

Rechnergestützte Musikethnologie am Beispiel historischer Aufnahmen mehrstimmiger georgischer Vokalmusik

Frank Scherbaum,¹ Meinard Müller,² Sebastian Rosenzweig²

Abstract: Historische Audioaufnahmen stellen in Fällen von mündlich überlieferten Musiktraditionen oft die einzigen Quellen über frühere musikalische Aufführungspraktiken dar. Wenn, wie im Fall der hier betrachteten mehrstimmigen georgischen Vokalmusik, schriftliche musiktheoretische Zeugnisse fehlen, stellt zum Beispiel schon die Rekonstruktion von früher verwendeten Tonleitern eine Herausforderung dar, die ohne historische Audioaufnahmen gar nicht lösbar wäre. Anhand eines Datensatzes von etwa 100 Audioaufnahmen traditioneller, mehrstimmiger georgischer Vokalmusik, die von dem gurischen Meistersänger Artem Erkomaishvili im Jahr 1966 eingesungen wurden, wird an ausgewählten Beispielen illustriert, wie die Auswertung und Interpretation musikethnologisch relevanten Audiomaterials mittels rechnergestützter Verfahren und Webtechnologie unterstützt werden kann.

Keywords: Musikethnologie. Musikinformatik. Georgische Vokalmusik. Audioverarbeitung. Aufführungspraxis.

1 Einleitung

Die traditionelle Praxis musikethnologischer Forschung besteht im wesentlichen in der Dokumentation musikalischer Aufführungen in Form schriftlicher Notizen sowie von Audio- und Videoaufzeichnungen [Ri14]. Das Ziel der Auswertung des aufgezeichneten Materials ist es dann, die Charakteristika der untersuchten Musiktradition in ihrem persönlichen, sozialen und kulturellem Kontext zu verstehen. Dabei spielt die musikalische Analyse des Audiomaterials, welche zunehmend mit Hilfe rechnergestützter Verfahren durchgeführt werden kann, eine ganz wesentliche Rolle.

Mehrstimmige georgische Vokalmusik, welche Gegenstand der hier vorgestellten Arbeit ist, stellt einen weltweit einmaligen Kulturschatz dar, der seit 2001 in die Liste des erhaltungsbedürftigen, immateriellen Kultererbes der Menschheit aufgenommen wurde. Anhand eines Datensatzes von etwa 100 historischen Audioaufnahmen traditioneller, mehrstimmiger georgischer Vokalmusik, die von dem gurischen Meistersänger Artem Erkomaishvili im Jahr 1966 eingesungen wurden, wird im Folgenden an ausgewählten Beispielen illustriert, wie die Auswertung und Interpretation musikethnologisch relevanten Audiomaterials mittels rechnergestützter Verfahren und Webtechnologie unterstützt werden

¹ Institut für Erd- und Umweltwissenschaften, Universität Potsdam, fs@geo.uni-potsdam.de

² International Audio Laboratories Erlangen, meinard.mueller@audiolabs-erlangen.de

kann. Dabei beschränken wir uns hier auf einen Aspekt der Analyse und Interpretation des melodischen Intervallgehaltes dieser Gesänge und diskutieren, welche Randbedingungen sich daraus hinsichtlich des „musikalischen Denkens“ von Artem Erkomaishvili ableiten lassen. Die hier vorgestellten Betrachtungen komplementieren die in [SMR17] beschriebenen Studien, welche sich ebenfalls auf den hier betrachteten Datensatz beziehen.

Der Artikel ist wie folgt gegliedert. Zunächst beleuchten wir in Abschnitt 2 den historischen Hintergrund zu den Tonaufnahmen von Artem Erkomaishvili und beschreiben deren Aufnahmeprozess. Daraufhin gehen wir in Abschnitt 3 auf Fragestellungen der Datenaufbereitung ein und beschreiben die für den Erkomaishvili-Datensatz erzeugten Annotationen. Weiterhin zeigen wir, wie man durch Einsatz von webbasierten Technologien den Zugriff auf die Tonaufnahmen und Annotationen erheblich vereinfachen kann. In Abschnitt 4 deuten wir das Potential rechnergestützter Verfahren für die musikethnologische Forschung an indem wir exemplarisch der Fragestellung nach der Beschaffenheit georgischer Skalenmodelle nachgehen. Der Artikel wird in Abschnitt 5 durch einen kurzen Ausblick abgerundet.

2 Hintergrund zu Erkomaishvilis Tonaufnahmen

Georgien hat eine lange und reiche Musikgeschichte, wobei die Ursprünge der traditionellen, mehrstimmigen Vokalmusik allerdings weitgehend ungeklärt sind. Für „westliche Ohren“ enthält diese oft als ungewöhnlich empfundene Intervalle, Harmonien und Harmoniefolgen und zeichnet sich durch eine beeindruckende harmonische und melodische Vielfalt aus [Zi09]. Ein großes Problem bezüglich ihrer Bewahrung in authentischer Form liegt darin, dass sie traditionell nur mündlich überliefert wurde und in der Vergangenheit praktisch nie „unverzerrt“ dokumentiert werden konnte, da hierfür die technischen Voraussetzungen fehlten. Es kann allerdings davon ausgegangen werden, dass diese Musik ursprünglich nicht in temperierter Stimmung, in welcher eine Oktave in 12 gleiche Intervalle eingeteilt ist, gesungen wurde.

