

institut für elektronische musik und akustik



IEM Report 42/07:

**Klangmodellierung - ein Beitrag zur musikalischen Akustik
Endbericht zum Projekt „Virtual Gamelan Graz (VGG)“**

Verfasser:

Alois Sontacchi, Franz Zotter, Gerhard Eckel, Robert Höldrich

Kooperationspartner:

Institut für Musikethnologie (IME)

Zukunftsfonds Steiermark (GZ: A3-25R9-05/19)

November, 2007

IEM - INSTITUT FÜR ELEKTRONISCHE MUSIK UND AKUSTIK

Vorstand: VProf. Dr. Gerhard ECKEL

A-8010 Graz, Inffeldgasse 10/3, Tel.:+43/(0)316/389 – 3170, FAX:+43/(0)316/389 – 3171

office@iem.at

<http://iem.at/>

Zusammenfassung

„Virtual Gamelan Graz (VGG)“ ist ein Projekt, das im Hinblick auf musikethnologische Forschung neue Möglichkeiten bieten soll, welche auch den gegenwärtigen Stand der Entwicklung im Bereich der spektralen und spatialen Klangsynthese sowie der algorithmischen Verfahren zur Generierung von musikalischen Strukturen nutzen. Es sollte also ein virtuelles Gamelan-Orchester erzeugt werden, welches selbständig Kompositionen der klassischen Musik Zentraljavas in verschiedenen Stimmungen und modifizierbaren Klangfarben generiert, und wodurch das Spiel auf unterschiedlichen Sets von Instrumenten simuliert werden kann.“

Für die sinnvolle Synthetisierung der besagten Klänge (es handelt sich grundlegend um Instrumente des lauten Still) ist es jedoch von absoluter Notwendigkeit, die genaue zeitliche spektrale Zusammensetzung eines jeden Klangs zu kennen. Darüber hinaus wird bei der Klangmodellierung auch deren erfasstes räumliches Abstrahlungsverhalten bei der Wiedergabe berücksichtigt. Dieses zusätzliche räumliche Attribut verleiht dem virtuellen Instrument eine weitere Komponente, die zur Erreichung einer realitätsnahen Abbildung und somit wahrgenommenen Natürlichkeit führt.

Inhaltsverzeichnis

Überblick.....	2
1. VGG Implementierungskonzept	3
2. Klangsyntheseverfahren.....	6
2.1. Anforderungen an die Klangmodellierung.....	6
2.2. Auswahl von geeigneten Klangmodellierungsverfahren	7
- FZ-ARMA	8
- (a)WSM	12
- EMD.....	15
3. Aufnahme und Wiedergabe der akustischen Abstrahlung	18
3.1. Hintergrund.....	18
3.2. Aufnahme der akustischen Abstrahlung.....	20
3.3. Analyse der Abstrahlungsmuster von Teiltönen.....	21
3.4. Wiedergabe der Akustischen Abstrahlung	22
3.5. Umsetzung und Optimierung des Wiedergabeverfahrens	23
Literatur.....	24
Anhang	26

Überblick

Der vorliegende Bericht beschäftigt sich vorwiegend mit den Ergebnissen für den Aspekt der Klangmodellierung des Projektes „Virtual Gamelan Graz“, die am Institut für Elektronische Musik und Akustik erarbeitet wurden.

Der Begriff „Klangmodellierung“ beschreibt beide Attribute *Klangfarbe* und *Raumklang*. Die Eigenschaft der Klangfarbe kann weitgehend durch klassische Klangsyntheseverfahren beschrieben und realisiert werden. Innerhalb dieses Projektes wurden dazu relevante Ansätze, welche die geforderten Eigenschaften (siehe Vorgaben Klangsynthese, Abschnitt 2.1) weitgehend erfüllen, untersucht und einander bzgl. Klangqualität und Implementierungsaufwand gegenübergestellt. Dazu wurden zwei Verfahren herausgegriffen und speziell für diese Anwendung weiterentwickelt. Die Umsetzung des Raumklangs wurde durch ein spezielles Aufnahme- und Wiedergabeverfahren realisiert. Dabei wird jedem Teilton neben seiner zeitlichen spektralen Entwicklung eine räumliche Abstrahlungsinformation zugeordnet.

Eng mit der Klangmodellierung und deren Umsetzung ist die zugrunde liegende Struktur für das Gesamtsystem verbunden. Die Aspekte für die Umsetzbarkeit und Implementierung eines virtuellen Orchesters wurden am IEM entwickelt und werden im Rahmen dieses Berichtes beschrieben.

Die Resultate zur automatischen Generierung von Kernmelodien (balungan) des javanesischen Gamelans basierend auf nichtwissensbasierte Methoden werden getrennt in einem anderen Bericht dargestellt [1].

Die Ergebnisse zur Ableitung der Parts und der Umsetzung einer Lösungsalternative des virtuellen Orchesters in Super Collider unserer Kollegen vom Institut für Musikethnologie finden sich in deren Abschlussbericht [2]

1 VGG Implementierungskonzept

In Abbildung 1 ist das Konzept für die Implementierung eines virtuellen Gamelan Orchesters dargestellt. Sämtliche Module mit Ausnahme jenes zur Generierung der Richtwirkung werden in Super Collider realisiert. Die Module Init/Control, Partableitung, Klangsynthese, Abstrahlcharakteristik befinden sich innerhalb eines Programms oder bei Verteilung auf unterschiedliche Computer kommunizieren über unidirektionale „Open Sound Control“ (OSC, siehe Anhang) Verbindungen.

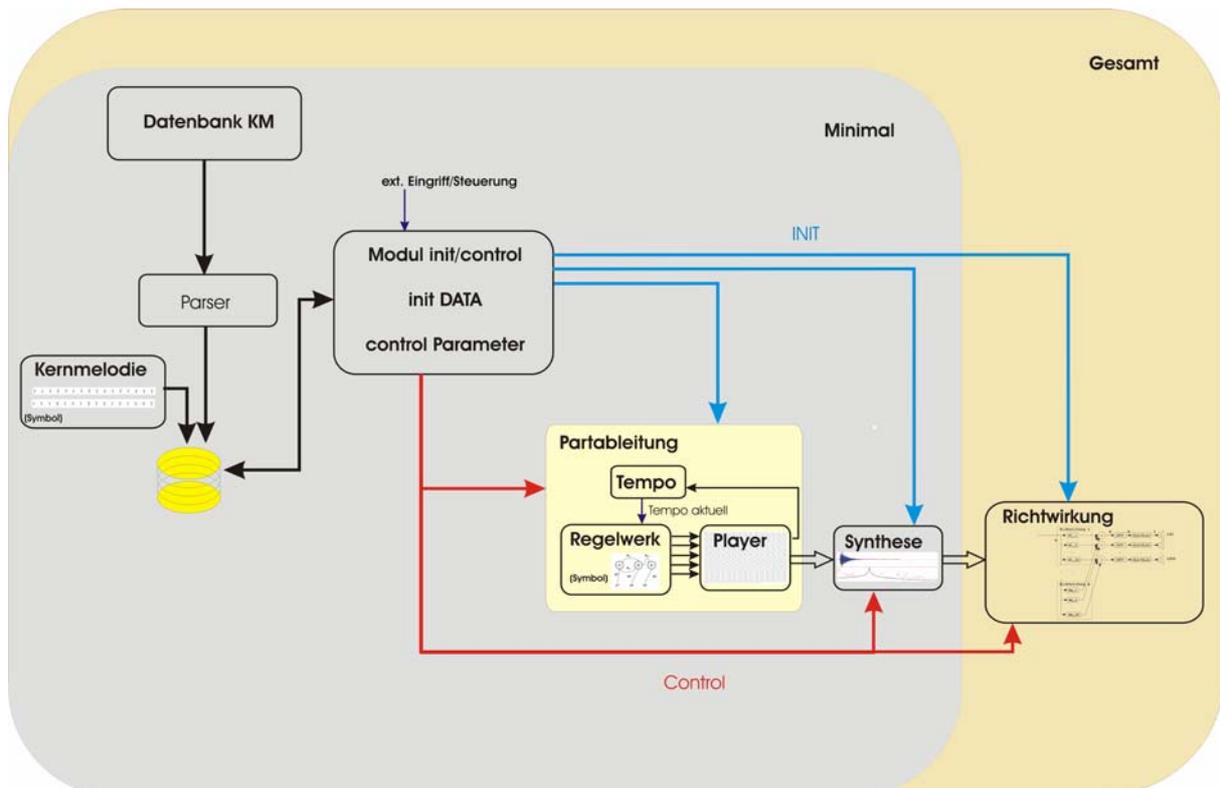


Abb.1: Strukturschaltbild der VGG-Implementierung

Modul Kernelmelodie

Die Ergebnisse des Moduls Kernelmelodie werden direkt in ein Archiv gespeichert. Dieses Modul ist während des Realisierungsprozesses nicht aktiv. Bestehende Kernelmelodienotationen aus Datenbanken können ebenfalls in dieses Archiv aufgenommen werden. Diese müssen jedoch zuvor mittels eines adäquaten Parser in das VGG – kompatible Datenformat konvertiert werden.

Die diversen Kernelmelodien werden als Listen in editierbaren Text-Dateien abgelegt und durch einen entsprechenden Filenamen bezeichnet. Der Aufbau dieser Liste könnte wie folgt aussehen:

Filename.txt : { [Subliste], [Subliste], [Subliste], [Subliste]...}

Wobei eine Subliste wie folgt definiert ist:

Alternative 1: [Tonhöhe_{relativ}, Oktavlage, Dauer]

Alternative 2: [Tonhöhe_{absolut}, Dauer]

Bsp f. 7-stufiges System: Alternative1 [1, 2, 4] bzw. Alternative2 [8, 4]

Modul Init/Control

Das Modul Init/Control sorgt nach Programmstart für das Laden der gewünschten Kernmelodie aus dem Datenarchiv und informiert die übrigen Module über das gewünschte Ausgangssetup (Initialisierung). Darüber hinaus werden während der Programmlaufzeit Eingabedaten, die das VGG-System modifizieren können erfasst und den jeweiligen Modulen mitgeteilt (Control).

Initialisierung (der Zustände 1 bis 6) */vgg/init/ZustandsNr „Daten“*

- 1) Laden der gewünschten Kernmelodie { *Pfad, Filename* }
- 2) Festlegen des Tonsystems { *5, 7* }
- 3) Vorgabe der Stimmung { *arg{Auswahl Tonsystem}, Intervalle, Referenzton¹* }
- 4) Festlegen der Besetzung { *[Instr.Name,Anzahl], [Instr.Name,Anzahl], ...* }
- 5) Verstimmungsangaben zu den Teiltönen { *[Instr.Name, Ton, Teiltonindex, Cent], [Instr.Name, Ton, Teiltonindex, Cent], ...* }
- 6) Tempo { *Wert* }
- 7) Wahl des Wiedergabesystems { *arg{ Auswahl: Mono, Stereo, Surround, Ikosaeder(Instr.Name)}, Ikosaeder(Instr.Name)} & Orch., Binaural* }

Control (der Zustände 3 bis 6) */vgg/control/ZustandsNr „Daten“*

Während der Laufzeit können folgende Zustände modifiziert werden:

- 3) Stimmung; Datenstruktur wie oben bei Initialisierung, wobei erster Eintrag *arg{Auswahl Tonsystem}* bleibt während der Ausführung unverändert. Die Mitteilung erfolgt nur an das Modul Klangsynthese, alle übrigen nehmen von dieser Information keine Notiz.
- 4) Die Änderungsangaben zur Besetzung erfolgen durch: *{[Instr.Name, ,ON' bzw. ,OFF']...}*
Die Mitteilung erfolgt nur an das Modul Klangsynthese, alle übrigen nehmen von dieser Information keine Notiz.
- 5) Verstimmungsangaben der Teiltöne; Datenstruktur wie oben bei Initialisierung. Die Mitteilung erfolgt nur an das Modul Klangsynthese, alle übrigen nehmen von dieser Information keine Notiz.
- 6) Tempo {rel.Änderung(+/-)} extern; Datenstruktur wie oben bei Initialisierung.

