

# Tonhöheninformierten Trennung in Solo- und Begleitspuren<sup>1</sup>

Estefanía Cano<sup>2</sup>

**Abstract:** Diese Veröffentlichung im Forschungsgebiet der Klangquellentrennung befasst sich speziell mit der Separierung von einzelnen Musikinstrumenten aus bereits gemischten Audio-Signalen. Eine Methode zur tonhöheninformierten Trennung in Solo- und Begleitspuren wird detailliert beschrieben. Die praktische Anwendbarkeit der vorgestellten Methode wird anhand der Integration in die Musiklernsoftware Songs2See erläutert. Des Weiteren werden die Ergebnisse einer Studie im Hinblick auf Anwendbarkeit für die harmonisch/perkussive Klangquellentrennung präsentiert, welche die Klänge verschiedener musikalischer Instrumente untersucht.

## 1 Einleitung

Wir sind den ganzen Tag unterschiedlichsten Geräuschen ausgesetzt. Verkehrslärm, Stimmen, Blätter, die sich im Wind wiegen. Alles wirkt gleichzeitig auf uns ein. Auch Musik besteht aus dem Zusammenspiel verschiedener Klangquellen. Was wir als Musikliebhaber heute auf CD oder als digitalen Download kaufen, ist ein fertiger Mix aus verschiedenen Instrumenten, Stimmen und Geräuschen, die alle gemeinsam erklingen und ihren Teil zum akustischen Gesamteindruck beitragen. Dadurch entstehen jedoch auch Nachteile für den Hörer: Nachträgliche Korrekturen der Audioaufnahme, wie beispielsweise die Reduzierung der Lautstärke einzelner Instrumente oder den Klang von alten Mono-Aufnahmen an heutige Stereo- oder Mehrkanalstandards anzupassen, wird durch die Verschmelzung der Klangquellen während der Aufnahme meist unmöglich.

Doch nicht nur für die nachträgliche Bearbeitung von fertigen Musiktiteln benötigt man idealerweise die einzelnen Instrumentenspuren. Auch für das Üben von Instrumenten können die getrennten Spuren den Lernprozess deutlich erleichtern. Viele Feinheiten werden erst wahrnehmbar, wenn die Melodiestimme isoliert gehört werden kann. Hat man die Melodie mit seinem eigenen Instrument einstudiert, gilt es als nächsten Schritt, das Erlernete im Band-Kontext anzuwenden. Hierfür eignen sich besonders Aufnahmen, in welchen das Melodieinstrument komplett ausgeblendet wurde. Die Erzeugung dieser Audiodaten ist bei bereits aufgenommenen und abgemischten Musikstücken jedoch nicht mehr möglich.

Die Klangquellentrennung, ein relativ neues Forschungsgebiet der Informatik, versucht hier Abhilfe schaffen und das fertige Musikstück wieder in seine einzelnen Bestandteile zu zerlegen.

---

<sup>1</sup> Englischer Titel der Dissertation: Pitch-informed Solo and Accompaniment Separation

<sup>2</sup> Technische Universität Ilmenau, cano@idmt.fraunhofer.de

## 2 Dissertationsschwerpunkte

Im Rahmen meiner Dissertation am Fraunhofer-Institut für Digitale Medientechnologie IDMT in Ilmenau galt es, dem Musikliebhaber genau diese Klangquellentrennung zu ermöglichen. Im Forschungsprojekt *Songs2See* entwickelten meine Kollegen und ich eine Musiklernsoftware, welche verschiedene Instrumente unterstützt und es dem Nutzer ermöglicht, jeden beliebigen Musiktitel als Übungsinhalt aufzubereiten. Dabei wurde auf bestehende Technologien des Fraunhofer IDMT aus dem Bereich des Music Information Retrieval (MIR) zur Informationsgewinnung aus dem Musiksignal zurückgegriffen. Diese beinhalten beispielsweise eine automatische Tonerkennung. Jedoch fehlte es an Technologien für die Aufteilung beliebiger Musikstücke in Begleit- und Melodieinstrumentenspuren, welchen sich die vorliegende Arbeit vorrangig widmet.