Wird für untemperiert gesungene Musik eine Transkription vorgenommen, so muss ihre mikrotonale Struktur bei der Interpretation berücksichtigt werden. Dies ist in der Regel sehr aufwändig und immer subjektiv und führt sehr schnell zu unvermeidbaren Transkriptionsproblemen. Dadurch kommt es dann leicht zu Mehrdeutigkeiten in der Interpretation der melodischen und harmonischen Strukturen der Musik. Trotz des seit den 80er Jahren des letzten Jahrhunderts auch international stetig zunehmenden Interesses an der Erforschung und der Aufführung der georgischen Vokalpolyphonie werden zentrale Fragen zum Beispiel bzgl. ihrer melodischen und harmonischen Struktur bzw. der ursprünglich verwendeten Tonleitern noch immer sehr kontrovers diskutiert, wie die Tagungsbände der seit 2002 durch das Internationale Forschungszentrum für traditionelle Polyphonie regelmässig durchgeführten Symposia eindrücklich dokumentieren.

In diesem Zusammenhang ist eine Sammlung von Audioaufnahmen von Artem Erkomaishvili (1887-1967) – einem der letzten Vertreter der Meistersänger georgischer Musik – von

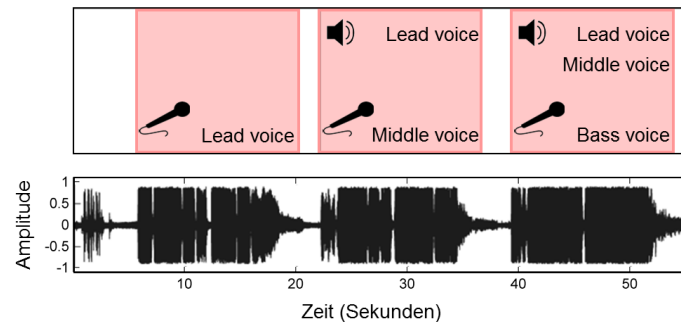


Abb. 1: Illustration des dreistufigen Aufnahmeprozesses von dreistimmigen georgischen Gesängen aufgeführt von Artem Erkomaishvili.

großer Bedeutung. Im Jahr 1966 wurden am „Tbilisi State Conservatory“ Tonaufnahmen angefertigt, bei denen der alternde Erkomaishvili dreistimmige Gesänge aufführte, indem er nacheinander die einzelnen Stimmen sang (siehe Abb. 1). Nach der Aufnahme der ersten Stimme (im Folgenden als „Lead-Stimme“ bezeichnet) wurde ein Tonbandgerät verwendet, um diese Stimme wiederzugeben, während ein zweites Tonbandgerät verwendet wurde, um die zweite Stimme („mittlere Stimme“) synchron aufzuzeichnen. Ähnlich spielte man dann die erste und zweite Stimme ab, um dazu synchron die dritte Stimme („Bassstimme“) aufzuzeichnen. Die daraus resultierende Sammlung von ca. 100 Audioaufnahmen stellt eine wichtige Quelle für das ursprüngliche georgische musikalische Denken dar und ist für die musikethnologische Forschung von großer Bedeutung [Sh14].

3 Datenaufbereitung und Annotation

Um die Aufnahmen für die musikethnologische Forschung zugänglich zu machen, müssen diese erst einmal in geeignete Darstellungsformen (z. B. durch Transkription in eine Notendarstellung oder durch Extraktion von geeigneten Merkmalen) gebracht werden. Dieser Schritt wird in traditioneller Musikforschung oft manuell durchgeführt und kann extrem aufwändig sein. Am Beispiel des Erkomaishvili-Datensatzes wurde in [Mü17] beschrieben, wie hier automatisierte Methoden sehr gewinnbringend zum Einsatz kommen können. Zunächst wurden für jede der Aufnahmen die Segmente des dreistufigen Aufnahmeprozesses manuell bestimmt. Unter Verwendung einer interaktiven grafischen Benutzeroberfläche wurden dann Grundfrequenzverläufe (im Folgenden auch als F0-Trajektorien bezeichnet, siehe [Mü15]) der drei Stimmen bestimmt. Hierbei wurde die F0-Trajektorie der Lead-Stimme innerhalb des ersten Segments, die F0-Trajektorie der mittleren Stimme innerhalb des zweiten Segments und die F0-Trajektorie der Bassstimme im dritten Segment bestimmt. Die Segment- sowie die F0-Annotationen wurden in einem einfachen CSV-Format abgespeichert und auf einer Webseite³ öffentlich zugänglich gemacht. Für eine ausführliche Beschreibung der

³ <https://www.audiolabs-erlangen.de/resources/MIR/2017-GeorgianMusic-Erkomaishvili>