¹ Angabe des „Kammertons“

Modul Partableitung

Das Modulpartableitung beginnt nach Bekanntgabe der Initialisierungsparameter (und durch abspielen der Kernmelodie) das Ableiten der Parts und geniert dadurch eine ‚online‘ Partitur. Ebenso werden während der Programmausführungsphase die Steuerdaten (s. Modul Init/Control) des variablen Zustands 6 (durch externe Einwirkung oder bedingt durch interne Konsequenz aus der Partitur folgend) in der Partableitung berücksichtigt.

Das Modul Partableitung liefert die Partitur als Listenaufträge an das Modul Klangsynthese via „Open Sound Control“ (OSC, Details s.u.) in folgender Form:

	/vgg/Instr.Name/Nummer	Ton	Anschlagstärke	Dämpfung	
Bsp:	/vgg/saronbarung/1	3	127	0	(max. Anschlagstärke)
	/vgg/saronbarung/1	3	0	1	(Ton wird gedämpft)

Modul Klangsynthese

Das Modul Klangsynthese erzeugt anhand der Vorgabe der Initialisierung und Modifikationen der Steuerdaten auf Anweisung durch das Modul Partableitung die entsprechenden Klänge. Diese Klänge werden anhand der Auswahl des Wiedergabemodus entweder direkt über einen (Mono), zwei (Stereo) oder fünf Lautsprecher+Subwoofer (Surround) bzw. Kopfhörer (Binaural) wiedergegeben. Die Wiedergabefähigkeit ist von der verfügbaren Rechenleistung und Audioausstattung abhängig.

Modul Richtcharakteristik

Sollte bei der Wiedergabe auch die Richtcharakteristik eines speziellen Instrumentes berücksichtigt werden, so ist neben dem zusätzlichen Modul Richtcharakteristik auch das speziell entwickelte Lautsprecherarray samt Versorgung erforderlich. Dieses Modul wird ebenfalls, wie bereits eingangs erwähnt, über die OSC-Verbindung angesprochen, wobei die Realisierung in PD erfolgt.

2 Klangsyntheseverfahren

2.1 Anforderungen an die Klangmodellierung

Primäres Ziel der Klangmodellierung ist die Steuerbarkeit der Klänge hinsichtlich deren Tonhöhen der einzelnen Klangkomponenten – den Teiltönen bzw. den Klangatomen. Dadurch wird natürlich auch die Klangfarbe des Gesamtklangs modifiziert. Die Akzeptanz und Präferenz der Klangfarbe als auch die Stimmung der einzelnen Klangkörper zueinander sind Untersuchungsgegenstand der Musikethnologie. Darüber hinaus sollte auch die Basis für die Steuerbarkeit der Abstrahlcharakteristik einzelner Teiltöne gelegt werden. Zusätzlich zur Steuerbarkeit wird auch eine Reduktion der Klangdaten angestrebt. Wobei Originaltreue die oberste Maxime ist. Die Originaltreue betrifft grundsätzlich die Hörwahrnehmung.

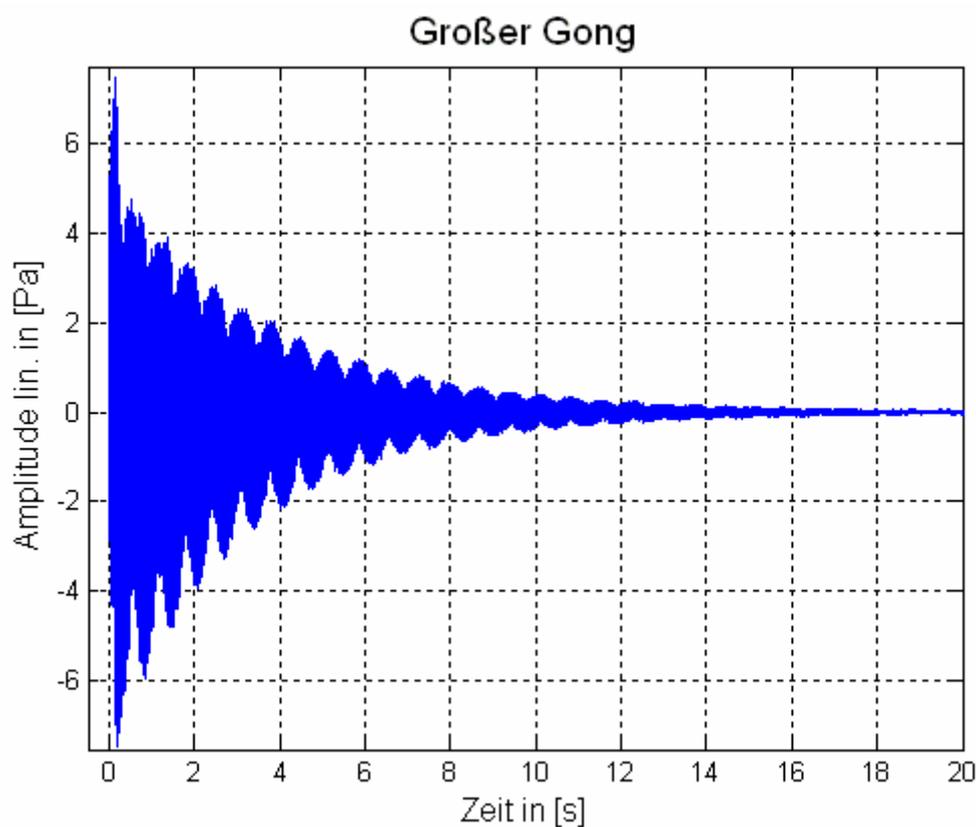


Abb. 2: Zeit- Lautstärke Diagramm für „großer Gong“

Gesucht wird eine mathematische Beschreibung – Modellierung der Klangsignale. Das aufgezeichnete akustische Schallsignal eines großen Gong kann man in einem Zeit- Lautstärke Diagramm (Abb. 2) oder aber auch in einem Zeit- Frequenz Diagramm darstellen (Abb. 3). Im Zeit- Frequenz Diagramm wird die Lautstärke der einzelnen Frequenzen zusätzlich durch Farben kartiert (rot entspricht laut, & blau bedeutet leise).

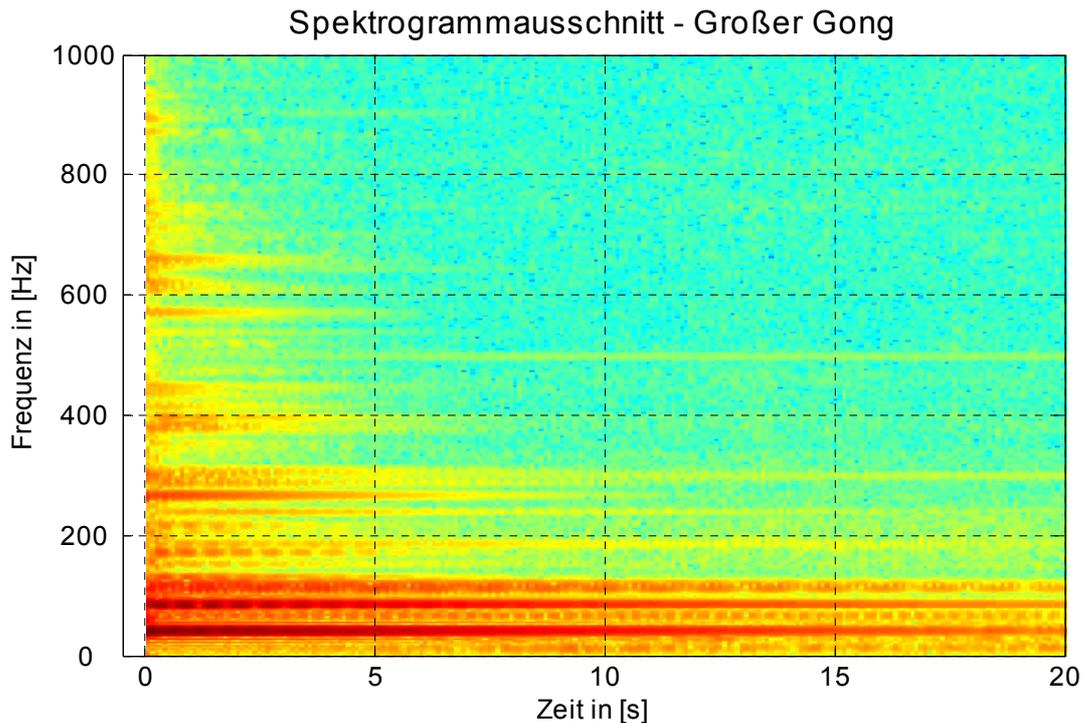


Abb. 3: Zeit- Frequenz Diagramm für „großer Gong“

Im Zeit-Frequenz Diagramm ergeben sich anhand der prominenten Frequenzen weitgehend horizontale Linienvläufe die als Klangspur oder im Fachterminus „Teiltonverlauf“ bezeichnet werden.

Somit kann man ein akustisches Zeitsignal $s(t)$ als Funktion von zeitlich veränderlichen Lautstärkeparameter und Frequenzparameter sowie weitere Parameter beschreiben werden d.h. : $s(t) = F\{a_i(n), f_i(n), \dots\}$, wobei t für die kontinuierliche Zeit steht; n für die diskreten Zeitpunkte, an denen die Parameter bestimmt werden und Index i steht für sämtliche differenzierbare Teilkomponenten.

2.2 Auswahl von geeigneten Klangmodellierungsverfahren

Basierend auf der mathematischen Signalbeschreibung können nun unterschiedliche Klangmodellierungsverfahren entwickelt werden. Folgende drei Verfahren, die unsere Vorgaben weitgehend abdecken, wurden näher behandelt.

- FZ-AMRA (Frequency Zooming Autoregressiv-Moving-Average)
- Warped Spectral Model
- Empirical Mode Decomposition

Diese Verfahren unterscheiden sich grundsätzlich in der Modellierung der Teiltonverläufe. Deren Grundform wird im folgenden Klangatom bezeichnet.

Klangmodellierungsverfahren - FZ-ARMA bzw. BFZ-ARMA

Bei „Frequency Zooming – Auto Regressive Moving Average“ (FZ-ARMA) Verfahren (vgl. [3]) setzt man auf ein Quelle-Filter-Model, wobei die Quelle durch einen Anschlag repräsentiert wird. Da die Instrumente mit unterschiedlichen Schlägeln gespielt werden, ergibt sich hierbei eine erste Anpassungsfähigkeit was zumindest das Anschlagsgeräusch betrifft. In Abbildung 4 sind unterschiedliche Schlägelausführungen dargestellt.