Im zweiten Teil der Arbeit wurde ein Experiment zur besseren Modellierung der Zusammenhänge zwischen Magnitude, Phase und Feinfrequenz von isolierten Instrumententönen durchgeführt. Diese Studie sollte das Verständnis des klanglichen Verhaltens fördern und dadurch zu einer Verbesserung der Klangquellentrennung beitragen. Während der Arbeiten an der Studie ergaben sich neue Erkenntnisse hinsichtlich der Phaseninformationen. Die Ergebnisse versprachen großen Nutzen zur Unterscheidung von harmonischen Instrumenten, beispielsweise Saxophon, und perkussiven Instrumenten, beispielsweise Schlagzeug, so dass zusätzlich ein System für die harmonisch-perkussive Trennung entwickelt und getestet wurde. Diese Art der Quellentrennung wird unter anderem bei der Wiederaufbereitung von Mono-Aufnahmen in Stereo-Signale verwendet [Fi11].

## 3 Stand der Technik

Die Klangquellentrennung hat sich erst in den letzten Jahren als Forschungsgebiet etabliert. Erste Varianten nutzten Equalizer Algorithmen. Diese dämpfen gezielt Frequenzbereiche, in denen das Instrument vermutet wird. Im Allgemeinen ist der Klang eines einzelnen Instrumentes jedoch nicht auf eine konkrete Frequenz, beispielsweise 400 Hz, zu begrenzen. Vielmehr setzt er sich dieser aus der Grundfrequenz sowie mehreren Obertönen zusammen, die einer Vielfachen der Grundfrequenz entsprechen (400 Hz, 800 Hz, 1200 Hz, etc.) [Br96]. Dies verdeutlicht die Nachteile, welche durch das Dämpfen einzelner Frequenzbereiche entstehen: Nicht ein einzelnes Instrument wird in der Lautstärke reduziert, sondern nur der Klang der Gesamtkomposition verändert.

Andere Ansätze verwenden Informationen, welche aus dem Stereoklangbild erlangt werden können: Oftmals klingt das Solo-Instrument auf beiden Kanälen gleich stark, während andere Instrumente nur links oder rechts zu hören sind. Durch Differenzbildung beider Kanäle bleibt nur das Gemeinsame - das Melodieinstrument - übrig. Von dieser Grundannahme kann jedoch nicht immer ausgegangen werden, da sich die tragende Stimme je nach Mix auch beispielsweise nur links befinden kann. Für ältere Aufnahmen aus der ersten Hälfte des 20. Jahrhunderts liegt zudem meist nur eine Monofassung vor, so dass dieser Ansatz für den Anwendungsfall auf beliebige Musiktitel nicht in Betracht kam und nur Verfahren für Einkanal-Aufnahmen berücksichtigt wurden. Neuere Lösungsversuche

auf diesem Forschungsgebiet nutzen beispielsweise Nicht-Negative Matrix-Faktorisierung [Le01], Tensor-Faktorisierung [FCC05] sowie verschiedene Filter [Li13]. Obwohl die publizierte Ergebnisse auf das Potential dieser Ansätze hindeuten, bringen sie zwei entscheidende Nachteile für den Einsatz in Songs2See mit sich: Sie benötigen umfangreiches Zusatzwissen über die zu trennende Quellen und sie verursachen aktuell einen enormen Rechenaufwand, welcher dem Endanwender einer Musiklernsoftware nicht zumutbar ist.

## 4 Eigener Ansatz

Im Rahmen dieser Dissertation wurde die tonhöhen-informierte Quellentrennung (pitch-informed source separation) als Ansatz gewählt. Hierbei werden lediglich Information über die gespielten Töne der Hauptstimme benötigt, um diese zu isolieren. Damit eignet sich dieser Ansatz dazu, das dominante Instrument aus einem Musikstück zu extrahieren, unabhängig von der Art des Instrumentes, der Begleitung und der Stilrichtung. Für die vorliegende Arbeit werden nur einstimmige Melodieinstrumente in Betracht gezogen. Die Musikaufnahmen liegen monaural vor, es kann also keine zusätzliche Information aus der Verteilung der Instrumente im Stereo-Panorama gewonnen werden.