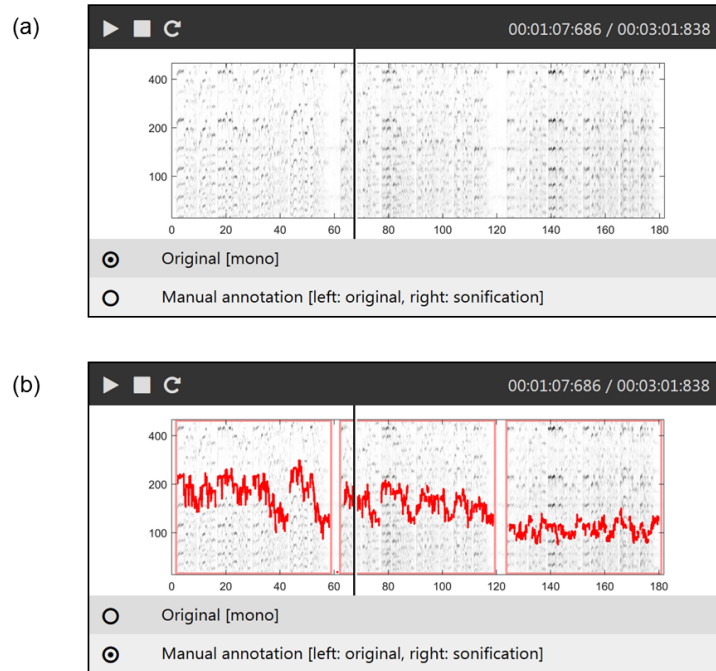


Abb. 2: Webbasierte Schnittstelle zur akustischen Wiedergabe von Audioaufnahmen und zeitsynchronen Anzeige der jeweiligen Zeitposition in Visualisierungen von Merkmalsdarstellungen und Annotationen. **(a)** Spektrogramm der Tonaufnahme „Die Engel im Himmel“ (ID 002 im Erkomaishvili-Datensatz). **(b)** Überlagerung des Spektrogramms mit Segment- und F0-Annotationen.

Benutzeroberfläche, der zugrundeliegenden Techniken und der resultierenden Annotationen verweisen wir auf [Mü17].

Neben der Verarbeitung der Audiodaten und der Erzeugung von Annotationen können computer-basierte Methoden auch für den Zugang der unterschiedlichen Datenarten enorme Hilfestellungen leisten. Exemplarisch zeigen wir nun, wie webbasierte Technologien eingesetzt werden können, um einen direkten und intuitiven Zugang zu den Tonaufnahmen und den erzeugten Annotationen zu ermöglichen. Als Beispiel zeigt Abb. 2a eine Zeit-Frequenzdarstellung (Spektrogramm, siehe [Mü15]) der Tonaufnahme „Die Engel im Himmel“ (ID 002) aus dem Erkomaishvili-Datensatz. Beim Abspielen der Aufnahme wird die entsprechende zeitliche Position im Spektrogramm durch eine mitlaufende vertikale Linie angegeben. Beim Umschalten auf die zweite Audiospur (siehe Abb. 2b) wird das Spektrogramm mit den erzeugten Segment- und F0-Annotationen überlagert. Darüber hinaus enthält der linke Kanal in dieser Spur die ursprüngliche Audioaufnahme, während der rechte Kanal eine Sonifikation der F0-Trajektorien bereitstellt. Hierdurch können die Annotationen leicht nachvollzogen und überprüft werden. Die Bereitstellung von

Benutzerschnittstellen, die sich gerade im Bereich der Webtechnologien mittels geeigneter Module häufig ohne tiefere Programmierkenntnisse realisieren lassen, können für Forscher einen großen Mehrwert bringen, um mit dem Audiomaterial und den Annotationen zu interagieren.

4 Rechnergestützte Analysen der Tonaufnahmen

Um das Potenzial der erzeugten Annotationen im Zusammenhang mit einer rechnergestützten Analyse der Tonaufnahmen anzudeuten, gehen wir nachfolgend kurz auf die Frage ein, welche Randbedingungen sich aus den Eigenschaften der erzeugten F0-Trajektorien hinsichtlich der „tonalen Organisation“ ([Ni15]) in den Tonaufnahmen und damit über Artem Erkomaishvilis Vorstellung von der authentischen georgischen „Stimmung“ ergeben. Diskussionen über Fragen der authentischen Aufführungspraxis sind natürlich kein rein georgisches Phänomen, werden aber im Zusammenhang mit georgischer Musik seit langem mit grosser Intensität und ohne, dass sich ein Konsens abzeichnet, geführt (siehe u. a. [Er02, Ge02, We02, Go04, Ka10, TV14, Er16]). Aufgrund des Fehlens schriftlicher musiktheoretischer Zeugnisse kommt der Analyse wie dem hier vorliegenden Datensatz deshalb eine große Bedeutung zu. Im Zusammenhang mit der Betrachtung der Stimmung vor allem von Vokalmusik ist es sinnvoll, zwischen dem melodischen und dem harmonischen Intervallinventar zu unterscheiden [Ni15]. Wir wenden uns zunächst dem melodischen (Abschnitt 4.1) und anschließend dem harmonischen Intervallgehalt (Abschnitt 4.2) der Tonaufnahmen zu.

4.1 Melodischer Intervallgehalt

Bei den Tonaufnahmen des Erkomaishvili-Datensatzes, wie in Abschnitt 2 beschrieben, wurde die Lead-Stimme zunächst isoliert gesungen (siehe auch Abb. 1). Unter der Annahme, dass in diesen Passagen die Tonhöhenverteilung dem *mentalen Skalenmodell* des Sängers am nächsten kommt, wurden daher aus der F0-Trajektorie der Lead-Stimme für jede dieser Aufnahmen ein Histogramm berechnet. Hieraus wurden durch Anwendung einer geeigneten Glättung, Interpolation und Normalisierung eine Wahrscheinlichkeitsdichte („probability density function“ oder PDF) abgeleitet. Abb. 3a zeigt ein Histogramm und eine abgeleitete PDF-Kurve für unser vorheriges Beispiel ID 002. Basierend auf jeder PDF-Kurve wurde dann eine Clusteranalyse mit Hilfe des *k*-Means-Algorithmus durchgeführt. Hierbei wurden die lokalen Maxima einer PDF-Kurve als Startwerte verwendet und $k = 9$ für die die Anzahl der Cluster gewählt. Abb. 3b zeigt das vorherige Histogramm des Beispiels ID 002, das nun entsprechend der Cluster-Zugehörigkeiten eingefärbt wurde. Diese Visualisierung zeigt, dass sich die F0-Werte für die Lead-Stimme in einer solchen Weise gruppieren, dass der Tonumfang einer Oktave (z. B. von 1300 bis 2500 cent) in sieben Intervalle eingeteilt wird. Dies ist konsistent mit der Annahme, dass es sich bei den Clusterzentren näherungsweise um die Tonhöhen des *mentalen Skalenmodells* des Sängers handelt.

