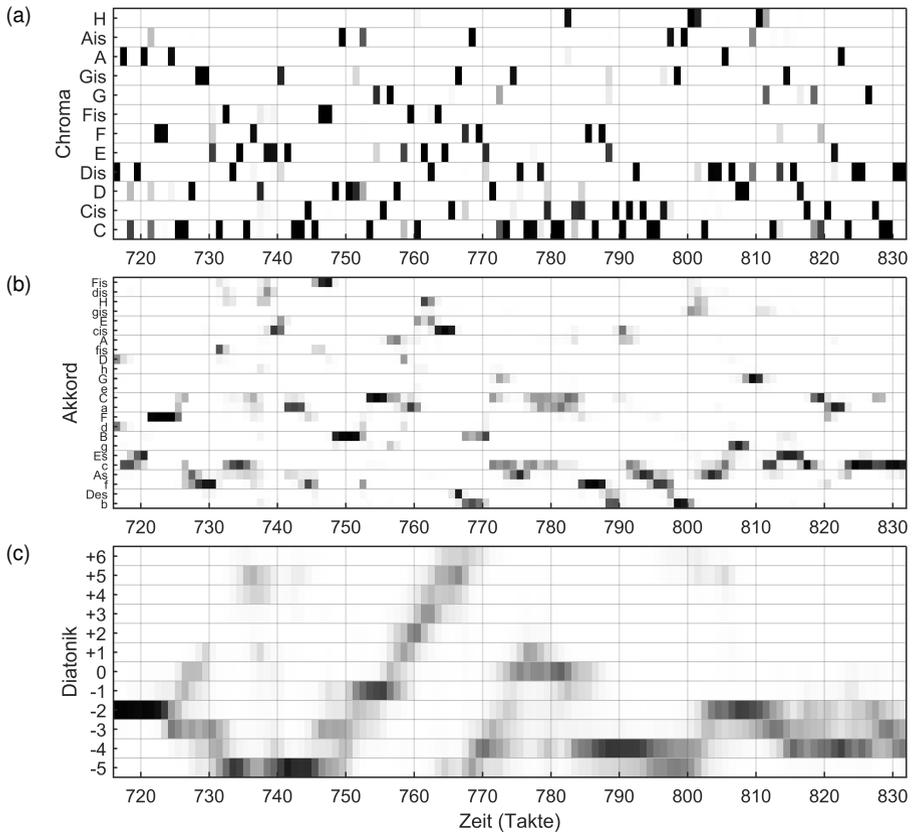


Abb. 2: Unterschiedliche Darstellungen für die Takte 716–832 aus dem dritten Akt der *Walküre*. **(a)** Chromagramm mit Fenstergröße  $w = 1$  Takt. **(b)** Zeit-Akkord Darstellung mit Fenstergröße  $w = 4$  Takte. **(c)** Zeit-Skalen Darstellung mit Fenstergröße  $w = 12$  Takte.

verfolgen. So wird beispielsweise in T. 719–720 ein *es* (enharmonisch verwechselt *dis*) gehalten, in T. 722–723 ein *f* und in T. 725–726 ein *c*, was in Abb. 2a deutlich sichtbar ist.

Die Betrachtung von Tonhöhenklassen hat jedoch nur begrenzte Aussagekraft für harmonisch komplexere Sachverhalte. Sinnvollere Kategorien für diesen Zweck stellen Intervalle, Akkorde oder Skalen dar. Im vorliegenden Beitrag konzentrieren wir uns bei der Akkordmessung auf die 24 Dur- und Molldreiklänge<sup>5</sup> und bei der Skalenmessung auf die zwölf diatonischen

<sup>5</sup> Aufgrund der enharmonischen Invarianz der Chromagramm-Darstellungen können Dreiklänge wie z. B. Cis-Dur und Des-Dur nicht unterschieden werden.