Das entwickelte System ist im folgenden Ablaufdiagramm dargestellt.

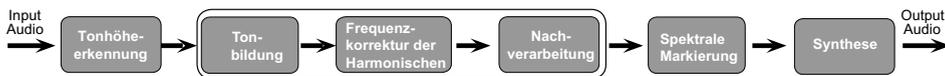


Abb. 1: Ablaufdiagramm des entwickelten Systems

### 4.1 Tonhöhererkennung - Pitch Detection

Bei der Pitch Detection wird zunächst die Tonhöhe des Melodieinstrumentes extrahiert. Tonhöhe bedeutet hierbei die als Grundton wahrgenommene Frequenz zu einem konkreten Zeitpunkt im Lied. Für den Ton C4 liegt die ideale Grundfrequenz bei 261,63 Hz. Für reale Instrumente ergeben sich leichte Abweichungen während des Ein- und Ausschwingvorganges des Tones. Eine kontinuierliche Folge dieser Frequenzen dient als Eingangsinformation für die folgenden Schritte, siehe Abbildung 2. Die Extraktion dieser Grundfrequenzen geschieht automatisch durch den Einsatz des Melodieerkennungsalgorithmus von K. Dressler [Dr11], welcher am Fraunhofer IDMT entwickelt wurde. Diese Erkennung wurde speziell für fertig gemischte Musik mit einer Vielzahl an gleichzeitig klingenden Instrumenten entwickelt. Die Qualität der Erkennung wurde in mehreren Evaluationskampagnen bestätigt [Sa13].

### 4.2 Tonbildung - Tone Formation

Im folgenden Schritt werden anhand dieser Grundfrequenzen Ton-Objekte erstellt, welche die Anfangszeit, die Dauer sowie die konkrete Tonhöhe beinhalten. Sie entstehen durch

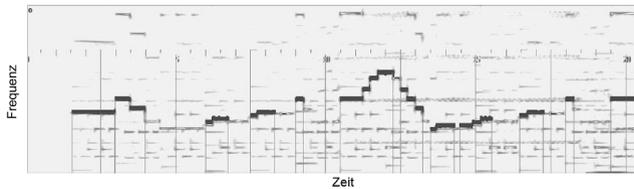


Abb. 2: Magnitudenspektrogramm eines Audiosignales und extrahierter Grundfrequenzverlauf in violett

die Zusammenfassung mehrerer aufeinanderfolgender ähnlicher Grundfrequenzen zu einzelnen Tönen. Die Ton-Objekte sind notwendig, um das Audio-Signal in musikalisch sinnvolle Bereiche zu unterteilen. Des Weiteren ermöglichen sie die Nutzung bekannter Charakteristiken von Tönen, beispielsweise die erwähnten Ein- und Ausschwingvorgänge, um das Ergebnis der Klangquellentrennung zu verbessern. Zusätzlich reduziert diese Zusammenfassung den Rechenaufwand, da weniger Daten verarbeitet werden müssen.

### 4.3 Frequenzkorrektur der Harmonischen - Harmonic Series Refinement

Das Ziel dieses Verarbeitungsschrittes besteht darin, eine Folge von Obertönen zu finden, welche das Solo-Instrument repräsentieren und zu den Parametern der erzeugten Ton-Objekte passen. Dafür werden folgenden Annahmen getroffen: (1) Jeder Oberton darf einen Viertelton von dem idealen Vielfachen der Grundfrequenz abweichen. (2) Die Obertonstruktur von Stimme, Blas- und Saiteninstrumenten sind bekannt.