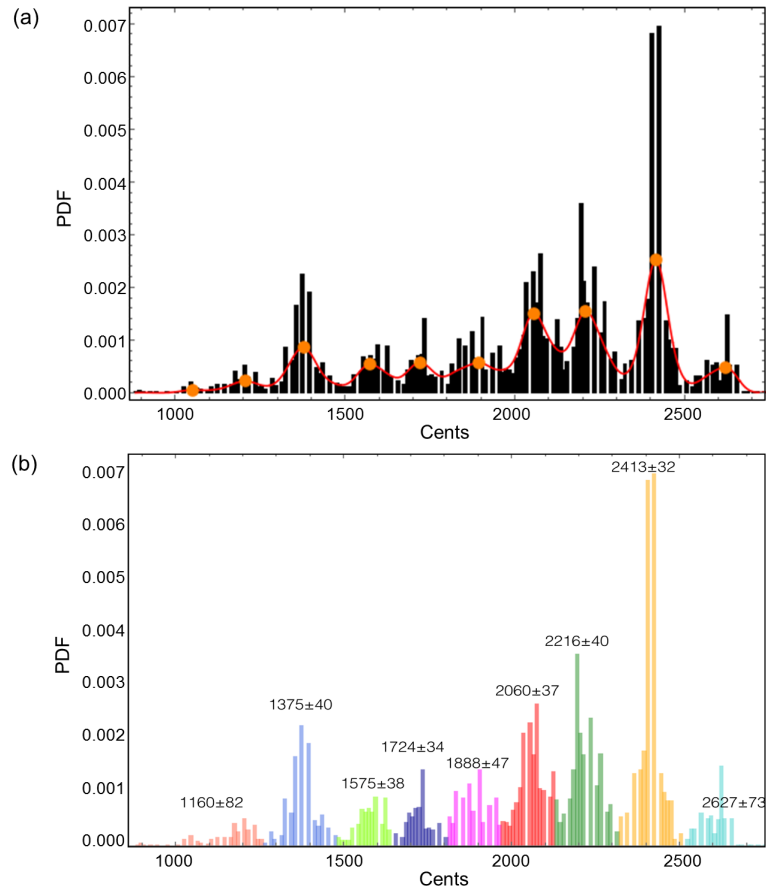


Abb. 3: **(a)** Histogramm der F₀-Verteilung für die Lead-Stimme des Beispiels ID 002. Die rote durchgezogene Linie entspricht der abgeleiteten PDF-Kurve, deren Maxima durch die Kreise markiert sind. Die Referenzfrequenz für die Cent-Berechnung ist 55 Hz (d. h., 0 Cent entsprechen 55 Hz). **(b)** Eingefärbtes Histogramm entsprechend der Cluster-Zugehörigkeiten nach Anwendung eines *k*-Means-Algorithmus mit *k* = 9. Die numerischen Werte der Clusterzentrus sind zusammen mit den jeweiligen Standardabweichungen oberhalb jedes Clusters angegeben.

Zur Erklärung der durch die Cluster-Analyse berechneten melodischen Intervallverteilungen (mentalen Skalenmodelle) wurden diese in einem weiteren Schritt mit unterschiedlichen bekannten Skalenmodellen (im Folgenden als *Testmodelle* bezeichnet) verglichen. Hierzu wurden 15 unterschiedliche Testmodelle herangezogen, von denen jeweils eine Oktave in Abb. 4 dargestellt ist. Dabei handelt es sich bei den Modellen JM und SM um das sogenannte „Joined Mode“ und „Split Mode“ Modell von [Er02], bei EQ um das gleichabständige Skalenmodell von [TV14], bei PY um die Pythagoräische und bei PT um die Ptolemäische

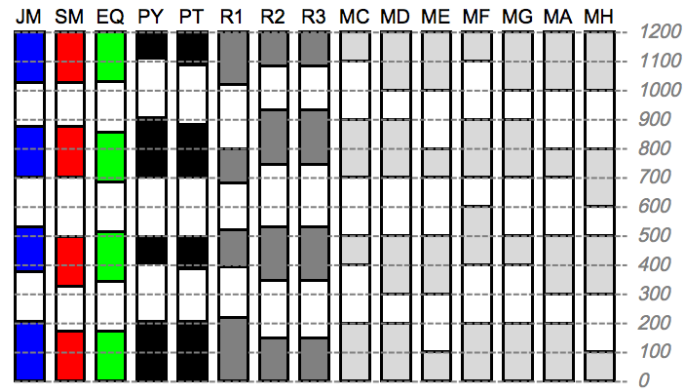


Abb. 4: Intervallverteilungen für die Skalenmodelle, die nachfolgend als Testmodelle zur Erklärung der beobachteten melodischen Intervallverteilungen im Erkomaishvili-Datensatz verwendet wurden. Die Zahlenwerte am rechten Rand sind in Cent angegeben.

Skala. Bei R1, R2 und R3 handelt es sich um zufallsgenerierte Skalen und bei MC bis MH um die Kirchentonarten auf den Tonika-Tönen C bis H.