Skalen, welche hier in Analogie zur jeweils entsprechenden Vorzeichnung benannt werden.<sup>6</sup> Bei der Berechnung der harmonischen Strukturen werden die Chromagramm-Darstellungen lokal mit binären Prototypen (templates) verglichen, welche den Akkorden [SE03, MD10] oder Skalen [WH14] entsprechen. Ein C-Dur-Dreiklang wird beispielsweise durch den Prototyp

$$\mathbf{P}_{\text{C-Dur}} = (1\ 0\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0)^T \quad (1)$$

modelliert. Über den lokalen Abgleich können Wahrscheinlichkeiten für das Auftreten dieser harmonischen Strukturen ermittelt werden. Im Falle diatonischer Skalen erhält man eine Matrix von errechneten Analyseergebnissen  $\mathcal{A} \in \mathbb{R}^{D \times M}$ . Für eine Diatonik mit Index  $d \in [0 : D - 1] := \{0, 1, \dots, D - 1\}$  mit  $D = 12$  und einen Taktindex  $m \in [1 : M]$  kann der Eintrag  $\mathcal{A}(d, m)$  als Wahrscheinlichkeit für das Auftreten der diatonischen Skala  $d$  im Takt  $m$  interpretiert werden, da die Ergebnisse vorher geeignet normiert wurden. Diese Analysematrix  $\mathcal{A}$  kann nun durch ein Graustufen-Schema visualisiert werden (wie in Abb. 2b), wobei hier dunkle Farben eine hohe Wahrscheinlichkeit andeuten. Zur Hervorhebung der wichtigen Strukturen können weitere Verarbeitungsschritte erforderlich sein. Beispielsweise kann eine zeitliche Glättung oder eine Anhebung der höheren Energiewerte sowie eine Unterdrückung der niedrigen, rauschartigen Werte zu deutlicheren visuellen Strukturen führen. Letzteres wird im vorliegenden Beitrag mittels einer exponentiellen Umskalierung der Energiewerte realisiert (ähnlich der Softmax-Funktion). Bei der grafischen Darstellung dieser Strukturen hat es sich als nützlich erwiesen, die vertikale Achse nach dem Quintenzirkel anzuordnen, um die tonalen Bezüge zwischen Akkorden und Skalen zu berücksichtigen. Für weitere Details zur Berechnung und Visualisierung verweisen wir auf [SE03, KMK13, WH14].

In Abb. 2 demonstrieren wir verschiedene Visualisierungskonzepte am Beispiel eines Ausschnitts aus der *Walküre*. Bei der Betrachtung der Skalenanalyse (Abb. 2c) lässt sich in den Takten 716 bis ca. 725 eine hohe Wahrscheinlichkeit für die  $-2$ -Diatonik erkennen. Daraus kann man auf die Tonarten B-Dur oder g-Moll in diesem Bereich schließen. Eine vergleichende Analyse des Notentexts zeigt tatsächlich einen deutlichen g-Moll-Charakter bis T. 721. Als weiteres Beispiel lässt sich im Notentext ein klarer f-Moll-Eindruck für die Passage T. 785–798 feststellen. In den Visualisierungen finden wir dies wieder, angedeutet durch eine hohe Wahrscheinlichkeit für die  $-4$ -Diatonik in Abb. 2c sowie durch die Dominanz des f-Moll-Dreiklangs in Abb. 2b.