Abb. 4: Diverse Schlägelausführungen eines Gamelan Orchesters.

Der Filter wird durch das Schwingungsverhalten des Klangkörpers beschrieben. Er setzt sich somit aus unterschiedlichen Resonanzen zusammen. Der Grundtypus d.h. das Klangatom, aus dem in Folge ein Gesamtklang zusammengesetzt wird, ist hier eine exponentiell abklingende Sinusschwingung. Zur mathematischen Beschreibung genügen eine Frequenz für die Sinusschwingung, ein Lautstärkestartwert und ein Dämpfungsfaktor für den Abklingvorgang.

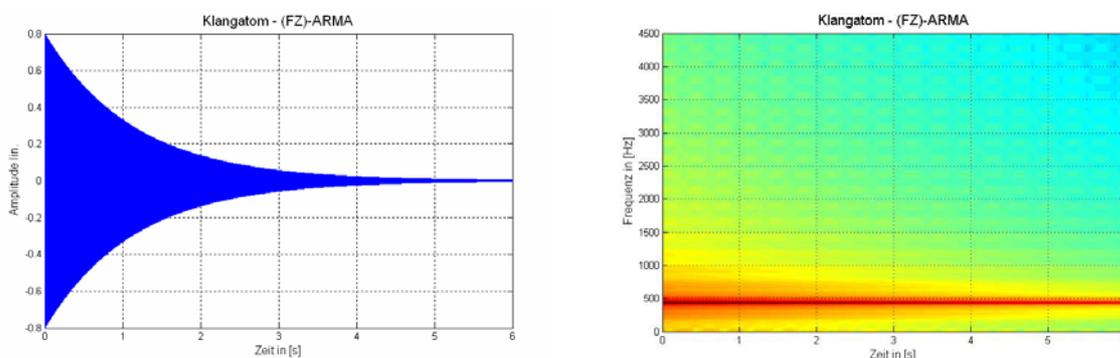


Abb. 5: Zeit- Lautstärke-Diagramm und Zeit- Frequenz-Diagramm für ein Klangatom zugehörig zum FZ-ARMA Verfahren.

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 werden die jeweils relevanten Frequenzkomponenten gegen 0Hz moduliert (Gl.1). Im nächsten Schritt erfolgt eine Bandbegrenzung und Abtastratenreduktion um einen Faktor K . Dadurch kann auf die einzelnen Resonanzen leichter zugegriffen werden. Gleichung 2 und 3 beschreiben die Modifikation der gesuchten Wurzeln z_i , die sich bei der beschriebenen Vorgangsweise ergibt (siehe Abb. 6). Die Rotation der Pollage (Modulation) und Stauchung des Polvektors (Abtastratenreduktion) muss anschließend nach Auffinden der Pole muss anschließend wieder rückgängig gemacht werden.

$$sig_m(n) = e^{j\Omega_m n} \cdot sig(n) \quad (\text{Gl.1})$$

mit $\Omega_m = 2\pi f_m / f_s$, wobei f_s Abtastrate und f_m Modulationsfrequenz

$$\Omega_{i,rot} = \Omega_i - \Omega_m = \arg(z_i) - \Omega_m \quad (\text{Gl. 2})$$

$$z_{i,zoom} = |z_i|^K \cdot e^{j(\arg(z_i) - \Omega_m)K} \quad (\text{Gl. 3})$$

Das Auffinden der Wurzeln kann z.B. durch das Verfahren der linearen Prediktion (Linear Prediction Coding, LPC) oder Steiglitz-McBright (vgl. [4]) und dergleichen gelöst werden.

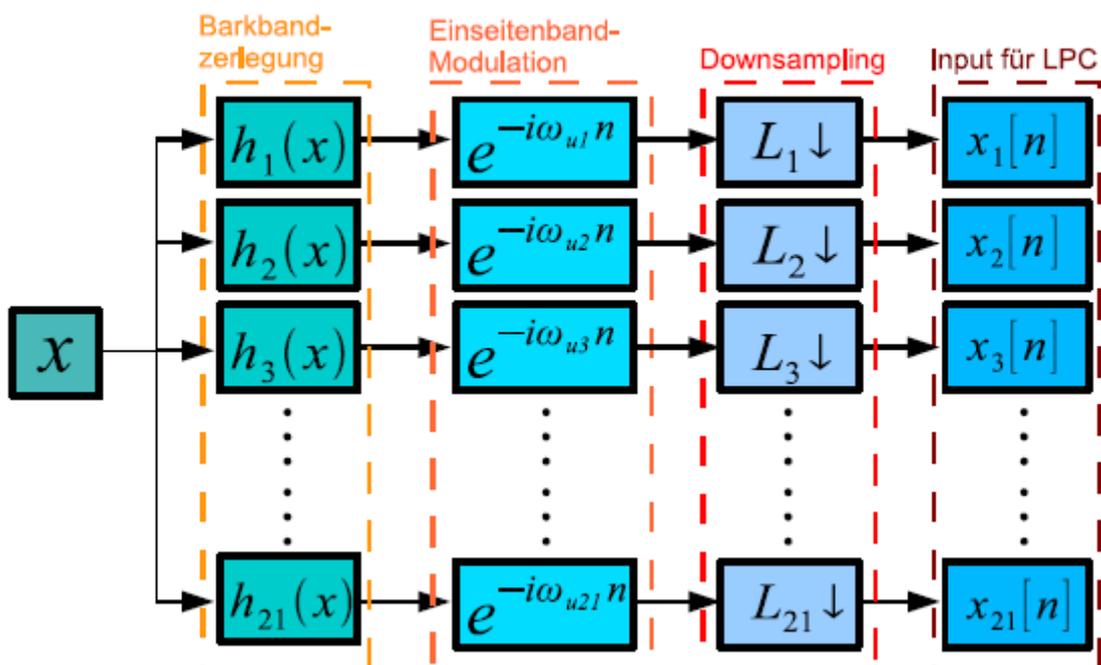


Abb. 6: Blockschaftbild der Analyse Filterbank.

Aus der Lage der Pole kann man deren Resonanzfrequenz f_i und Dämpfung a_i bestimmen. Die Startamplitude und Startwinkel der einzelnen exponentiell abklingenden Sinusschwingungen (der einzelnen Teiltöne) kann man aus einem (least squares) LS-Ansatz durch lineare Kombination (mit komplexe Gewichte) der Einzelschwingungen bestimmt werden.

Im Rahmen des Projektes Vitruval Gamelan Graz wurde dieser Ansatz wie folgt erweitert und verbessert:

Zum einen wurden zusätzlich auch die Eigenschaften des Gehörs bzgl. Maskierungseigenschaften (spektrale als auch temporale Verdeckung, irrelevante nicht hörbare Signalkomponenten werden durch dieses Verfahren nicht abgebildet, siehe Abb.6) miteinbezogen. Dadurch erfolgt eine Analyse innerhalb von den gehörspezifischen Barkbänder (daher auch die neue Bezeichnung *Bark Frequency Zooming - ARMA Model, BFZ-ARMA*). Darüber hinaus wurde durch einen weiteren Beschreibungsparameter, der die Einschwingphase der einzelnen Teiltöne betrifft erweitert. Physikalisch gesehen werden nunmehr Pole 2.Ordnung verwendet die neben einem exponentiellen Abfall/Ausklang einen linearen Einschwingvorgang/Anklang besitzen. Genauere Beschreibungen dazu finden sich in folgenden Publikationen [9],[14],[21] und [26].

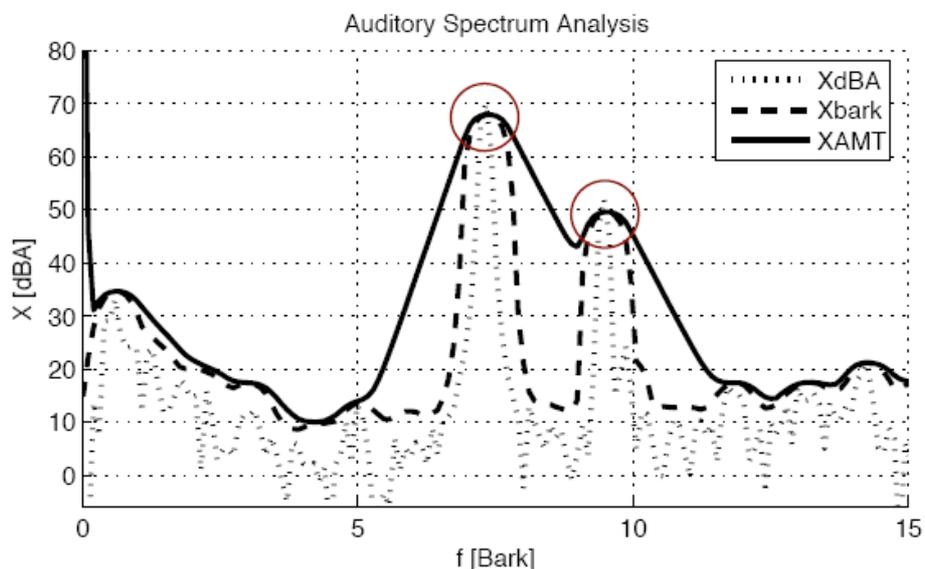


Abb. 7: Verlauf der Mithörschwelle für ein bestimmtes spektrales Erregungsmuster, sowie Detektierung der relevanten spektralen Komponenten.

Da für Feldversuche nur eine limitierte Hardwareausrüstung (Laptop) üblich ist, wurde dafür dieses Verfahren speziell weiterentwickelt, das auch mit extrem stark reduzierten Systemkomponenten (Rechenleistungs- und Speicherplatzbedarf im RAM) in der Praxis zurechtkommt. Für einen Klangelement eines Gamelan Orchesters (im lauten Still existieren davon ca. 120 unterschiedliche Klangelemente)

werden nunmehr pro Teilton 6 Parameter (*Frequenz, Startphase, Startamplitude, Steigung und Dauer der Anklingphase, sowie die Ausklingkonstante*) benötigt, wobei pro Klangelement mit 20 bis ca. 40 Teiltönen zu rechnen ist. Die Synthesebank für die deterministischen Komponenten (Teiltöne) ist in Abbildung 8 dargestellt.