Für jede Tonaufnahme des Erkomaishvili-Datensatzes wurde das berechnete mentale Skalenmodell mit jedem der 15 unterschiedlichen Testmodelle gekreuzkorreliert. Die Testmodelle mit dem jeweils größten Korrelationskoeffizienten wurden als *Gewinnermodelle* angesehen. Für diese wurden anschließend sogenannte *Modellresiduenvektoren* wie folgt berechnet. Zunächst wurde die als mentale Skalenmodell angenommene F0-Verteilung so gegenüber dem Testmodell verschoben, dass die bei dieser Verschiebung berechnete Kreuzkorrelation dem Maximalwert der Kreuzkorrelation beider Skalen entspricht. Anschließend wurden die Distanzen der jeweiligen Intervallgrenzen der Testskala zu ihren nächsten Nachbarn auf der als mentale Skalenmodell angenommenen F0-Verteilung berechnet. Die Wertemenge der sich ergebenden Modellresiduen gibt Aufschluss darüber, wie gut (in einem absoluten Sinn) das Gewinnermodell die zu modellierende F0-Verteilung erklärt.

In diese Betrachtung wurden nur die Tonaufnahmen einbezogen, bei denen die F0-Cluster der obersten Stimme mindestens eine Oktave überstreichen. In Abb. 5 sind die Ergebnisse dieser Betrachtung in Form von Box-Whisker-Diagrammen für 46 der verbleibenden Tonaufnahmen des Erkomaishvili-Datensatz dargestellt. Eine detaillierte Analyse von Abb. 5 zeigt, dass bezüglich der Korrelation die Modelle JM und SM von [Er02] mit insgesamt 29 Gewinnermodellen sowie das Modell EQ von [TV14] mit insgesamt 12 Gewinnermodellen am erfolgreichsten sind. In einem Fall korreliert das Ptolemäische Skalenmodell (PT) und in vier Fällen eines der Zufallsmodelle (R1, R2, R3) am besten. In keinem der Fälle gehört eine der Kirchentonarten zu den Gewinnermodellen. Die Mittel- und Medianwerte der Residuenvektoren für die Gewinnermodelle liegen für einige der Tonaufnahmen in einem Bereich von etwa 10 Cent, was nahe der wahrnehmbaren Schwelle

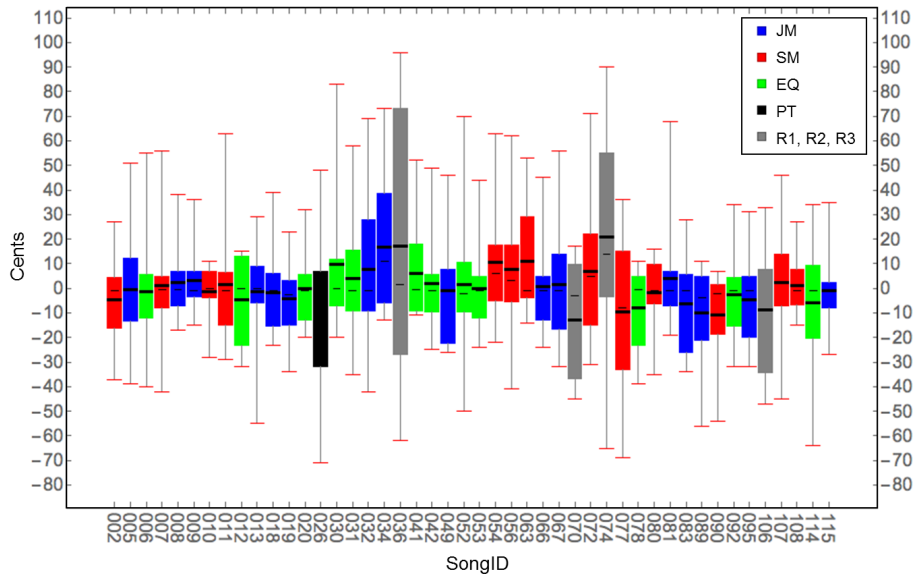


Abb. 5: Box-Whisker-Darstellung der Modellresiduenvektoren für die Gewinnermodelle (kodiert wie in Abb. 4). Die oberen und unteren Querbalken (whisker) der Box-Whisker-Diagramme stellen die maximalen Modellresiduen dar. Die oberen und unteren Begrenzungen der Rechtecke (box) entsprechenden 75%- und 25%- Quantilen. Die dicken und dünnen horizontalen Linien geben die jeweiligen Mittel- und Medianwerte des Residuenvektors an.

ist. Dies gilt allerdings nicht für die maximalen Abweichungen. Diese sind in allen Fällen sehr viel größer, sodass man nicht davon ausgehen kann, dass Artem Erkomashvili sie nicht wahrgenommen und korrigiert hätte, wenn er dies gewollt hätte.

Für die großen Maximalwerte der Modellresiduenvektoren sind mehrere Gründe denkbar. Einzelne F0-Cluster, vor allem solche mit einer geringen Anzahl von F0-Samples, wie z. B. das Cluster in Abb. 3b mit dem Zentrum bei 1160 cent können eine so große Standardabweichung haben, dass die Annahme, dass es sich dabei um eine Tonhöhe des mentalen Skalenmodells des Sängers handelt, nicht mehr gerechtfertigt ist. Hier könnte eine detailliertere Betrachtung der einzelnen Audioaufnahmen mit individuell angepassten Parametern bei der Clusteranalyse, die hier allerdings nicht angestrebt war, möglicherweise zu einer Verbesserung der Modellanpassung führen. Zu bedenken ist allerdings auch, wie man aus der in [SMR17] diskutierte Analyse der tatsächlich gesungenen melodischen Intervalle (im Unterschied zu den hier betrachteten F0-Clustern) schließen kann, dass eine melodische Präzision gar nicht angestrebt wurde, sondern dass vielmehr, und darauf deuten die harmonischen Intervallverteilungen in [SMR17] hin, im Zweifelsfall der harmonischen Präzision der Vorzug vor der melodischen gegeben wurde. Ein weiterer Grund für große Modellresiduen können Wechsel der „Tonart“ (oder genauer gesagt des „Modus“) innerhalb