Um eine Suche nach Obertönen zu ermöglichen, werden das Magnituden- und Phasen-Spektrogramm mittels Kurzzeit-Fourier-Transformation (STFT) mit einer Fenstergröße von 2048 Samples und einer Sprungweite von 256 Samples aus dem Audiosignal berechnet. Beide Spektrogrammdarstellungen werden später zur Manipulation und Resynthese ebenfalls genutzt. Für jeden Oberton wird eine iterative Suche im erwarteten Bereich des Magnitudenspektrums durchgeführt (siehe Abbildung 3). Das lokale Maximum mit der höchsten Amplitude im Suchbereich wird als harmonischer Oberton in Betracht gezogen. Als Ergebnis erhält man für jeden Frame (alle 256 Samples) der Ton-Objekte das harmonische Spektrum des Solo-Instrumentes.

### 4.4 Nachverarbeitung - Post-Processing

Folgende Nachverarbeitungsschritte verfeinern das harmonische Spektrum des Solo-Instrumentes und sind entscheidend für die klangliche Qualität der erzeugten Audio-Dateien:

(1) Zuerst werden die Notenanfänge verfeinert, da diese während der Pitch Detection oft verpasst werden. Viele Instrumente weisen ein breitbandiges Spektrum auf, bevor ein stabiler Ton erklingt. Entsprechend reagiert die Erkennung leicht verzögert. Für die Separierung ist die korrekte Zuordnung dieser Anschläge zu der Solospur jedoch wichtig, da dies stark den Klang des Instrumentes prägt. Zur Korrektur wird die harmonische Struktur

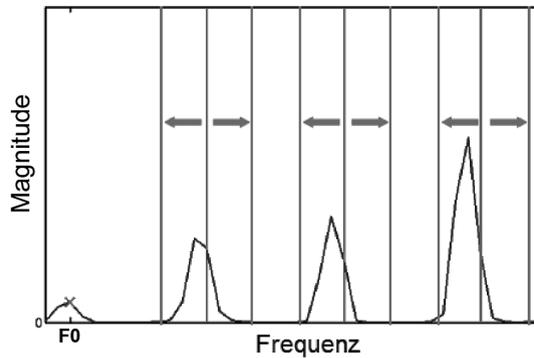


Abb. 3: Grundfrequenz (F0) und drei Obertöne, rot die ideale Position, grün der Suchbereich

des ersten Frames auf einen Bereich von 70 Millisekunden vor dem erkannten Anschlagszeitpunkt ausgedehnt. Dieser Wert wurde in Studien als durchschnittliche Dauer für den Einschwingvorgang verschiedener Instrumente gefunden [LC65].

(2) Anschließend wird versucht, die Anteile perkussiver Instrumente zu reduzieren, welche fälschlicherweise dem Solo-Instrument zugeordnet wurden. Diese Artefakte entstehen durch die Überlagerung der Klänge und sind im Spektralbereich als vertikale Linien sichtbar. In Anlehnung an [Fi10] wird ein Median-Filter über die Zeit-Achse des Spektrogramms angewendet, welcher diese Störungen beseitigt soll.

(3) Ein Merkmal musikalischer Klangquellen ist eine Folge von Obertönen, welche nicht unabhängig voneinander verlaufen, sondern zeitlich korrelieren. Dieses Phänomen wird als *Common Amplitude Modulation (CAM)* bezeichnet und ist wichtig für die auditive Wahrnehmung [WLW08]. Für jeden Oberton wird die Magnituden-Hüllkurve mit den Hüllkurven der anderen Obertöne korreliert. Anschließend wird die Hüllkurve mit der größten durchschnittlichen Korrelation ausgewählt und alle anderen damit gewichtet. Dadurch werden die Hüllkurven der Obertöne aneinander angepasst, um das erwähnte Phänomen besser zu erfüllen.