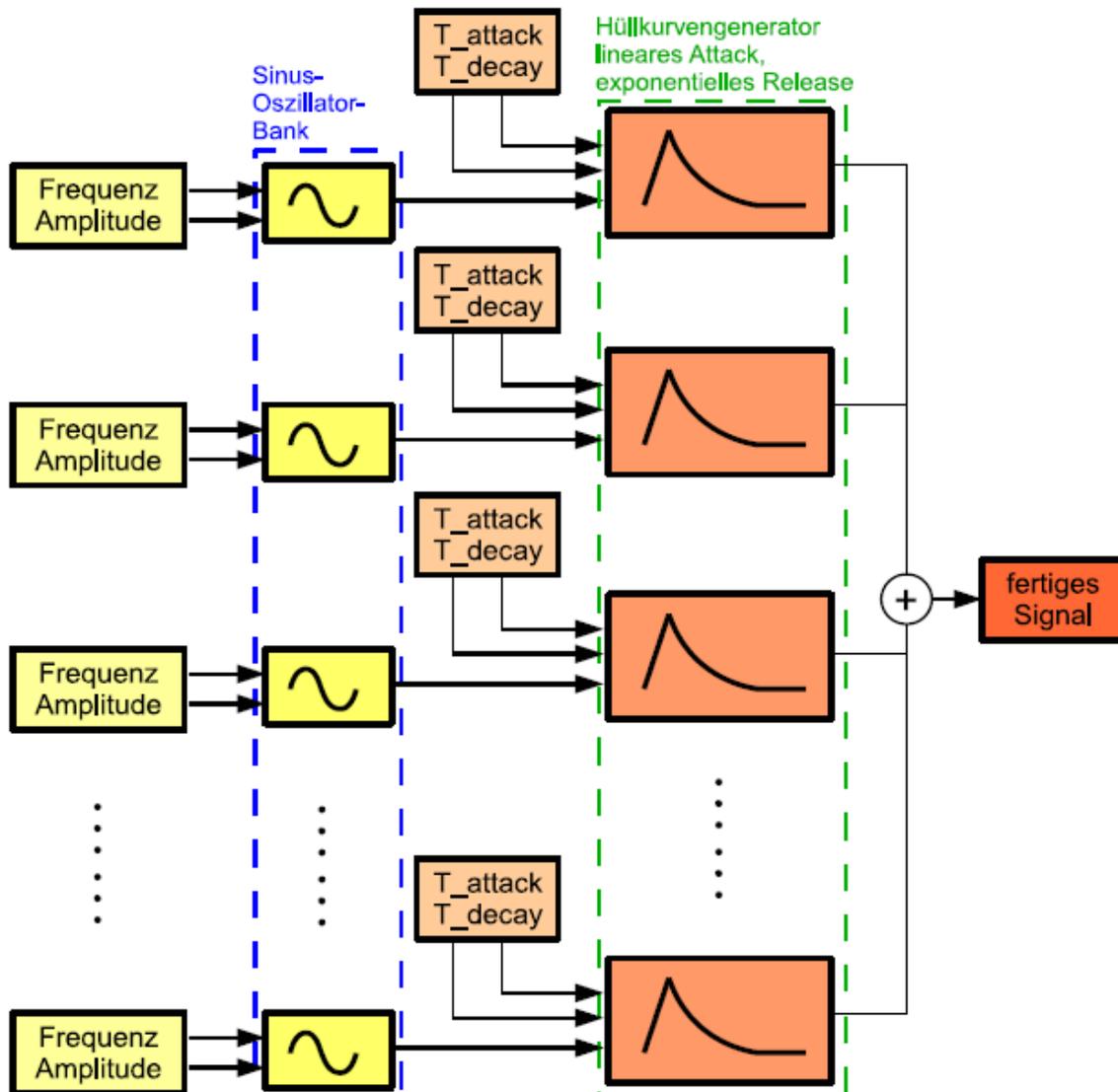


Abb. 8: Blockdiagramm der Teiltongeneration (vgl. [21])

Anhand eines ausgewählten Klanges (Bonang Barung) sollte hier zumindest auf Basis der spektralen ähnlichen Gestalt, die Qualität des beschriebenen Resyntheseverfahrens dokumentiert und dargestellt werden.

Weitere Veröffentlichungen sind mit den neu gewonnenen wissenschaftlichen Kooperationspartnern am CNMAT, Universität von Berkeley, derzeit in Bearbeitung (siehe [17]).

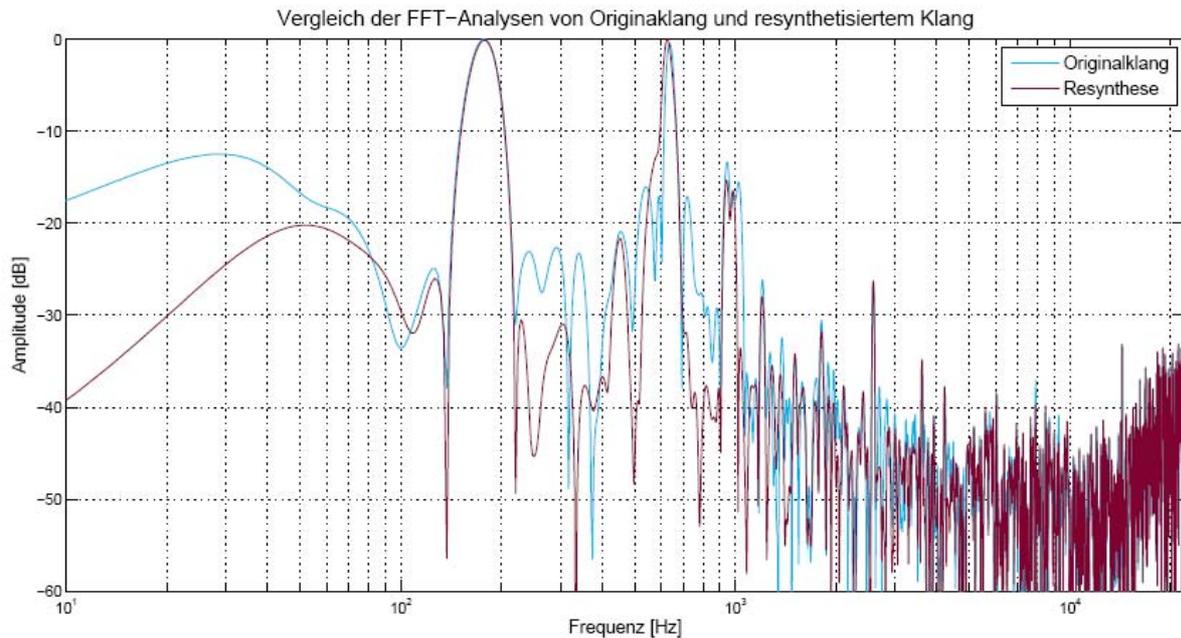


Abb. 9: Vergleich eines Originalklanges des Instruments Bonang Barung mit seinem resynthetisierten Adäquat (vgl. [21]).

Klangmodellierungsverfahren – (A)WSM

Beim „auditory Warped Spectral Model“(A)WSM wird der Gesamtklang durch einzelne Klangspuren beschrieben. Wobei im Gegensatz zum statischen Model von vorher hier die Frequenz und die Lautstärke zeitlich variieren können. Der Einschwingvorgang für den Klang wird hier ebenfalls explizit behandelt.

Zusätzlich wird bei diesem Verfahren in Anlehnung an unsere Hörwahrnehmung die Hörbarkeit von diesen Klangspuren untersucht und ggf. aufgegriffen oder verworfen. Abbildung 10 zeigt ein einfaches mögliches Klangatom als exemplarisches Beispiel für das Zeit- Lautstärke-Diagramm und das Zeit-Frequenz Diagramm. Hier wurde nur der zeitliche Verlauf der Lautstärke verändert und auf die Veränderlichkeit der Frequenz verzichtet.

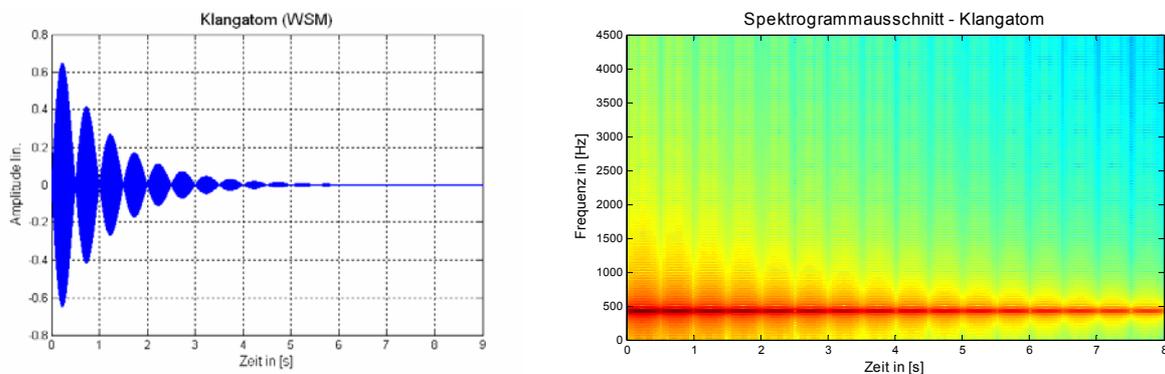


Abb. 10: Zeit- Lautstärke-Diagramm und Zeit- Frequenz-Diagramm für ein Klangatom zugehörig zum WSM Verfahren.

Im entwickelten WSM Verfahren wird die Frequenzauflösung des menschlichen Gehörs bei der Nachbildung von Klängen herangezogen. Dadurch kann die Anzahl der Teiltöne drastisch gesenkt werden. Die Auswertung erfolgt in zwei Richtungen: Teiltonzerlegung und Verfolgung von Amplitude und Frequenz (deterministischer, schmalbandiger Anteil), sowie Auswertung des Restsignals (stochastischer, breitbandiger Anteil).

Aufbauend auf die Arbeit von Härmä (vgl. [5]) wurde am IEM ein Modell entwickelt [26], das sich weitgehend an einer gehörgerechten Analyse orientiert. Das Modell umfasst folgende wesentliche Verarbeitungsschritte:

1. Berechnung der frequenzachsenverzerzte Kurzzeit Fourier Transformation (STFT)
2. Berücksichtigung der Außen-&Mittelohrübertragungsfunktion
3. Berücksichtigung der (simultanen & zeitlichen) Maskierungseigenschaften
4. Berechnung der Mithörschwelle
5. Kriterienabhängige Trajektorien für Teiltönverläufe
- [6.] Transientenextrahierung mittels voltage controlled notch filters (VCNF)

Bei der Bestimmung der gewarpten Eingangsspektren wurden 2 grundlegende Methoden (A und B) untersucht deren Strukturschaltbilder in Abbildung 11 dargestellt sind.