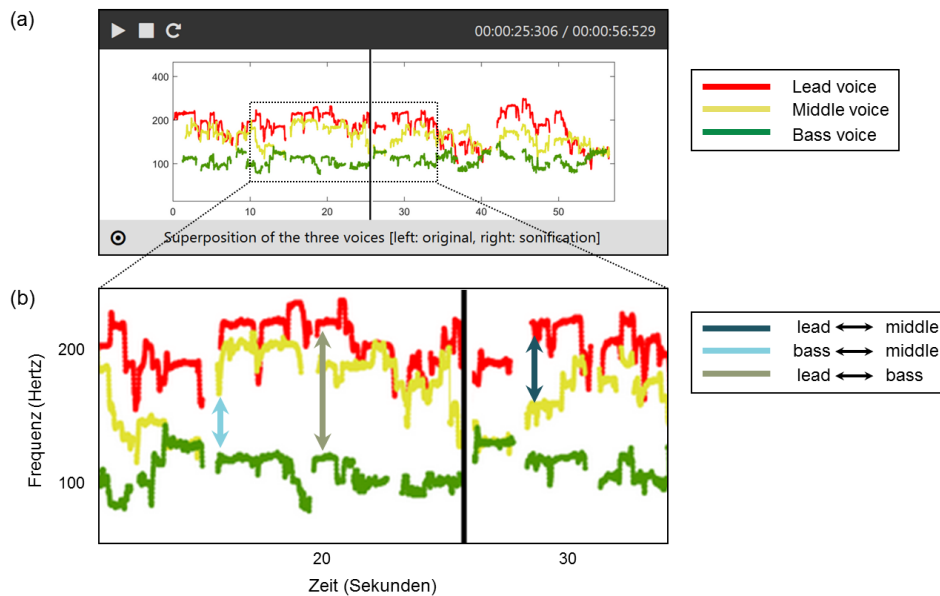


Abb. 6: Webbasierte Schnittstelle zur akustischen Wiedergabe der Überlagerung aller drei Stimmen und ihrer F0-Trajektorien. (a) Überlagerte Darstellung der F0-Trajektorien der Lead-, Mittel-, und Bassstimme für das Beispiel ID 002 (Fortsetzung von Abb. 2). (b) Vergrößerter Bereich (zoom) mit einer Illustration der Berechnung der Intervalle zwischen den unterschiedlichen Stimmen.

einer Aufnahme sein. Auch hier wäre zur besseren Analyse eine detailliertere Betrachtung der einzelnen Audioaufnahmen notwendig.

4.2 Harmonischer Intervallgehalt

Abschließend soll noch kurz auf den *harmonischen Intervallgehalt* der Tonaufnahmen eingegangen werden. Im Gegensatz zu Instrumentalmusik von Instrumenten mit festgelegtem Tonvorrat ist für mehrstimmige Vokalmusik nicht notwendigerweise damit zu rechnen, dass die verwendeten harmonischen Intervalle mit den melodischen übereinstimmen. Dies zeigt sich beispielsweise sehr deutlich in der Analyse von Feldaufnahmen von Dorfsängern aus den Bergregionen Nordwest-Georgiens (Swanetien), welche mit Kehlkopfmikrofonen aufgenommen wurden [Sc16].

Auch für die hier untersuchten Tonaufnahmen des Erkomaishvili-Datensatzes sind die F0-Trajektorien für die Einzelstimmen verfügbar (siehe Abschnitt 3). Neben der Analyse des melodischen Intervallgehalts (horizontale Sichtweise) gehen wir nun kurz auf die Analyse des harmonischen Intervallgehalts (vertikale Sichtweise) einzelner Aufnahmen sowie des gesamten Datensatzes ein, siehe auch [Mü17]. Zu diesem Zweck wurden die geschätzten