#### 4.5 Spektrale Maskierung - Spectral Masking

Durch die vorhergehenden Methoden wird eine sehr gute Annäherung des Spektrums eines Solo-Instrumentes erreicht. Hieraus wird eine Maske im Zeit-Frequenz-Bereich erstellt, welche die spektralen Bereiche markiert, die als Solo-Instrument eingestuft werden (siehe Abbildung 4). Das Magnituden-Spektrogramm des Soloinstrumentes wird mit Hilfe des Hadamard-Produktes aus dem ursprünglichen Magnituden-Spektrogramm des Audiomischsignals und der Maske erzeugt. Das Magnituden-Spektrogramm der Begleitspur entsteht durch die Subtraktion des Solo-Spektrogramms vom Spektrogramm des Audiomischsignals.

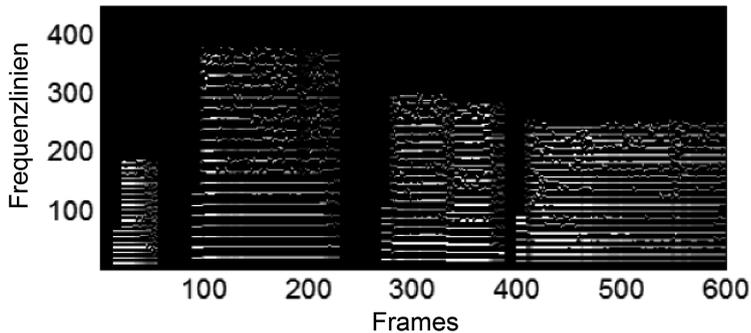


Abb. 4: Beispielhafte spektrale Maske

#### 4.6 Synthese - Re-synthesis

Das erhaltene Magnituden-Spektrogramm des Solo-Instrumentes sowie das Phasen-Spektrogramm des Audiomischsignals werden verwendet, um das komplexe Spektrogramm zu erstellen. Aus diesem wird mittels der inversen STFT das Audiosignal der Solostimme gewonnen. Gleiches gilt für die Begleitspur.

## 5 Evaluation

### 5.1 Nutzerstudie

Als ein Anwendungsszenario für die Klangquellentrennung in Solo- und Begleitspur wurde ein Hörtest mit 10 Personen im Alter zwischen 27 und 34 Jahren mit unterschiedlichem musikalischem Vorwissen durchgeführt, der die Qualitätsanforderungen an die Quellentrennung im Kontext von Musiklernsoftware evaluieren sollte. Die Ergebnisse dieses Hörtests verdeutlichen, dass die Möglichkeit, die Lautstärke von Solo- und Begleitspur zu verändern ein hilfreiches Feature für Musiklernsoftware darstellt. Außerdem zeigte sich, dass für Solo- und Begleitspur unterschiedliche Qualitätskriterien gelten.

### 5.2 Automatische Evaluation

Des Weiteren wurde die vorgestellte Methode zu den Evaluationskampagnen SiSEC 2011 und SiSEC 2013 eingereicht [ANV12]. Dort konnten vergleichbare Ergebnisse im Hinblick auf perzeptuelle Bewertungsmaße erzielt werden. Im Gegensatz zu den anderen eingereichten Algorithmen benötigte das hier beschriebene Verfahren nur einen Bruchteil der Verarbeitungszeit. Da der Datensatz dieser Kampagnen nur Gesang als Melodieinstrument beinhaltet, wurde eine zusätzliche Evaluation auf einem Datensatz mit einer Vielzahl von

Melodieinstrumenten durchgeführt. Die Qualität des State of the Art-Referenzalgorithmus [RP13] wurde hierbei übertroffen.

## 6 Instrumentenanalyse

Um die Schätzung der spektralen Komponenten von Solo-Instrumenten und damit die Qualität der Quellentrennung zu verbessern, wurde in einer weiteren Studie das Verhalten der Parameter Magnitude, Phase und Frequenz untersucht. Diese lassen sich mit Hilfe der bereits erwähnten STFT berechnen. Für die ersten zehn Obertöne aller Instrumente wurden die jeweiligen Differenzen der Magnitude zwischen aufeinanderfolgenden Frames analysiert. Dabei zeigte sich, dass die durchschnittliche Änderung kleiner als 2% ist. Nur bei Gitarre und Klavier ergaben sich größere Abweichungen. Dies erklärt sich durch die Art der Tonerzeugung: Nur bei diesen beiden Instrumenten wird die Saite zur Tonerzeugung angeschlagen, was zu zeitlich weniger stabilen Tönen führt.