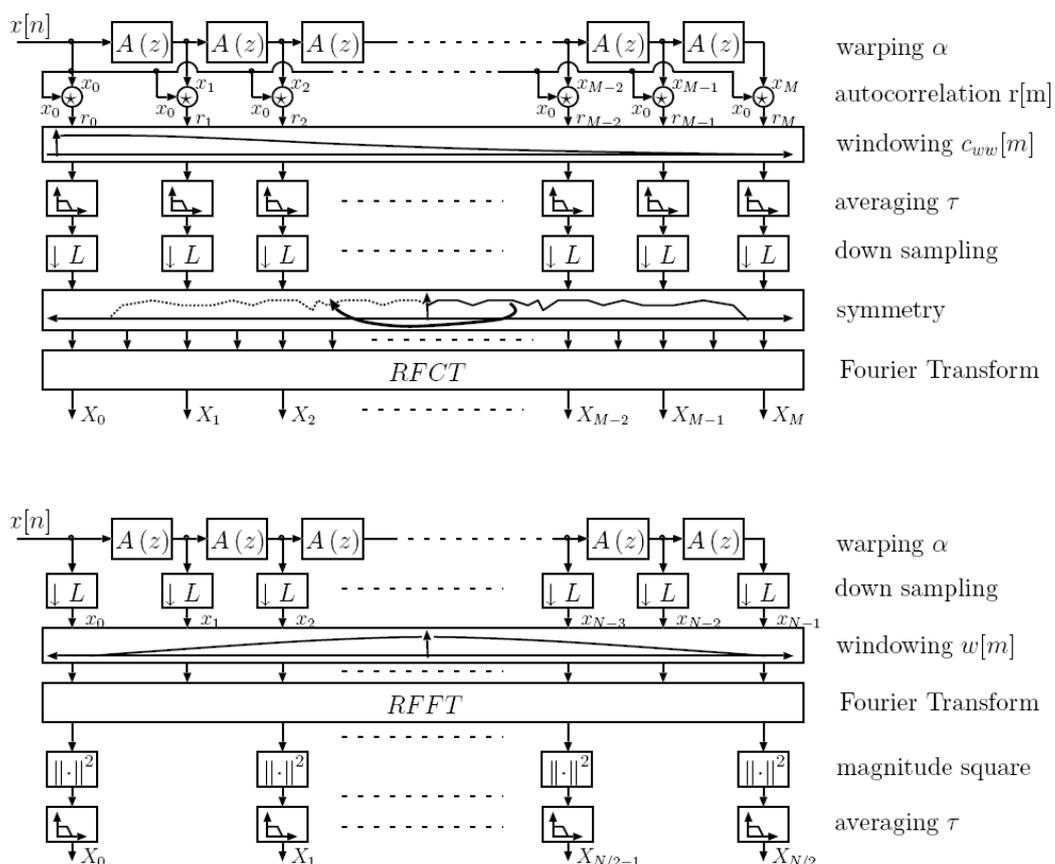


Abb. 11: Strukturschaltbilder für die Signalverarbeitungsschritte
oben: Methode A und unten: Methode B

Methode A) Das Leistungsdichtespektrum wird über die Autokorrelation bestimmt:

Der Vorteil dieser Variante liegt in der nahezu latenzfreien Verfolgung von zeitlichen Signalveränderungen. Jedoch stellt sich dabei dem Algorithmus innewohnenden Nachteil einer geringen Dynamik des detektierten Spektrums ein, d.h. leise jedoch hörbare Signalkomponenten werden bereits vorzeitig eliminiert.

Methode B) Leistungsdichtespektrum wird über das Betragsquadrat der warped-STFT-Spektren gebildet:

Bei dieser Variante steht zwar der gewünschte große Dynamikbereich weiterhin zur Verfügung, jedoch ergibt sich durch die zeitliche Blockung der STFT eine inhärente Zeitauflösung.

Die Berücksichtigung der Simultanmaskierung erfolgt nach der Methode von Zwicker [6] und jene der zeitlichen Maskierung nach dem Modell von Baumgarte [7].

Bei der Aufzeichnen der Teiltonspuren bzw. deren zeitliche Verfolgung erfolgte eingangs mit Hilfe von Übergangsbewertungsmatrizen, welche die Attribute wie Lautstärke der einzelnen Teiltöne, Alter einer Trajektorie, Frequenzänderung eines Teiltöns usw. für jede einzelne Spur iterativ bewerten. In [16] wurde eine verbesserte Abschätzung des Trajektorienverlaufs basierend auf Frequenz- und Zeitreassignment (mit Erweiterung der Terme 2.Ordnung) vorgestellt. Anhand dieses neuen Ansatzes wird die Abschätzung des Verlaufes vereinfacht (vgl. Abb. 12).

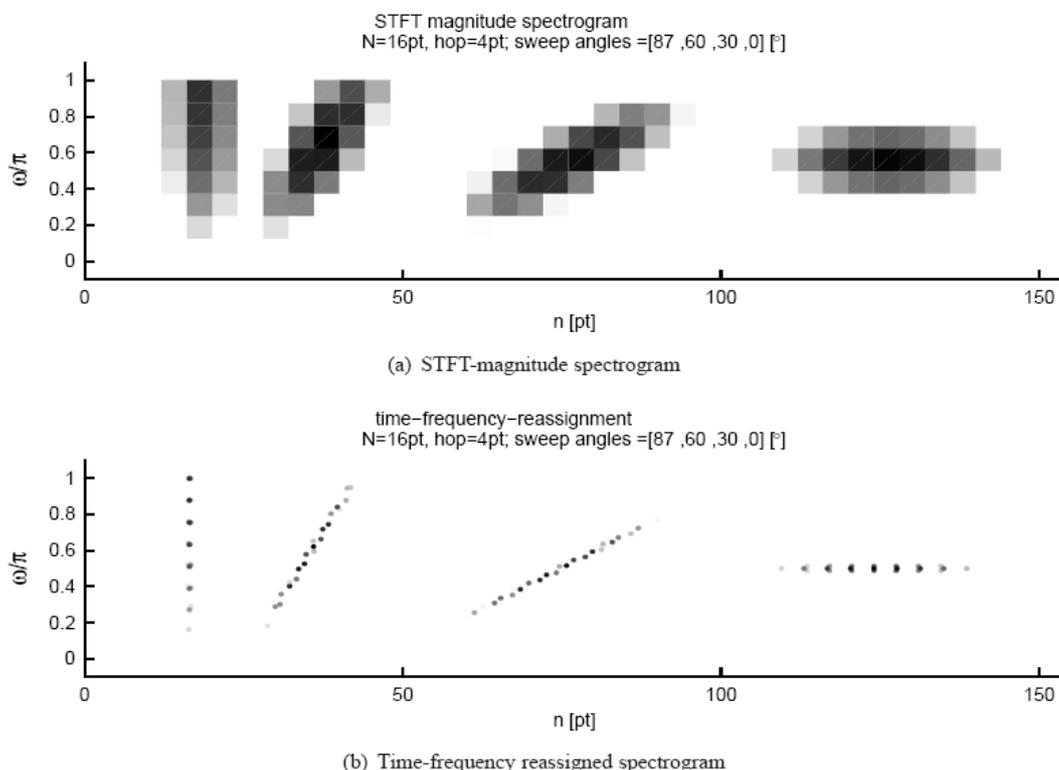


Abb. 12: Verbesserter Darstellungsqualität für unterschiedliche Signalformen (vgl. [17]).
oben: herkömmliche Kurzzeitspektren (Fourier Transformation) **unten:** Reassigned Spektrum.

Kommt zur Berechnung der gewarpten Eingangsspektren die Berechnungsmethode B zum Einsatz so werden aufgrund der geringen zeitlichen Auflösung - die vor allem am Beginn der Klänge von Bedeutung ist - für den transienten Anschlag Residuen bestimmt. 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.

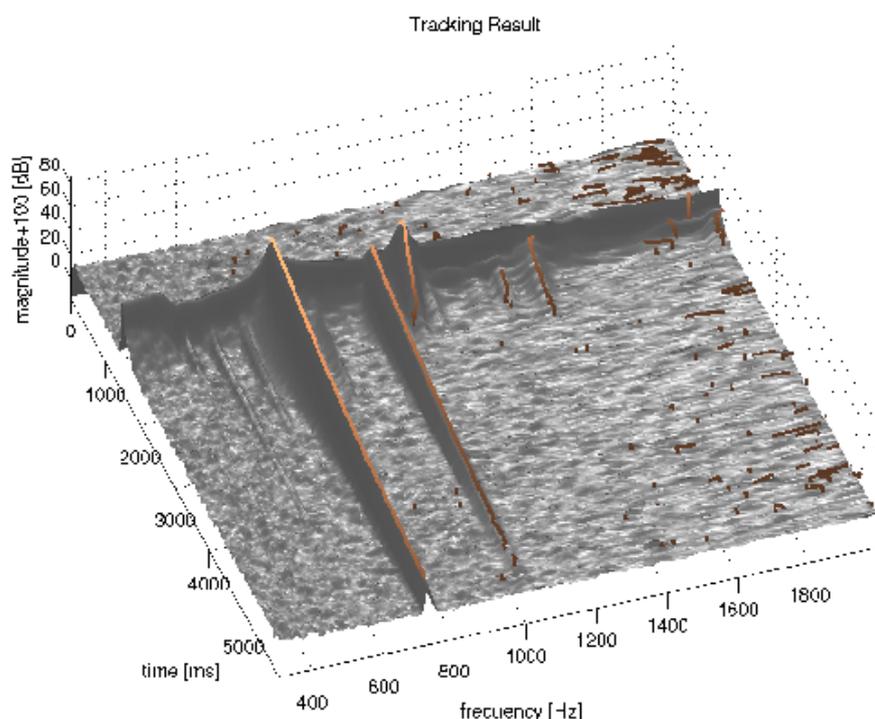


Abb. 13: Vergleich der verfolgten (getrackten) Teiltöne (eingefärbte Linien) und Wasserfall-Spektrogrammdarstellung.

Klangmodellierungsverfahren - EMD

Das Empirical Mode Decomposition (EMD) Verfahren erlaubt eine Selbstorganisation der Analyse in Abhängigkeit von der Signalform des Klages selbst. Es gibt hier jedoch keine a priori Beschreibung eines möglichen Klangatoms. In ungünstigen Fällen kann die Darstellung und Modellierung erst durch eine größere Datenmenge als die ursprüngliche erfolgen.

Die EMD Methode zerlegt das zu untersuchende Signal in so genannte Intrinsische Mode Funktionen (IMF). Für die Berechnung der tatsächlich vorhandenen Spektralkomponenten werden aus den reellen IMFs analytische Signale gebildet und durch Ableitung der Phase nach der Zeit die Momentanfrequenzen errechnet

(Hilbert-Huang-Spektrum). Die Problematik der begrenzten Frequenzauflösung wird durch diesen Ansatz scheinbar gelöst.

Die EMD Methode wurde bereits in den 90iger Jahren von N.E. Huang et al. [8] entwickelt. Gegenüber der *Fourier Transformation* muss das Analysesignal weder lineare noch stationäre Eigenschaften besitzen.

Es gilt jedoch zu beachten, dass für das Auffinden der Momentanfrequenz der IMF kein Gleichanteil bzw. Trend überlagert ist (mittelwertfrei) und weiters muss der Signalrauschabstand (d.h. es muss tatsächlich ein Mode vorliegen) entsprechend groß sein.

Der Siebungsprozess zur Bestimmung der IMF kann wie folgt beschrieben werden:

1. Lokalisierung aller Extremwerte (Maxima und Minima) des Analysesignals $x(t)$
2. Interpolation aller Maxima respektive Minima. Daraus erhält man zwei das Eingangssignal einhüllende Kurvenverläufe ($e_{\max}(t)$ und $e_{\min}(t)$)

3. Bildung des Mittelwerts $m(t) = \frac{e_{\min}(t) + e_{\max}(t)}{2}$ Gl. (4)

4. Subtraktion des lokalen Mittelwerts (aus 3) vom ursprünglichen Signal:

$$h_1(t) = x(t) - m(t) \quad \text{Gl. (5)}$$

um lokale Mittelwertfreiheit zu erlangen.

5. Überprüfung, ob $h_1(t)$ die Bedingungen einer IMF (mittelwertfrei und #Extremwerte gleich #Nulldurchgänge ± 1) erfüllt.

Werden die Kriterien nicht erfüllt, so müssen die Schritte 1 – 4 wiederholt werden. Stellt sich heraus, dass $h_1(t)$ eine IMF ist, so wird diese vom Originalsignal $x(t)$ subtrahiert und man erhält ein Residuum, welches nun als neues „Analysesignal“ verwendet und sämtliche Schritte 1-5 wiederholt werden.