## 4 Versionsübergreifende Visualisierung

Als Besonderheit sogenannter „klassischer“ Musik liegen Werke oft in einer Vielzahl verschiedener Einspielungen (Versionen) vor. Diese können in der musikalischen Gestaltung hinsichtlich des Tempoverlaufs, der Phrasierung oder der Artikulation erheblich voneinander abweichen. Der tonale Gehalt der Musik sollte von aufführungsspezifischen Besonderheiten

<sup>6</sup> Die Skala  $+1$  bezieht sich beispielsweise auf eine Diatonik mit einem  $\sharp$  (entspricht dem Tonvorrat von G-Dur oder e-Moll),  $-3$  auf eine Diatonik mit drei  $b$  (Es-Dur oder c-Moll), und  $0$  auf die Diatonik ohne Vorzeichen (C-Dur oder a-Moll), welche den weißen Tasten des Klaviers entspricht.

hingegen weitgehend unberührt bleiben und gegenüber Detailvariationen wie Vibrato oder Feinheiten in der Intonation stabil sein.<sup>7</sup> Unsere Idee ist es, das Vorliegen mehrerer Versionen für die audiobasierte Harmonieanalyse auszunutzen, indem die Analyse separat für mehrere Einspielungen durchgeführt wird. Dazu berechnen wir die Analysen auf Basis einer musikalischen Zeitachse (taktweise, siehe oben), was einen Vergleich der Ergebnisse über verschiedene Versionen hinweg ermöglicht [Za17]. Auf dieser Basis stellen wir Gemeinsamkeiten und Unterschiede in den Ergebnissen grafisch dar. Bereits in [KMK13] wurde eine solche Strategie zur Analyse einer Klaviersonate auf Basis mehrerer Versionen herangezogen. Im Folgenden führen wir ein Farbschema ein, welches im Speziellen den Vergleich von Analyseergebnissen über drei Versionen ermöglicht. Dabei können sowohl die Konsistenz der Versionen als auch die Ergebnisse der Einzelversionen visuell nachverfolgt werden.

Für diese versionsübergreifende Darstellung berechnen wir die Analysematrix  $\mathcal{A}$  zunächst für verschiedene Einspielungen separat. Für  $N$  verschiedene Einspielungen (Versionen) erhalten wir somit  $N$  Analysematrizen  $\mathcal{A}_n$  mit  $n \in [0 : N-1]$ . Einen beliebigen Eintrag dieser Matrizen für die verschiedenen Versionen kürzen wir im Folgenden durch  $a_n := \mathcal{A}_n(d, m)$  ab. Die Konsistenz dieser Einträge  $a_n$  über unterschiedliche Versionen  $n$  wird nun durch eine Farbkodierung dargestellt. In unserem Fall von  $N = 3$  Versionen nutzen wir hierfür die Parametrisierung des additiven RGB-Farbraums (Rot, Grün, Blau). Eine Farbe wird im Folgenden durch einen dreielementigen Vektor  $\mathbf{c} := (c_r, c_g, c_b)^T \in \mathbb{R}^3$  modelliert. Jeder Eintrag nimmt Werte im Bereich  $[0, 1]$  an, wobei Schwarz mit  $\mathbf{c} = (0, 0, 0)^T$  und Weiß mit  $\mathbf{c} = (1, 1, 1)^T$  die Grenzfälle darstellen.

Für die Farbmodellierung fassen wir  $\mathbf{c}$  im Folgenden als eine vektorwertige Funktion  $\mathbf{c} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  auf, welche die drei Analyseergebnisse  $a_n$  auf den RGB-Farbraum abbildet. Für die Funktionswerte  $\mathbf{c}(a_0, a_1, a_2)$  betrachten wir folgende Randbedingungen: Bei Übereinstimmung aller Versionen für einen Eintrag der Analysematrix ( $a := a_0 = a_1 = a_2$ ) sollen sich die Farben zu einem Grauwert addieren. Für den Maximalwert  $a = 1$  (hohe Wahrscheinlichkeit) soll sich der Farbvektor für Schwarz ergeben:

$$\mathbf{c}(1, 1, 1) := (0, 0, 0)^T. \quad (2)$$

Falls alle Versionen übereinstimmend den Minimalwert  $a = 0$  liefern, soll der Farbvektor für Weiß resultieren:

$$\mathbf{c}(0, 0, 0) := (1, 1, 1)^T. \quad (3)$$

Für Zwischenwerte  $0 < a < 1$ , welche jedoch über die Versionen hinweg übereinstimmen, soll sich entsprechend ein mittlerer Grauwert ergeben. Stimmen die Analyseergebnisse der Einzelversionen nicht überein, soll sich nun ein echter (nicht grauer) Farbwert ergeben. Falls