Für die Frequenzanalyse wurden die Positionen der ersten zehn Obertöne untersucht und die durchschnittliche Abweichung von der zu erwartenden Frequenz berechnet. Alle Instrumente außer Gitarre und Piano wiesen Abweichungen von weniger als  $\pm 1\%$  auf. Höhere Obertöne wichen dabei stärker ab als die ersten Vielfachen der Grundfrequenz. Auch diese Ergebnisse unterstreichen, dass sich Saiteninstrumente, bei welchen die Saite zur Tonerzeugung angeschlagen wird, inharmonischer verhalten.

Hinsichtlich der Phase wurden die Parameter *Phase Expectation* und *Phase Coupling* analysiert. Für die Phase Expectation wird angenommen, dass die Phasenänderung einer harmonischen Klangquelle vorhergesagt werden kann. Das Phase Coupling beschreibt, dass sich die Phasen-Abweichung zwischen zwei Frames in drei harmonisch zusammengehörenden Obertönen gleicht. Mit Hilfe der Phase Expectation konnten Phasenänderungen mit großer Genauigkeit vorhergesagt werden. Die Überprüfung des Phase Coupling erfolgte, indem die Phase jedes Obertones durch die Phasen der zwei harmonisch zugehörigen Obertöne berechnet wurde. Für alle Instrumente nahm der Schätzungsfehler für Obertöne höherer Ordnung zu, trotzdem blieb der maximale Fehler unter 0,4 rad.

## 7 Harmonisch/Perkussive Klangquellentrennung

Die Ergebnisse der Phasenanalyse, speziell der Phase Expectation, wurden genutzt, um einen neuen Ansatz der Quellentrennung zwischen harmonischen und perkussiven Instrumenten zu testen. Der grundlegende Ansatz baut darauf auf, dass die Phasenänderung zwischen aufeinander folgenden Frames für harmonische Instrumente vorhergesagt werden kann, während die Phasenänderung perkussiver Instrumente ein höherer Prädiktionsfehler entsteht. Durch diese Analyse werden die einzelnen Bereiche des Magnituden-Spektrogramms den zwei Kategorien zugeordnet. Durch das gleiche spektrale Maskierungsverfahren wie bei der tonhöhen-informierten Quellentrennung werden die Spektrogramme von Solo- und Begleitstimme erhalten und anschließend in den Zeitbereich zurück transformiert. Eine Evaluation ergab eine verbesserte perzeptuelle Qualität der harmonischen und perkussiven Signale gegenüber vergleichbaren Methoden [Fi10] nach dem Stand der Technik.

## 8 Fazit

In dieser Dissertation wurden neuartige Ansätze zur Klangquellentrennung entwickelt, welche die Ergebnisse in verschiedenen Anwendungsbereichen verbessern. Die Ergebnisse der Instrumentenanalyse wurden genutzt, um die Trennung harmonischer und perkussiver Quellen anhand der Phaseninformation akkurat durchzuführen. Der vorgestellte Algorithmus der tonhöhen-informierten Quellentrennung zeigte sich als robuste und schnelle Methode für die Separation von Solo-Instrument und Begleitung heraus, was durch Nutzerstudien sowie wissenschaftliche Evaluationskampagnen unterstützt wurde. Die Verwendung des entwickelten Systems in der mittlerweile kommerziell verfügbaren Musiklernsoftware Songs2See demonstriert die Praxistauglichkeit. Die Bedeutung der Marktreife von Technologien aus dem Gebiet des Music Information Retrieval belegt die Auszeichnung des Songs2See-Teams, bestehend aus Christian Dittmar, Sascha Grollmisch und Estefania Cano Ceron, mit dem Innovations- und Entrepreneurpreis 2012 durch die Gesellschaft für Informatik.