Aufgrund dieser Vorgangsweise erhält man eine hierarchische *Top-Down* Analyse: zuerst werden die hochfrequenten und zum Schluss die tieffrequenten Signalkomponenten extrahiert.

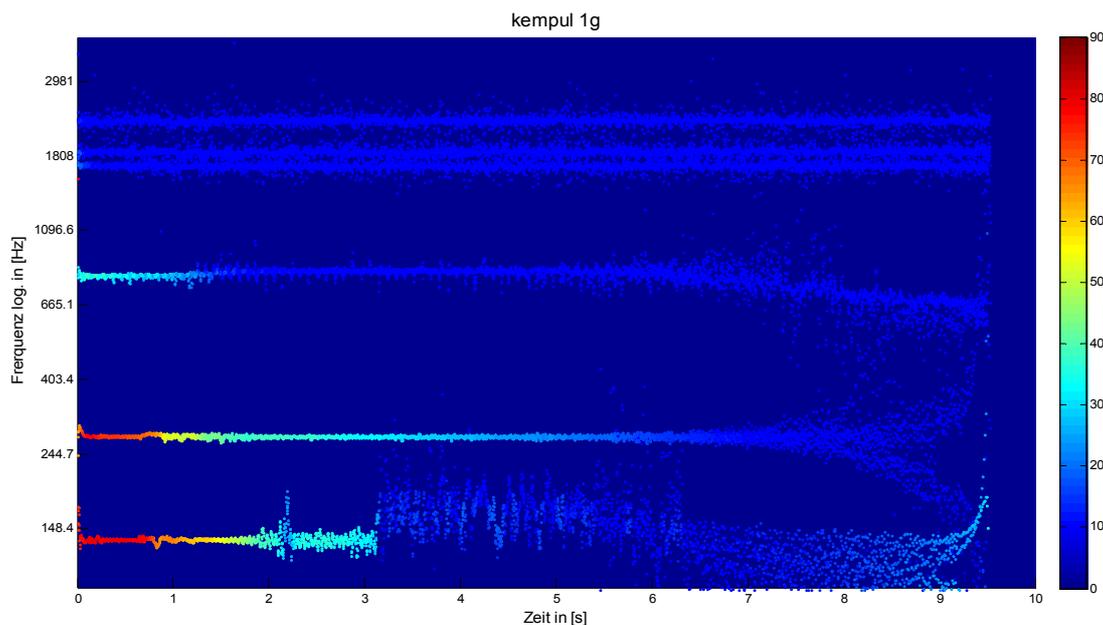


Abb. 13: Analyseergebnis des mod. EMD Verfahrens anhand des Klanges „kempul 1g“ („glides“ bei ca. 1sek).

Im Allg. ist der Fehler bei den extrahierten IMFs dann größer, je näher deren Frequenz bei Hälfte der Abtastrate liegt. Stimmt die Frequenz exakt mit einem ganzzahligen Verhältnis der Abtastrate überein 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 tiefer-frequenten Signalkomponente detektiert, solange der Frequenzabstand wiederum nicht zu groß (abh. von den Amplitudenverhältnissen) wird.

Im Rahmen des VGG Projektes wurde der bestehende EMD Algorithmus für die Gamelan Instrumentalklänge verbessert. Der modifizierte EMD-Ansatz geht davon aus, dass die relevanten Frequenzkomponenten weitgehend 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. Befinden sich in einer IMF tatsächlich nur eine Signalkomponente, dann kann anschließend die Momentanfrequenz und der Amplitudenverlauf ohne Verlust von Information unterabgetastet (datenreduziert) werden.

Wie in Abbildung 13 ersichtlich wurden mit dem modifizierten EMD Verfahren tatsächlich auch beeindruckende Ergebnisse realisiert. Bedingt durch die oben geschilderten Tatsachen und der erforderlichen Anpassungen an den Signaltypus wurde dieses Verfahren jedoch verworfen.

3 Aufnahme und Wiedergabe der akustischen Abstrahlung

Dieser Teil in Virtual Gamelan Graz befasst sich mit dem Entwurf eines umfassenden Ansatzes zur Klangerfassung und Klangabgabe. In diesem Zusammenhang werden vor allem zwei Kernfragen dazu beantwortet:

Wie können Mikrofone auf eine Weise angeordnet werden, die das räumliche Klangbild eines Instrumentes möglichst vollständig aufnehmen können? Und wie sieht eine geeignete Lautsprecheranordnung und ein geeignetes Wiedergabeverfahren Wiedergabe aus?

Eine rein mathematische, aber vielseitige Beschreibung der Klänge im Raum, kann in der akustischen Wellengleichung gefunden werden. Eine genaue Kenntnis dieser Lösungen gibt auch eine Antwort auf diese Frage, und lässt sich leicht über das vorgeschlagene Seifenblasenmodell erklären.

Für ein kompakteres Klangmaterial wurden zudem auch einfache Aufnahmen der Klänge gemacht, die insbesondere für die Klangzerlegungen und Modifikationen, sowie als Sample-Datenbank wertvolles Material sind.

3.1 Hintergrund

Bisweilen wird zumeist ohne zu hinterfragen hingenommen, dass Klänge von Musikinstrumenten mit einem Mikrophon und einem Lautsprecher aufgenommen- und wiedergegeben werden können. Natürlich ist dieser Ansatz nicht gänzlich falsch, besitzt aber bei weitem die Möglichkeit, die Klangabgabe des Instrumentes, und damit die von der Richtung abhängige Klangfarbe zu beschreiben. Ebenso bietet die Wiedergabe einem einfachen Lautsprecher keinerlei Unterstützung dieser Eigenschaften. Sämtliche richtungsabhängige Klangfärbung lässt sich einzig und alleine auf den jeweiligen Lautsprecher zurück führen, und ist somit in den meisten Fällen stark abweichend. In der Akustik wird diese Eigenschaft of als *akustische Abstrahlungswirkung* oder *Abstrahl-Charakteristik* bezeichnet.

Schon seit längerem gibt es, insbesondere im Bereich der elektronischen Musik, die Absicht, Klänge von natürlichen Instrumenten mit Verfahren aufnehmen und wiedergeben zu können, die dazu in der Lage sind, den gesamten Instrumentalklang rund um ein Instrument erfassen, und wiedergeben zu können. Zudem besteht das Interesse diese Verfahren auch zur Verfremdung und als künstlerischen Effekt einzusetzen. Allerdings haben bislang nur äußerst wenige Institutionen diese nun schon fast 30 Jahre alte Absicht auch wissenschaftlich erforscht.

VGG hat es ermöglicht, diese Forschungsarbeit mit der hoch interessanten Erforschung von Gamelan-Instrumenten zu verknüpfen. Dabei ist die Zielsetzung entstanden, mit geeigneten Verfahren ein Gamelan-Ensemble akustisch nachzubilden. Besonders aufregend dabei ist, die räumlich differenziert angeordneten Schwebungen zwischen den Instrumenten nachzubilden. Im Zuge der ausgedehnten Untersuchungen im Projekt VGG hat sich eine geeignete Herangehensweise zur Lösung der Fragestellungen herauskristallisiert. Sie basiert auf einer sowohl höchst mathematischen, aber auch abstrakten wissenschaftlichen Herangehensweise:

Das Seifenblasenmodell der akustischen Abstrahlungs:

Dieses abstrakte Modell hat nicht nur bei der tatsächlichen Umsetzung der Zielsetzungen geholfen, es hilft auch bei folgender Erläuterung unseres Lösungsweges. Als Vorstellungshilfe für die Schallausbreitung wird angenommen, wir könnten eine Seifenblase um ein Instrument und seine(n) MusikerIn herum aufblasen. Erklängt das Instrument, so schwingt die Oberfläche dieser Seifenblase in jedem Punkt im Gleichklang mit dem in diese Richtung abgegebenen Luftschall. Die Stärke der Schwingung zeigt dabei an, wie laut der Klang des Instrumentes in die jeweilige Richtung strahlt. Des Weiteren gibt die Schwingungsform Auskunft darüber, wie der Klang zusammengesetzt, also welche Klangfarbe er in unterschiedlichen Richtungen besitzt.

Um einen Gamelan-Klang nun auch vollständig im räumlichen Sinne beschreiben zu können müsste die diese Seifenblasenschwingung für jeder Stelle an der Blase bestimmt, also mit unzähligen Mikrofonen aufgenommen werden. Sofern bestimmte Annahmen gelten, reichen auch weniger Mikrofone aus. Ein besonders wertvolles Hilfsmittel dazu sind die Schwingungsmoden einer Seifenblase. Erst diese ermöglichen eine Idee einer begrenzten Auflösung. Das erleichtert, ja ermöglicht in vielerlei Hinsicht erst, die Anforderungen für ein geeignetes Aufnahme- und Wiedergabeverfahren zu beschreiben. ²

² *Anm: Das schöne an diesen Schwingungsformen ist, dass sie gleichzeitig auch "Lösungen der akustischen Wellengleichung" sind, somit unmittelbar die akustische Schallausbreitung mit beschreiben können.*

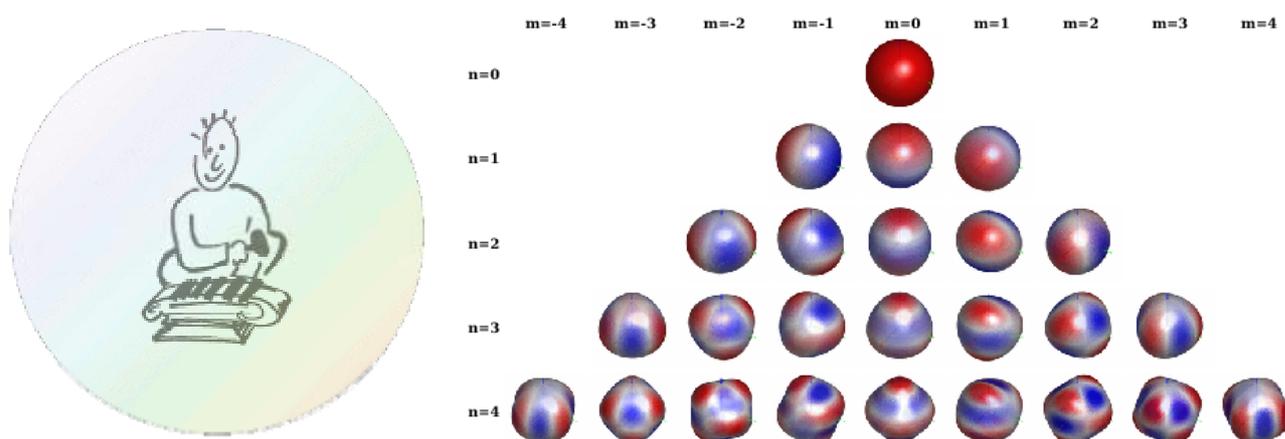


Abb.14: links: Seifenblasenmodell, und rechts: Schwingungsmoden einer Seifenblase / Kugel bis zum 4. Grad (*Kugelflächenfunktionen*).