<sup>7</sup> Globale Aspekte wie eine Transposition oder abweichende Grundstimmungen stellen größere Probleme für die versionsübergreifende Harmonieanalyse dar. Solche Unterschiede müssen vor der Analyse detektiert und entsprechend berücksichtigt werden.

die erste Version den Maximalwert  $a_0 = 1$  liefert (die anderen beiden  $a_1 = a_2 = 0$ ) soll der Farbvektor für Rot resultieren:

$$\mathbf{c}(1, 0, 0) := (1, 0, 0)^T. \quad (4)$$

In entsprechender Weise soll die zweite Version ( $n = 1$ ) zu Blau und die dritte Version ( $n = 2$ ) zu Grün korrespondieren:

$$\mathbf{c}(0, 1, 0) := (0, 1, 0)^T, \quad (5)$$

$$\mathbf{c}(0, 0, 1) := (0, 0, 1)^T. \quad (6)$$

Damit stellen wir ein möglichst einfaches Gleichungssystem (linear in den Parametern) auf

$$\mathbf{c}(a_0, a_1, a_2) = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \end{pmatrix} + a_0 \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{pmatrix} + a_1 \begin{pmatrix} \gamma_0 \\ \gamma_1 \\ \gamma_2 \end{pmatrix} + a_2 \begin{pmatrix} \delta_0 \\ \delta_1 \\ \delta_2 \end{pmatrix} + \begin{pmatrix} a_1 a_2 \cdot \epsilon_0 \\ a_0 a_2 \cdot \epsilon_1 \\ a_0 a_1 \cdot \epsilon_2 \end{pmatrix}, \quad (7)$$

mit insgesamt 15 Parametern  $\alpha_k, \beta_k, \gamma_k, \delta_k$  und  $\epsilon_k$ , wobei  $k \in \{0, 1, 2\}$ . Die gemischten Terme sind erforderlich, um die 15 Freiheitsgrade der Randbedingungen (2)–(6) erfüllen zu können. Durch Lösung des Gleichungssystems folgt die folgende Berechnungsvorschrift:

$$\mathbf{c}(a_0, a_1, a_2) = \begin{pmatrix} c_r \\ c_g \\ c_b \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} a_1 + a_2 - a_1 a_2 \\ a_0 + a_2 - a_0 a_2 \\ a_0 + a_1 - a_0 a_1 \end{pmatrix}. \quad (8)$$

Diese Berechnung wird nun für jeden Eintrag der Analysematrizen  $\mathcal{A}_n(d, m)$  durchgeführt, mit  $d \in [0 : D - 1]$ ,  $m \in [1 : M]$  sowie  $n \in [0 : N - 1]$  wobei  $N = 3$ . Als Resultat erhält man einen Tensor dritter Stufe  $C \in \mathbb{R}^{D \times M \times N}$  mit Einträgen  $C(d, m, n)$ , deren letzter Index  $n \in \{0, 1, 2\}$  sich auf die drei Dimensionen des RGB-Farbraums bezieht.

## 5 Diskussion und Ausblick

Wir wollen nun die Eigenschaften des Visualisierungskonzeptes diskutieren. Mit Hilfe des eingeführten Farbschemas stellen wir in Abb. 3 verschiedene Analysen einer Passage aus dem dritten Akt der *Walküre* auf Basis dreier Einspielungen (dirigiert von Karajan, Haitink und Barenboim) dar. Abb. 3a zeigt ein Chromagramm mit einer Fensterbreite  $w = 1$  Takt. Das grafische Resultat ergibt ein unübersichtliches Bild, da sich die beobachteten Muster von Takt zu Takt meist stark ändern. Diese Wechselhaftigkeit ist nicht weiter verwunderlich, stellt das Chromagramm doch Tonhöhen-eigenschaften einzelner Töne über einen Zeitraum von mehr als 100 Takten dar. Die hohe Farbigkeit impliziert eine vergleichsweise geringe Konsistenz der Analyse über die Versionen hinweg. Dies ist zunächst verwunderlich, da ein sehr ähnlicher tonaler Verlauf in allen Einspielungen zu erwarten ist und globale Effekte