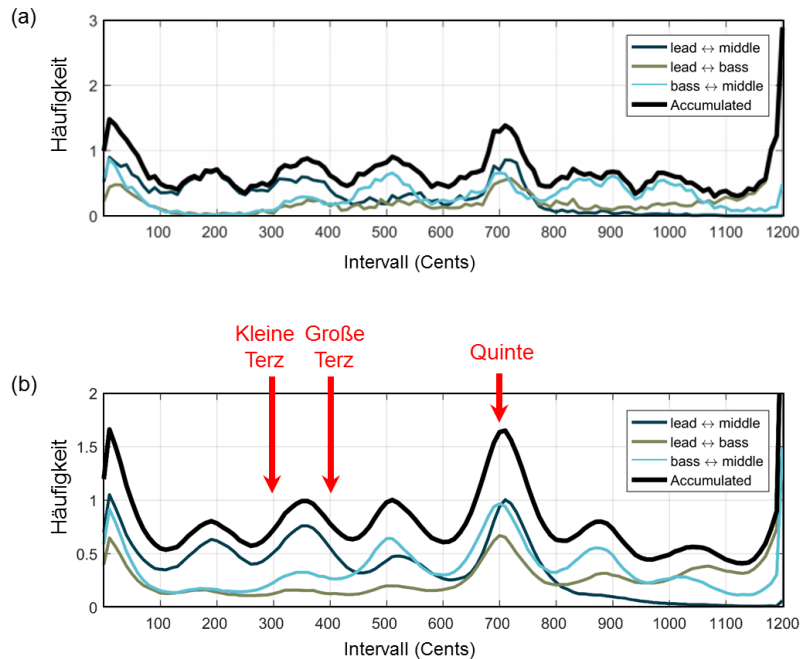


Abb. 7: Intervallverteilungen für drei Fälle (Lead- und Mittelstimme, Lead- und Bassstimme, Mittel- und Bassstimme) sowie die Gesamtverteilung über die drei Fälle („Accumulated“). **(a)** Intervallverteilung für die Aufnahme „Die Engel im Himmel“ (Fortsetzung des Beispiels ID 002 von Abb. 6). **(b)** Intervallverteilung gemittelt über alle der Aufnahmen des Erkomaishvili-Datensatzes.

F0-Trajektorien der Lead-, Mittel- und Bassstimme zeitlich überlagert (siehe Abb. 6a). Dann wurden für jede Zeitposition die Intervalle (gegeben in Cent⁴) zwischen den F0-Trajektorien der Lead- und Mittelstimme, der Lead- und Bassstimme sowie der Mittel- und Bassstimme berechnet (siehe Abb. 6b). Schließlich wurde für jeden der drei Fälle eine Intervallverteilung berechnet, in dem die auftretenden Intervalle über die Zeit integriert wurden (siehe Abb. 7a). Diese Intervallverteilungen wurden für alle Aufnahmen des Erkomaishvili-Datensatzes bestimmt.⁵ Eine Mittelung dieser Verteilungen über alle Aufnahmen hinweg ist in Abb. 7b zu sehen. Diese Abbildung zeigt, dass die Quinte (abgesehen von dem Unisono) das häufigst vorkommende harmonische Intervall im gesamten Datensatz ist. Weiterhin kann man feststellen, dass die Quinte mit etwas über 700 Cent auch sehr „rein“ gesungen wurde. Eine interessante Beobachtung ist, dass Erkomaishvili nicht zwischen kleiner Terz (300 Cent, Moll-Terz) und großer Terz (400 Cent, Dur-Terz) – Intervalle der 12-tönigen gleichstufigen Stimmung – unterscheidet. Stattdessen wird die Terz quasi „neutral“ (also zwischen der

⁴ Cent ist eine Maßeinheit für den Vergleich musikalischer Intervalle, siehe [Mü15]. Hierbei wird ein Halbtonintervall in 100 Cent unterteilt. Eine Oktave entspricht also 1200 Cent.

⁵ Die Ergebnisse sind auf der Webseite <https://www.audiolabs-erlangen.de/resources/MIR/2017-GeorgianMusic-Erkomaishvili> verfügbar.

Moll- und Dur-Terz) gesungen – eine Praxis, die sich in gleicher Weise auch heute noch in den Gesängen swanetischer⁶ Sänger wiederfinden lässt [Sc16].

5 Diskussion und Ausblick

Rechnergestützte Musikethnologie (*Computational Ethnomusicology* [GHGM13, SB13, Tz07]), ist ein vergleichsweise junges Forschungsgebiet, welches sich salopp gesprochen dadurch beschreiben lässt, dass rechnergestützte Methoden zur Behandlung musikethnologischer Fragestellungen eingesetzt werden. Nach [Tz14] geschieht dies zur Zeit allerdings noch überwiegend explorativ. Mit zunehmender Entwicklung des Gebietes und bei zunehmend stärkerer Interaktion zwischen Musik- und Computerwissenschaftlern könnten rechnergestützte Methoden jedoch vermehrt auch zum systematischen, datengetriebenen Testen musikwissenschaftlicher Hypothesen und damit zur Lösung musikwissenschaftlicher Fragen mittels eines naturwissenschaftlich verankerten Methodeninventars eingesetzt werden [Tz14]. Dies wurde im vorliegenden Fall am Beispiel des Vergleichs der in Abb. 4 dargestellten Skalenmodelle mit dem melodischen Intervallgehalt des Erkomaishvili-Datensatzes angedeutet. Auch wenn die Frage nach der authentischen musikalischen Aufführungspraxis traditioneller georgischer Vokalmusik nicht anhand der hier exemplarisch vorgestellten Analysen geklärt werden kann, so ist doch bereits aus den Betrachtungen dieses Beitrags deutlich geworden, dass beispielsweise die europäischen Kirchentöne für die Beschreibung des Erkomaishvili-Datensatzes völlig ungeeignet sind. Sicher ließe sich auch die Analyse der hier als mentale Skalenmodelle angenommenen Tonhöhencluster noch verbessern, indem z. B. nur solche Clusterverteilungen betrachtet werden, bei denen bestimmte definierte Qualitätskriterien (wie z. B. beschränkte Standardabweichungen der Cluster) erfüllt sind.

Der Vergleich des *melodischen* und *harmonischen Intervallgehalts* im Erkomaishvili-Datensatz stellt ein weiteres Beispiel für das Potential rechnergestützter Verfahren in der Musikethnologie dar. Einerseits wären Betrachtungen wie die des über den gesamten Datensatz gemittelten harmonischen Intervallgehalts (Abb. 7b) ohne Rechnerunterstützung kaum möglich; andererseits verdeutlicht der Vergleich, dass das Konzept einer einzigen „eindimensionalen“ Einteilung des Tonvorrats in wie auch immer geartete Tonklassen der Komplexität der georgischen mehrstimmigen Vokalmusik, und möglicherweise auch anderen Vokalmusiktraditionen, nicht gerecht wird.