3.2 Aufnahme der akustischen Abstrahlung

Zunächst wurde für die Aufnahme der Abstrahlung die benötigte Infrastruktur konzipiert und beschafft. Dabei wurde am IEM eine akustisch bedämpfte Messkoje geplant und eingerichtet, und Ausrüstung zur gleichzeitigen Erfassung von vielen Mikrofonsignalen gekauft.

Zur Aufnahme selbst wurden zweierlei Anordnungen von Mikrofonen entworfen und zwei Verfahren mit unterschiedlichem Einsatzbereich im Hinblick auf die Gamelan-Instrumente gefunden:

- **10 Mikrofone im Halbkreis für hängende, rotationssymmetrische Instrumente,**
- **26 Mikrofone in einer Halbkugel für Instrumente, die am Boden stehen.**

Für den zweiten Fall wurden die Mikrofone besonders sorgfältig verteilt, da es streng genommen keine regelmäßig angeordneten Punkte auf der Halbkugel gibt, welche die Anzahl 10 überschreiten können [10].

Bereits im Sommer 2006 konnten mit beiden Verfahren und unter der Mithilfe des Instituts für Musikethnologie Klängaufnahmen von mehr als 100 unterschiedlicher Gamelan-Klängen zu je 26, bzw. 10 gleichzeitigen "Tonspuren" am IEM aufgezeichnet werden.

Des Weiteren wurde bereits im Frühjahr 2006 einfache einkanalige Aufnahme der Klänge durchgeführt, die beinahe das vollständige Instrumentarium von Gamelan-Instrumenten an der Kunstuniversität abdecken. Auch dieses Mal mit großer

Unterstützung des Institutes für Musikethnologie (IME). Die Vervollständigung des einfach aufgezeichneten Klangmaterials erfolgte durch das IME.

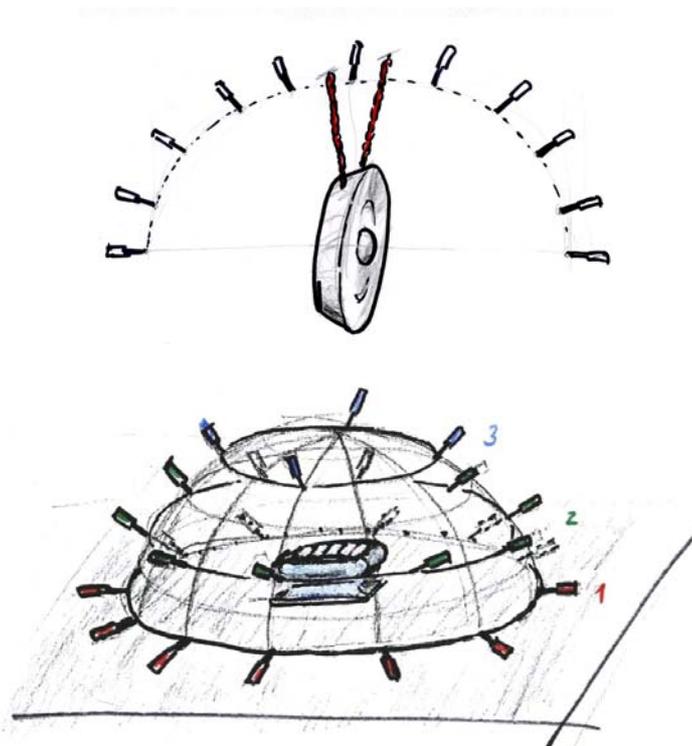


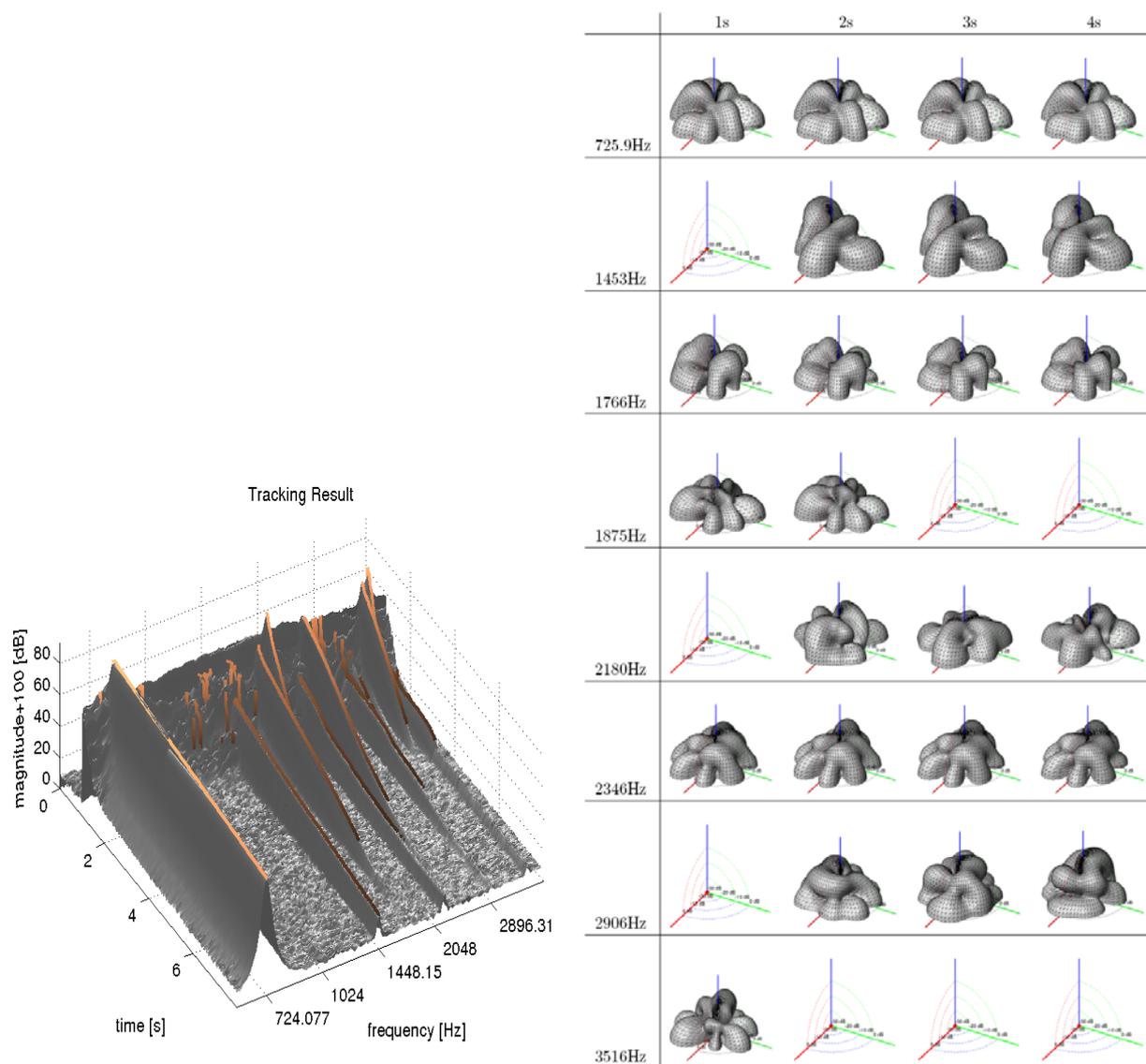
Abb. 15: Aufnahmeanordnung mit 10 Mikrofonen für Gongs (*kempul, gong suwukan+gong ageng*) 26 Mikrofonen in 3 Ringen zu 12, 10 und 4 Mikrofonen für die gleichmäßige Erfassung der akustischen Abstrahlung von am Boden stehenden Gamelan-Instrumenten.

3.3 Analyse der Abstrahlungsmuster von Teiltönen

Eines der Klangzerlegungsverfahren aus den vorigen Abschnitten wurde nun ganz spezifisch dahingehend weiter entwickelt, auch auf den Fall mit mehreren Mikrofonen anwendbar zu sein. Das beinhaltet ein Konzept des *abgestrahlten Gesamtklanges* eines Instrumentes, das in der folgenden Abbildung 16 gezeigt wird.

So gewonnene, neuartige Diagramme können für jeden Klang nun zeigen, ob die Abstrahlung der Teiltöne sich im Laufe des Ausklingens verändert. Im dargestellten Beispiel ist dies im Wesentlichen für keinen Teilton der Fall³. Es ist auch wunderbar möglich deutlich zu machen, warum dieses Instrument nicht in jede Richtung gleich klingt. So strahlt etwa der Grundton des Instrumentes wesentlich stärker in der Horizontalen ab, als nach oben. Daher ist bei diesem Ton ein sehr voluminöser Klang auf gleicher Höhe zum Instrument wahrnehmbar. Nach schräg oben ergeben sich gänzlich verschiedene Klangeigenschaften, in dem die höheren Teiltöne eine vorherrschende Rolle einnehmen.

³ Dazu sei gesagt, dass sich die Diagramme der sehr leisen Teiltöne sehr wohl ändern, es handelt sich hierbei aber um Messungenauigkeiten aufgrund der Umgebungsgeräusche.

**Abb. 16:**

(links) Abgestrahlter Gesamtklang des Instrumentes *saron barung*, Ton Nummer 3, *patet slendro*, lauter Anschlag. Dargestellt ist die Frequenz als x-Achse, die Zeit als y-Achse (aus dem Bild heraus), und die Stärke der jeweiligen Klangkomponente in der Höhe auf der z-Achse. Die graue Fläche beschreibt den Gesamtklang, die Kupferfarbenen Linien sind die extrahierten Klangbestandteile.

(rechts) Abstrahlungsmuster aller relevanten Klangkomponenten (vgl. Diagramm links), wobei die dargestellte Fläche ein so genannter *3D-Richtplot* ist. Darin ist die Abstrahlungslautstärke für jede Richtung als Abstand vom Zentrum aufgetragen. (Vorstellungshilfe: Eine das Instrument umgebende lautstärke-empfindliche Kaugummi-Blase, die sich an lauten Stellen nach außen dehnt).

3.4 Wiedergabe der Akustischen Abstrahlung

Ziel ist nun natürlich auch, den Klang an jener Stelle wieder herstellen zu können, an der er aufgezeichnet wurde. Plakativ gesprochen, wo vorher die Schwingung im Seifenblasenmodell mit Mikrofonen aufgenommen wurde, soll nun dieselbe Schwingung wieder hergestellt werden, um damit die aufgezeichneten Instrumentalklänge wieder vollständig herzustellen.

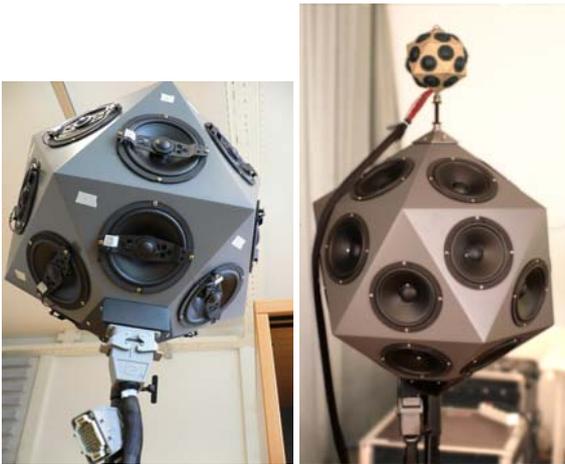


Abb. 17: IEM - Iksaeder Lautsprecher

Für diese herausfordernde Aufgabenstellung wurde zu erst ein ikosaederförmiger Wiedergabelautsprecher mit einer besonderen Koaxial-Lautsprecherkonstellation angefertigt⁴ für alle Arten von Kugelwellen (siehe Abbildung 14) mit je 20 Lautsprechern entworfen und gebaut. Im Zuge der Arbeiten an unserer *mathematischen* Beschreibung dieses Gerätes sind auch unbekannte Grenzen sichtbar gemacht worden; als funktionierende Lösung konnte darauf hin mit einem zusätzlichen zweiten, kleineren System reagiert werden.

