

Klangmodellierungsansätze und Ergebnisse für Gamelaninstrumente des lauten Stils

Alois Sontacchi, Franz Zotter und Robert Höldrich

Institut für Elektronische Musik und Akustik, 8010 Graz, Österreich, Email: sontacchi@iem.at

Einleitung

Dieser Beitrag beschäftigt sich mit der Klangmodellierung ausgewählter Idiophone aus dem Gamelaninstrumentarium. Dabei müssen die Steuerbarkeit der Stimmung und Klangfarbe der einzelnen Instrumente, sowie die Zuordnung der Teiltöne zu unterschiedlichen räumlichen Abstrahlungseigenschaften (siehe auch [1]) gewährleistet werden.

Für die Musikethnologie sollte als Ergebnis eine echtzeitfähige interaktive Implementierung eines virtuellen Gamelanorchersters stehen, die speziell auch für Feldforschungen eingesetzt werden kann und durch mobiles Equipment (Laptop) realisierbar sein sollte. Daher schließt sich zusätzlich der Wunsch nach einer datenreduzierten Klangrepräsentation mit originalgetreuer Klangqualität an.

Methodik

Die Idiophone werden angeschlagen und zeigen im Ausklang zum Teil nichtlineares Verhalten wie *pitch glides*. Bedingt durch die spezielle Signalform wurden folgende drei Techniken untersucht, die von unterschiedlichen Voraussetzungen an das Signalmodell ausgehen: Auto Regressive Moving Average Modell mit Frequency Zooming (FZ-ARMA), Warped Spectral Model (WSM) und Empirical Mode Decomposition (EMD).

Die Unterschiede und Auswirkungen bei Analyse und Synthese werden aufgezeigt und die klanglichen Ergebnisse einander gegenübergestellt.

Modellierungsverfahren

FZ Autoregressive Moving Average (FZ-ARMA)

Anhand von Untersuchungen an Lamellophonen (handgezapfte Metallzungen) [2] konnte gezeigt werden, dass deren Klang als eine Summe von exponentiell gedämpften Sinusschwingungen hinreichend modelliert werden kann. Bei den vorliegenden Klängen ist die Signalgrundform entsprechend ähnlich. Bedingt durch die unterschiedliche Bauformgröße der einzelnen Instrumente werden jedoch einerseits weitaus größere Bandbreiten abgedeckt, andererseits liegen Teiltöne sehr eng nebeneinander. Beim Gong Ageng z.B. liegen die Frequenz des 2. Schwingungsmodus (87.7Hz) und die 2. Harmonische des ersten Modus (43Hz) (hervorgerufen durch die Nichtlinearität der Rückstellkraft, vgl. [3]) nur um 1.7Hz auseinander.

Bei einer Modellierung des Signals durch ein Zähler- und ein Nennerpolynom (ARMA Modell) würde eine extrem hohe Anzahl von Pol- und Nullstellen erforderlich sein. In der Praxis zeigt sich meist ein numerisches Problem bei der Berechnung der komplexen Wurzeln dieser Polynome höherer Ordnung. Beim FZ-ARMA-Modeling (vgl. [4]) werden die jeweils relevanten Frequenzkomponenten gegen 0Hz moduliert. Darauf folgt eine Bandbegrenzung und Abstratenreduktion um einen Faktor K . Dadurch kann auf

die einzelnen Resonanzen vereinfacht zugegriffen werden. Gleichung 1 beschreibt die Modifikation der gesuchten Wurzeln z_i , die sich bei der beschriebenen Vorgangsweise ergibt. Die Rotation der Pollage und Stauchung des Polvektors muss nach Auffinden der Pole wieder rückgängig gemacht werden.

$$z_{i, zoom} = |z_i|^K \cdot e^{j(\arg(z_i) - \Omega_m)K} \quad (1)$$

Aus der Lage der Pole kann man deren Resonanzfrequenz f_i und Dämpfung a_i bestimmen. Startamplitude und -winkel der einzelnen exponentiell abklingenden Sinusschwingungen kann man aus einem LMS-Ansatz durch lineare Kombination (komplexe Gewichte) der Einzelschwingungen bestimmen [5].

Durch die Subbandanalyse des originalen Klangsample ist die Verteilung der Pole pro Frequenzband steuerbar. Dadurch kann trotz geringer Gesamtanzahl von Analyseparameter ein perceptiv weitgehend originalgetreues Klangabbild der Synthese erreicht werden.