Mit den hier diskutierten Beispielen kann das Potential rechnergestützter Verfahren und Methoden der automatisierten Musikverarbeitung [Mü15] für den Einsatz in der Musikethnologie [Ri14] natürlich nur ansatzweise illustriert werden. Der Erfolg einer „rechnergestützten Musikethnologie“ wird sicherlich sehr stark davon abhängen, inwieweit Musikethnologen (welche letztendlich die Fragestellungen liefern) und Computerwissenschaftler (welche über ein reichhaltiges Methodeninventar der Audioverarbeitung, der Algorithmik und des maschinellen Lernens verfügen) bereit sind, in einen interdisziplinären Dialog zu treten.

⁶ Swanetien ist eine Bergregion Georgiens im Großen Kaukasus.

Danksagungen

Die International Audio Laboratories Erlangen sind eine Gemeinschaftseinrichtung der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) und des Fraunhofer-Institut für Integrierte Schaltungen IIS. Wir danken Stefanie Kämmerle, Vlora Arifi-Müller und Stefan Balke für die Hilfe bei der Erstellung der Annotationen und der Webseiten.

Literaturverzeichnis

- [Er02] Erkvanidze, Malkhaz: On georgian scale system. In: The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia. S. 178–185, 2002.
- [Er16] Erkvanidze, Malkhaz: The Georgian Musical System. In: 6th International Workshop on Folk Music Analysis, Dublin, 15-17 June, 2016. S. 74–79, 2016.
- [Ge02] Gelzer, Stuart: Testing a scale theory for Georgian folk music. In: The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia. S. 194–200, 2002.
- [GHGM13] Gómez, Emilia; Herrera, Perfecto; Gómez-Martin, Francisco: Computational Ethnomusicology: perspectives and challenges. *Journal of New Music Research*, 42(2):111–112, 2013.
- [Go04] Gogotishvili, Vladimir: On authentic and plagal types of monotonic (non-octave) scales in Georgian traditional vocal polyphony. In: Proc. of the Second International Symposium on Traditional Polyphony, 23-27 September, 2004, Tbilisi, Georgia. S. 218–226, 2004.
- [Ka10] Kawai, Norie; Morimoto, Masako; Honda, Manabu; Onodera, Eiko; Oohashi, Tsutomu: Study on sound structure of Georgian traditional polyphony. Analysis of its temperament structure. In: The Fifth International Symposium on Traditional Polyphony, 4-8 October, 2010, Tbilisi, Georgia. Jgg. 1, S. 532–537, 2010.
- [Mü15] Müller, Meinard: *Fundamentals of Music Processing*. Springer Verlag, 2015.
- [Mü17] Müller, Meinard; Rosenzweig, Sebastian; Driedger, Jonathan; Scherbaum, Frank: Interactive Fundamental Frequency Estimation with Applications to Ethnomusicological Research. In: Submitted to the AES Conference on Semantic Audio. Erlangen, Germany, 2017.
- [Ni15] Nikolsky, Aleksey: Evolution of tonal organization in music mirrors symbolic representation of perceptual reality. Part-1: Prehistoric. *Frontiers in Psychology*, 6(OCT):1–36, 2015.
- [Ri14] Rice, Timothy: *Ethnomusicology: A very short introduction*. Oxford University Press, 2014.
- [SB13] Savage, Patrick E.; Brown, Steven: Toward a New Comparative Musicology. *Analytical Approaches To World Music*, 2(2):148–197, 2013.
- [Sc16] Scherbaum, Frank: On the benefit of larynx-microphone field recordings for the documentation and analysis of polyphonic vocal music. Proc. of the 6th International Workshop Folk Music Analysis, 15 - 17 June, Dublin/Ireland, S. 80–87, 2016.

-
- [Sh14] Shugliashvili, David: Georgian Church Hymns. Georgian Chanting Foundation, Tbilisi State Conservatoire, 2014.
- [SMR17] Scherbaum, Frank; Müller, Meinard; Rosenzweig, Sebastian: Analysis of the Tbilisi State Conservatory Recordings of Artem Erkomaishvili in 1966. In: International Workshop on Folk Music Analysis (FMA). Málaga, Spain, 2017.
- [TV14] Tsereteli, Zaal; Veshapidze, Levan: On the Georgian traditional scale. In: The Seventh International Symposium on Traditional Polyphony: 22-26 September, 2014, Tbilisi, Georgia. S. 288–295, 2014.
- [Tz07] Tzanetakis, George; Kapur, Ajay; Schloss, W Andrew; Wright, Matthew: Computational Ethnomusicology. *Journal of Interdisciplinary Music Studies*, 1(2):1–24, 2007.
- [Tz14] Tzanetakis, George: Computational ethnomusicology: a music information retrieval perspective. In: Music Technology meets Philosophy - From Digital Echos to Virtual Ethos: Joint Proceedings of the 40th International Computer Music Conference (ICMC) and the 11th Sound and Music Computing Conference (SMC). Michigan Publishing, Athens, Greece, 2014.
- [We02] Westman, Johan: On the problem of the tonality in Georgian polyphonic songs: The variability of pitch, intervals and timbre. In: The First International Symposium on Traditional Polyphony: 2-8 September, 2002, Tbilisi, Georgia. S. 212–220, 2002.
- [Zi09] Ziegler, Susanne: Georgien. In: *Musik in Geschichte und Gegenwart*, S. 1271–1278. 2009.