liegen, aber auch in Ungenauigkeiten der Intonation. Häufiger anzutreffen sind auch Terz- oder Quint-Verwechslungen wie in T. 757 (*d, f, a*). Diese können unter anderem von den Tönen eines Dreiklangs herrühren, welche in verschiedenen Einspielungen aufgrund der individuellen Klangbalance zwischen den Instrumenten unterschiedlich stark präsent sind. Außerdem kann die individuelle Klangfarbe der Instrumente oder Sänger zu einer abweichenden Präsenz von Obertönen führen, welche sich im Chromagramm niederschlägt. Solche und ähnliche Effekte werden durch die exponentielle Umskalierung noch verstärkt, was zu den starken versionsabhängigen Unterschieden in der Visualisierung führt.

In Abb. 3b gehen wir nun zu einer Dreiklangsanalyse über, mit Fensterbreite  $w = 1$  Takt. Auch hier zeigt sich eine kleinteilige Struktur mit taktweise stark wechselnden Analyseergebnissen. Versionsabhängige Unterschiede treten oft zwischen Akkorden mit gemeinsamen Tönen auf, welche in der gewählten Anordnung meist benachbart sind. Ein Beispiel ist in T. 724 zu beobachten (a-Moll, F-Dur und d-Moll). Auch Verwechslungen zwischen Variantklängen treten auf wie in T. 742 (a-Moll und A-Dur). In T. 771–785 ist eine längere Phase mit niedriger Konsistenz zu beobachten. Diese Stelle zeichnet sich durch weitgehend einstimmigen Gesang mit stark wechselnder Harmonik aus, was oft zu instabilen Analyseergebnissen führt. In ähnlicher Weise sind viele Beobachtungen anhand des Klavierauszugs gut nachzuvollziehen. Die Übersichtlichkeit der Darstellung ist bei einer solchen kleinteiligen Analyse jedoch begrenzt, da sich kaum Trends über größere Zeiträume erkennen lassen.

Um die globalen Strukturen stärker hervorzuheben, gehen wir nun zu größeren Analysefenstern über. Abb. 3c zeigt eine Dreiklangsanalyse mit einer Fensterbreite von  $w = 12$  Takte. Das visuelle Ergebnis gewinnt dadurch an Deutlichkeit. Beispielsweise lässt sich in T. 723–733 ein f-Moll-Dreiklang erkennen oder in T. 818–829 ein c-Moll-Akkord, beide mit hoher Konsistenz. Die grobe zeitliche Glättung mag zunächst irritieren, erklingt doch selten ein einzelner Akkord über einen Zeitraum von zwölf Takten. Es zeigt sich jedoch, dass geglättete Chromagramme die dominanten Tonhöhen über einen solchen Zeitraum gut hervorheben und auf diese Weise zentrale tonale Strukturen erkennen lassen, welche oft mit Dreiklängen korrespondieren. Beispielsweise lässt der Klavierauszug im Abschnitt T. 723–733 neben f-Moll auch mehrere verwandte Akkorde wie Des-Dur oder ein Dominantseptakkord auf *c* erkennen. Diese stehen jedoch in einem engen klanglichen Verhältnis mit dem f-Moll-Akkord und bestätigen diesen als tonales Zentrum (vgl. Abb. 4). Interessanterweise führt die zeitliche Glättung auch zu höherer Konsistenz zwischen den Versionen. Lediglich tonal instabile Stellen, welche sich von einer Dreiklangsharmonik entfernen, führen zu Inkonsistenzen wie in T. 771–785. Letzteres ist in der Wagnerschen Musik aufgrund ihrer harmonischen Komplexität häufig der Fall, weshalb eine Analyse hinsichtlich Dur- und Moll dreiklängen hier teilweise an ihre Grenzen stößt. Allerdings hat sich herausgestellt, dass die Dreiklangsanalyse besonders für harmonisch stabile Stellen zu Konsistenzen in den Visualisierungen führt. Somit kann ein guter Überblick über die tonalen Ankerpunkte der Musik gewonnen werden.