Empirical Mode Decomposition (EMD)

Die EMD Methode wurde in den 90iger Jahren von Huang et al. [6] entwickelt. Gegenüber der *Fourier Analyse* muss das Analysesignal weder lineare noch stationäre Eigenschaften besitzen. Das Analyseverfahren erfolgt signalangepasst und somit adaptiv. Durch einen Siebungsprozess (vgl. [6]) werden so genannte Intrinsische Modenfunktionen (IMF) aus dem Signal rekursiv extrahiert. Zuerst werden die hochfrequenten und zum Schluss die tieffrequenten Signalkomponenten extrahiert. Die IMF stellen reine frequenz- und amplitudenmodulierte Signale dar, für die zu jedem Zeitpunkt die Momentanfrequenz und Amplitude bestimmt werden kann. Es gilt jedoch zu beachten, dass für das Auffinden einer physikalisch sinnvoll zu interpretierenden Momentanfrequenz der Signalausgang entsprechend groß sein muss (d.h. es liegt tatsächlich ein Mode vor). Allgemein ist der Fehler der Momentanfrequenzschätzung umso größer, je näher diese bei Hälfte der Abtastrate liegt (vgl. [7]). Liegt die Frequenz eines Modes exakt mit einem rationalen Verhältnis der Abtastrate, geht der Fehler weitgehend zurück.

Ebenfalls kann der Algorithmus zwei Signalkomponenten nur dann vollständig isolieren, wenn deren Frequenzabstand bei gleicher Amplitude größer als ein Faktor 2 ist. Ist die Signalamplitude der höherfrequenten Signalkomponente deutlich geringer, dann wird nur die tieffrequenten Signalkomponente detektiert, solange der Frequenzabstand wiederum nicht zu groß (abh. von den Amplitudenverhältnissen) wird.

Der hier verwendete modifizierte EMD-Ansatz geht davon aus, dass die relevanten Frequenzkomponenten bekannt sind. Der Bereich um diese Komponenten wird mittels Bandpass (Güte = 10) gefiltert und mit Einseitenbandmodulation gegen tiefe Frequenzen verschoben. Diese Subbänder werden bzgl. IMFs analysiert. Anschließend werden diese

wieder in ihr ursprüngliches Band zurückgeschoben. Befindet sich in einer IMF tatsächlich nur eine Signalkomponente, dann kann die Momentanfrequenz und der Amplitudenverlauf ohne Verlust von Information unterabgetastet (datenreduziert) werden.

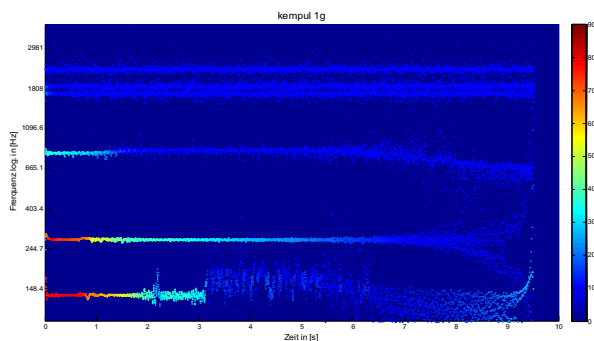


Abbildung 1: Analyseergebnis des mod. EMD Verfahrens anhand des Klages „kempul 1g“ („glides“ bei ca. 1sek).

Warped Spectral Modelling

Aufbauend auf die Arbeit von Härmä (vgl.[8]) wurde ein Modell entwickelt, das sich weitgehend an einer gehörgerechten Analyse orientiert. Das Modell umfasst folgende Verarbeitungsschritte:

- Frequenzverzerrte Kurzzeit-FT(warped STFT)
- Berücksichtigung der Außen/Mittelohr-Übertragungsfunktion
- Berücksichtigung der Simultanmaskierung
- Berechnung der Mithörschwelle
- Kriterienabhängige Trajektorien für Teiltonverläufe
- Transientenextrahierung

Das Erzeugung der zeitlich verlaufenden Teiltonspuren erfolgt mit Hilfe von Übergangsbewertungsmatrizen, welche Attribute wie Lautstärke der Teiltöne, Alter der Trajektorie, Frequenzänderung eines Teiltons usw. für jede einzelne Spur iterativ bewerten.

Aufgrund der geringen zeitlichen Auflösung - die vor allem am Beginn der Klänge von Bedeutung ist - werden für den transienten Anschlag Residuen bestimmt werden. Diese erhält man durch Filterung des Originalsignals durch zeitvariante Kerbfilter, die durch die Teiltonverläufe angesteuert werden. Der frühe Ausschnitt des resultierenden Residuums wird als Anschlagkomponente den Teiltonkomponenten hinzugefügt.