## Literaturverzeichnis

- [ANV12] Araki, Shoko; Nesta, Francesco; Vincent, Emmanuel: The 2011 Signal Separation Evaluation Campaign (SiSEC2011): Audio Source Separation. In: Latent Variable Analysis and Signal Separation. S. 414–422, 2012.
- [Br96] Brown, Judith C.: Frequency Ratios of Spectral Components of Musical Sounds. Journal of the Acoustical Society of America, 99(2):1210–1218, 1996.
- [Dr11] Dressler, Karin: Pitch Estimation by the Pair-Wise Evaluation of Spectral Peaks. In: AES 42nd International Conference on Semantic Audio. Ilmenau, Germany, S. 1–10, 2011.
- [FCC05] Fitzgerald, Derry; Cranitch, Matt; Coyle, Eugene: Non-Negative Tensor Factorisation for Sound Source Separation. In: Irish Signals and Systems Conference. Dublin, Ireland, 2005.
- [Fi10] Fitzgerald, Derry: Harmonic/Percussive Separation Using Median Filtering. In: 13th International Conference on Digital Audio Effects (DAFx-10). Graz, Austria, S. 10–13, 2010.
- [Fi11] FitzGerald, Derry: Upmixing from mono; a source separation approach. In: 17th International Conference on Digital Signal Processing. Corfu, Greece, 2011.
- [LC65] Luce, David; Clark, Melville Jr.: Durations of Attack Transients of Nonpercussive Orchestral Instruments. Journal of the Audio Engineering Society (AES), 13(3):194–200, 1965.
- [Le01] Lee, Daniel D; Laboratories, Bell; Hill, Murray; Y, H Sebastian Seung: Algorithms for Non-negative Matrix Factorization. 1, MIT Press, Lee2001, S. 556–562, 2001.
- [Li13] Liutkus, Antoine; Durrieu, Jean-Louis; Daudet, Laurent; Richard, Gaël: An Overview of Informed Audio Source Separation. In: 14th International Workshop on Image and Audio Analysis for Multimedia Interactive Services. Paris, France, S. 3–6, 2013.
- [RP13] Rafii, Zafar; Pardo, Bryan: REpeating Pattern Extraction Technique (REPET): A Simple Method for Music / Voice Separation. IEEE Transactions on Audio, Speech and Language Processing, 21(1):73–84, 2013.

- [Sa13] Salamon, Justin; Gómez, Emilia; Ellis, Daniel P. W.; Richard, Gaël: Melody Extraction from Polyphonic Music Signals : Approaches, Applications and Challenges. IEEE Signal Processing Magazine, 2013.
- [WLW08] Woodruff, John; Li, Yipeng; Wang, DL: Resolving Overlapping Harmonics for Monaural Musical Sound Separation using Pitch and Common Amplitude Modulation. In: 9th International Society for Music Information Retrieval Conference (ISMIR 2008). Philadelphia, USA, S. 538–543, 2008.



**Estefanía Cano** absolvierte ein Studium der Elektrotechnik (B.Sc.) an der Universidad Pontificia Bolivariana - Kolumbien, ein Musikstudium (B.A.) an der Universidad de Antioquia - Kolumbien und ein Musikingenieursstudium an der University of Miami - USA.

Seit 2009 forscht Dr. Cano am Fraunhofer-Institut für Digitale Medientechnologie IDMT, wo sie ihre Dissertation mit dem Titel *Tonhöheninformierten Trennung in Solo- und Begleitspuren* im Jahr 2014 erfolgreich verteidigte.

Ihr Forschungsschwerpunkt ist die Klangquellentrennung, insbesondere in den Bereichen Tonhöheninformierten Klangquellentrennung und Harmonisch/Perkussive Klangquellentrennung. Ihre weiteren Forschungsinteressen liegen in der Analyse und Modellierung von musikalischen Klängen, der Nutzung von Phaseninformationen im Music Information Retrieval, der Klassifikation von Musikinstrumenten und dem Einsatz von Techniken des Music Information Retrieval für musikwissenschaftliche Analysen.