Neben dem zeitlichen Filtern durch größere Analysefenster haben wir bereits ein „musikalisches Filtern“ angewandt, indem wir von einzelnen Tonhöhenklassen zu Akkorden



welche aus einer kleinteiligen Analyse nur schwer ersichtlich ist. Die beschriebene Stelle zeigt auf, wie die vorgestellten Visualisierungen in der musikwissenschaftlichen Praxis von Nutzen sein können. Tatsächlich trat der besondere tonale Charakter dieser Passage bei der gemeinsamen Projektarbeit zwischen Musik- und Ingenieurwissenschaften zum Vorschein. Die Grafiken stellten hier den Ausgangspunkt dar und motivierten eine nähere Betrachtung der betreffenden Stelle. Im Gegensatz zur Verifizierung bereits bestehender Hypothesen, wie sie in [WKM16] durchgeführt wurde, können so neue Sachverhalte aufgespürt und untersucht werden. Wie im vorliegenden Beitrag angedeutet wurde, spielt dabei die Wahl der Analyseparameter eine entscheidende Rolle. Eine universell anwendbare, „richtige“ Parameterwahl gibt es dabei häufig nicht, da diese üblicherweise vom jeweiligen musikalischen Kontext abhängt. Die schrittweise Einstellung insbesondere der zeitlichen und musikalischen Glättungsparameter kann jedoch zum Hervortreten charakteristischer Muster führen oder diese in detaillierte Bestandteile auflösen. Um dies in einer interaktiven Arbeitsweise umsetzen zu können, arbeiten wir an einer Benutzerschnittstelle, mit der sich wesentliche Parameter auf intuitive Weise ändern lassen (siehe Abb. 4). Eine solche Arbeitsmethode ermöglicht die schnelle Navigation innerhalb des *Rings*, die Fokussierung auf verschiedene harmonische Konzepte sowie einen visuellen Überblick über großformatige Strukturen und kann so zu neuen Erkenntnissen in der Musikwissenschaft führen.

### **Danksagungen**

Die International Audio Laboratories Erlangen sind eine gemeinsame Einrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer-Instituts für Integrierte Schaltungen IIS. Die vorgestellte Arbeit wurde durch die Deutsche Forschungsgemeinschaft gefördert (DFG MU 2686/7-1, DFG KL 864/4-1). Wir bedanken uns bei Vlora Arifi-Müller und den studentischen Hilfskräften für die Erstellung der Musikannotationen.

### **Literaturverzeichnis**

- [Br02] Breig, Werner: Wagners Begriff der „dichterisch-musikalischen Periode“. In (Döge, Klaus; Jost, Christa; Jost, Peter, Hrsg.): „Schlagen Sie die Kraft der Reflexion nicht zu gering an.“ Beiträge zu Richard Wagners Denken, Werk und Wirken, S. 158–172. Schott, Mainz, 2002.
- [CB14] Cho, Taemin; Bello, Juan Pablo: On the Relative Importance of Individual Components of Chord Recognition Systems. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(2):477–492, 2014.
- [Da65] Dahlhaus, Carl: Wagners Begriff der „dichterisch-musikalischen Periode“. In (Salmen, Walter, Hrsg.): Beiträge zur Geschichte der Musikanschauung im 19. Jahrhundert. G. Bosse, Regensburg, S. 179–187, 1965.
- [K114] Kleinertz, Rainer: Richard Wagners Begriff der „dichterisch-musikalischen Periode“. *Die Musikforschung*, 67:26–47, 2014.
- [KMK13] Konz, Verena; Müller, Meinard; Kleinertz, Rainer: A Cross-Version Chord Labelling Approach for Exploring Harmonic Structures—A Case Study on Beethoven’s *Appassionata*. *Journal of New Music Research*, 42(1):61–77, 2013.