Diskussion und Ergebnisse

Der FZ-ARMA Algorithmus liefert bzgl. der Datenkompression die besten Ergebnisse. Diese liegen im Bereich über 90%. Beim Vergleich der Resynthesesignale mit den Originalen wird ein perceptiv weitgehend zufriedenstellendes Ergebnis erreicht, wobei bei nichtlinearen Klangplatten (*gliding modes*) das Modell aufgrund seiner Annahmen scheitert.

Obwohl der EMD-Algorithmus auf den ersten Blick durch seinen adaptiven Charakter vielversprechend erscheint, zeigt sich bei näherer Betrachtung, dass a priori Wissen über das zu analysierende Signal vorhanden sein muss. Durch diese Zusatzinformationen können klanglich respektable Ergebnisse erzielt werden, wobei eine Datenreduktion nur dann möglich ist, wenn in jeder IMF wirklich nur ein Schwingungsmodus zu finde ist. Allgemein können auch

gliding modes mit dem modifizierten EMD-Ansatz weitgehend abgebildet werden (vgl. Abb.1).

Die besten klanglichen Ergebnisse, bei ebenso erfolgreicher Datenreduktion (ca. 90%), liefert die Warped Spectra Modellierung unter Zuhilfenahme der Residuen. Sie stellt unter Betrachtung der diversen Signalformen das flexibelste der drei untersuchten Verfahren dar (vgl. Abb.2).

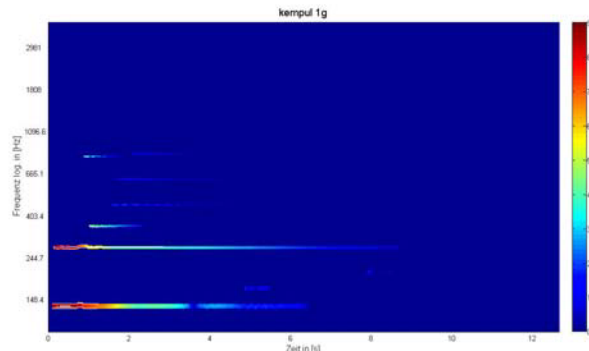


Abbildung 2: Analyseergebnis des WMS Verfahrens für Klang „kempul 1g“, ohne Resid. („glides“ bei ca. 1sek).

Danksagung

Die Arbeit entstand im Rahmen des Forschungsprojekts „Virtual Gamelan Graz“, das durch Fördermittel des Zukunftsfonds Steiermark ProjektNr. 3027 ermöglicht wird.

Literatur

- [1] F. Zotter, A. Sontacchi, R.Höldrich, “Modeling a Spherical Loudspeaker System as Multiple Source”, DAGA 2007.
- [2] G. Grupe, A. Sontacchi und R. Höldrich, ”Pitch Perception in Lamellophones and the Development of a Computer Based Tool for Interactive Experiments on Tunings and Tonal Systems”, CIM04, Graz, 2004.
- [3] H. Fleischer, Schwingungsverhalten an Gongs, H. Fleischer u. H. Fastl, Inst. f. Mechanik, Fakultät f. Luft- und Raumfahrttechnik, Universität der Bundeswehr München, 85577 Neubiberg, Deutschland, 2001.
- [4] Matti Karjalainen, Paulo A. A. Esquef, Poju Antsallo, Aki Mäkivirta, and Vesa Välimäki, „AR/ARMA Analysis and Modeling of Modes in Resonant and Reverberant Systems“, 112th AES Conv., 2002.
- [5] S. Lawrence Marple, Jr., “Digital Spectral Analysis with Applications”, Prentice Hall, Signal Processing Series, Prentice-Hall Inc. Englewood Cliffs, New Jersey 07632, 1982. ISBN 0-13-214149-3 025
- [6] N. E. Huang, Z. Shen, M.C.Wu, S.R. Long, H. H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis”, in Proc. R. Soc. Lond. A, London, UK, 1998.
- [7] G. Rilling, P. Flandrin and P. Goncalvès, “On Empirical Mode Decomposition and its Algorithms,” IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing NSIP-03, Grado (I), 2003.
- [8] A. Härmä, “Frequency-warped autoregressive modeling and filtering,” PhD-Thesis, Helsinki University of Technology, P.O. BOX 1000, FIN-02015 HUT, 2001.