- [LGS15] Liem, Cynthia C. S.; Gómez, Emilia; Schedl, Markus: PHENICX: Innovating the Classical Music Experience. In: Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW). Turin, Italien, S. 1–4, 2015.
- [Lo24] Lorenz, Alfred O.: Der musikalische Aufbau des Bühnenfestspiels Der Ring des Nibelungen. In: Das Geheimnis der Form bei Richard Wagner. M. Hesse, Berlin, 1924.
- [Mc14] McVicar, Matt; Santos-Rodriguez, Raúl; Ni, Yizhao; Bie, Tijn De: Automatic Chord Estimation from Audio: A Review of the State of the Art. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 22(2):556–575, 2014.
- [MD10] Mauch, Matthias; Dixon, Simon: Simultaneous Estimation of Chords and Musical Context from Audio. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(6):1280–1289, 2010.
- [Mü15] Müller, Meinard: *Fundamentals of Music Processing*. Springer, Cham, Schweiz, 2015.
- [NS06] Noland, Katy; Sandler, Mark: Key Estimation Using a Hidden Markov Model. In: Proceedings of the International Society for Music Information Retrieval Conference (ISMIR). Victoria, Kanada, S. 121–126, 2006.
- [PP11] Papadopoulos, Hélène; Peeters, Geoffroy: Joint Estimation of Chords and Downbeats From an Audio Signal. *IEEE Transactions on Audio, Speech, and Language Processing*, 19(1):138–152, 2011.
- [Sa05] Sapp, Craig Stuart: Visual Hierarchical Key Analysis. *ACM Computers in Entertainment*, 3(4):1–19, 2005.
- [SE03] Sheh, Alexander; Ellis, Daniel P.W.: Chord Segmentation and Recognition Using EM-trained Hidden Markov Models. In: Proceedings of the International Society for Music Information Retrieval Conference (ISMIR). Baltimore, USA, S. 185–191, 2003.
- [We16] Weiß, Christof; Arifi-Müller, Vlora; Prätzlich, Thomas; Kleinertz, Rainer; Müller, Meinard: Analyzing Measure Annotations for Western Classical Music Recordings. In: Proceedings of the International Conference on Music Information Retrieval (ISMIR). New York, USA, S. 517–523, 2016.
- [WH14] Weiß, Christof; Habryka, Julian: Chroma-Based Scale Matching for Audio Tonality Analysis. In: Proceedings of the Conference on Interdisciplinary Musicology (CIM). Berlin, S. 186–173, 2014.
- [WKM16] Weiß, Christof; Kleinertz, Rainer; Müller, Meinard: Möglichkeiten der computergestützten Erkennung und Visualisierung harmonischer Strukturen – eine Fallstudie zu Richard Wagners “Die Walküre”. In (Auhagen, Wolfgang; Hirschmann, Wolfgang, Hrsg.): Bericht zur Jahrestagung der Gesellschaft für Musikforschung (GfM) 2015 in Halle/Saale. Schott Campus, Mainz, 2016.
- [Za17] Zalkow, Frank; Weiß, Christof; Prätzlich, Thomas; Arifi-Müller, Vlora; Müller, Meinard: A Multi-Version Approach for Transferring Measure Annotations Between Music Recordings. In: Proceedings of the AES International Conference on Semantic Audio. Erlangen, 2017.