3.5 Umsetzung und Optimierung des Wiedergabeverfahrens

Das Verfahren des *radialen Beamforming* konnte zur Steuerung der Wiedergabe (Synthese) der akustischen Abstrahlungsmuster entwickelt werden. Es ermöglicht, nicht nur auf der Oberfläche des Lautsprechersystems Muster im Luftschall zu erzeugen, sondern dieses Muster auf eine größere, den Lautsprecher umgebende, Kugel mit frei wählbarem Abstand zu projizieren.

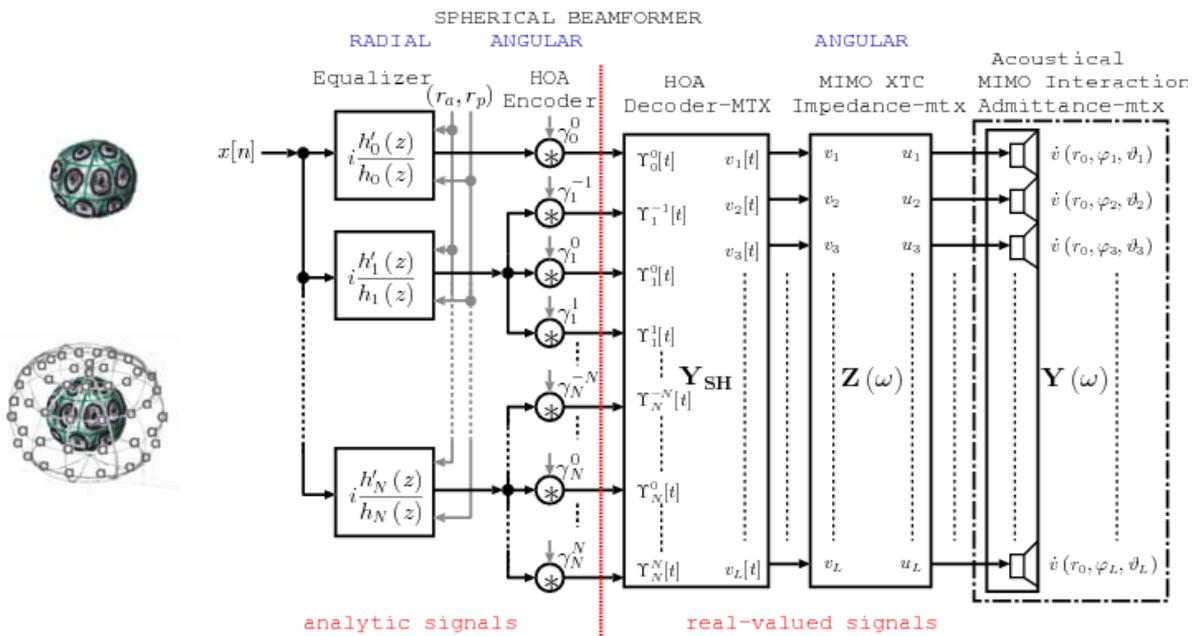


Abb. 18:

(links) Konzept der Projektion des Synthese-Musters auf eine größere Kugel.

(rechts) Blockdiagramm der Umsetzung dieses Verfahrens. Dabei werden im Signalweg links der roten Linie die Teiltöne mit variablen räumlichen Verteilungen (Parameter r_p und γ) zugespielt. Die rechte Seite des Bildes wird dem Lautsprechersystem zugeordnet. Dabei ist Y_{SH} die Zerlegung der Kugelmoden auf die Lautsprecher und Z eine Kompensation des Lautsprecherinneren [19][20]. Y ist das physikalische Modell des Lautsprechersystems.

⁴ Herzlicher Dank gilt der Firma ITEC-Audio in Gleisdorf, die uns großzügig mit ihrem Produkt ausgestattet hat.

Die dabei zum Einsatz kommenden hoch effizienten Signalfilter sind besonders auf die Anwendung hin optimiert worden, um Rechenleistung zu sparen, und konnten bereits in der Arbeit [15] veröffentlicht werden.

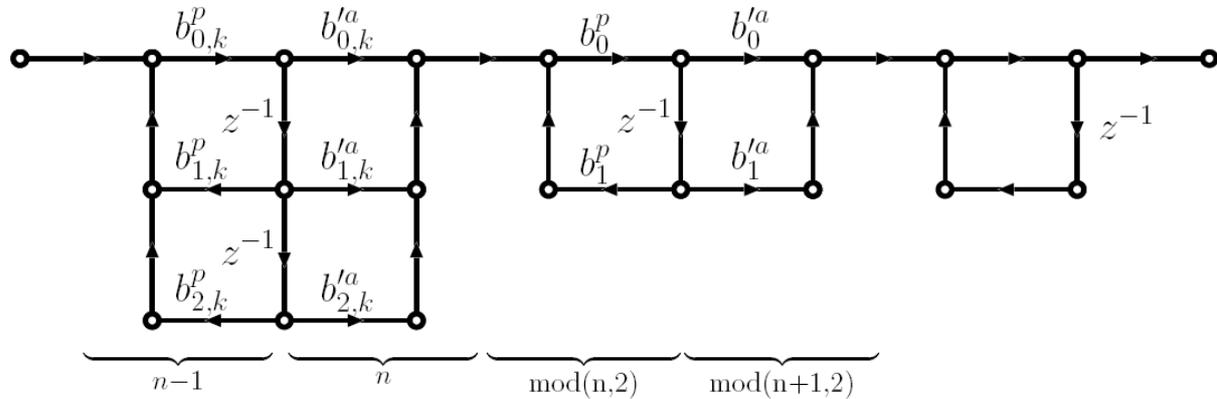


Abb. 19: Effiziente Filterimplementierung für den Radial-Beamformer.

Literatur

- [1] Gerhard Nierhaus, „Zur Modellierung von Kernmelodien im Rahmen des Virtual Gamelan Graz (VGG)“, VGG Endbericht IEM Report 41/07, 2007.
- [2] Gerd Grupe, Rainer Schütz, „Abschlussbericht zum Projekt „Virtual Gamelan Graz (VGG)“ IME VGG Endbericht, 2007.
- [3] Matti Karjalainen et al., „AR/ARMA Analysis and Modeling of Modes in Resonant and Reverberant Systems“, 112th AES Convention, Convention Paper 5590, May 10-13 Munich, Germany, 2002.
- [4] S. Lawrence Marple, Jr., “Digital Spectral Analysis with Applications”, Prentice Hall, Signal Processing Series, Alan.V. Oppenheim, Series Editor, Prentice-Hall Inc. Englewood Cliffs, New Jersey 07632, 1982. ISBN 0-13-214149-3 025.
- [5] A. Härmä, “Frequency-warped autoregressive modeling and filtering,” PhD-Thesis, Helsinki University of Technology, P.O. BOX 1000, FIN-02015 HUT, 2001.
- [6] Zwicker, E., and Fastl, H. (1990). “Psychoacoustics - Facts and models,” Springer-Verlag.
- [7] Frank Baumgarte, „Ein psychophysiologisches Gehörmodell zur Nachbildung von Wahrnehmungsschwellen für die Audiocodierung“, Dissertation, Fachbereich Elektrotechnik und Informationstechnik, Universität Hannover, 2000.
- [8] 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.

- [9] Alois Sontacchi, Franz Zotter: IEM-Report 35/06: "Virtual Gamelan Graz". 2006.
- [10] Franz Zotter, Alois Sontacchi, Markus Noisternig und Robert Höldrich: "Capturing the Radiation Characteristics of the Bonang Barung". 3rd Congress of the Alps Adria Acoustics Association, Graz, Sept. 2007.
- [11] Franz Zotter, Alois Sontacchi: IEM-Report 39/07: "Icosahedral Loudspeaker Array". Jan. 2007.
- [12] Franz Zotter, Alois Sontacchi, Robert Höldrich: "Modeling Spherical a Loudspeaker Systems as Multipole Source". Fortschritte der Akustik, DAGA Stuttgart, March 2007.
- [13] Franz Zotter, Robert Höldrich: "Modeling Radiation Synthesis with Spherical Loudspeaker Arrays". International Congress on Acoustics, Madrid, Sept. 2007.
- [14] Alois Sontacchi, Franz Zotter, Robert Höldrich: "Klangmodellierungsansätze und Ergebnisse für Gamelaninstrumente des lauten Stils". Fortschritte der Akustik, DAGA, Stuttgart, March 2007.
- [15] Franz Zotter, Markus Noisternig: "Near- and Far-Field Beamforming Using Spherical Loudspeaker Arrays". 3rd Congress of the Alps Adria Acoustics Association, Graz, Sept. 2007.
- [16] Franz Zotter, Robert Höldrich: "Chirp-Rate Estimation for Time-Frequency Reassigned Spectrograms". 3rd Congress of the Alps Adria Acoustics Association, Graz, Sept. 2007.

Veröffentlichungen in Vorbereitung:

- [17] Franz Zotter, Adrian Freed, Alois Sontacchi: "Auditory Resonance Model of Percussive Sounds". Fortschritte der Akustik, DAGA, Dresden, März. 2008.
- [18] Franz Zotter, Andrew Schmeder, Markus Noisternig: "Crosstalk Cancellation for Spherical Loudspeaker Arrays". Fortschritte der Akustik, DAGA, Dresden, März. 2008.

Unsere besondere Anerkennung gilt jenen StudentInnen, die sich im Rahmen ihrer wissenschaftlichen Arbeiten für das Projekt VGG begeistern konnten:**abgeschlossene Arbeiten:**

- [19] Peter Reiner, Christian Jochum: Projekt ToningenieurIn: "Erfassung und Evaluierung der Übersprecheigenschaften des Ikosaederlautsprechers". IEM, KUG, 2007.
- [20] Christian Jochum, Peter Reiner: Projekt ToningenieurIn: "Ansteuerungsfilter für den Ikosaederfilter". IEM, KUG, 2007.
- [21] Margherita Jammer: Projekt ToningenieurIn: "Klangmodellierung von Gamelaninstrumenten basierend auf exponentiell ausschwingenden Sinuskomponenten". IEM, KUG, 2007.

laufende Arbeiten:

- [22] Fabian Hohl: Projekt ToningenieurIn: "Klang- und Abstrahlungsanalyse von mehrkanaligen Gamelan Klangsamples". IEM, KUG, 2007.

[23] Hannes Pomberger: Diplomarbeit: "Nah- und Fernfeldbeamforming für den Ikosaederlautsprecher mit Absolutbetragsoptimierung". IEM, KUG, 2007/2008.

[24] Nino Skiljic: Diplomarbeit: "Sound Source Directivity in Binaural Synthesis and Room Simulation". IEM, KUG, 2007/2008.

[25] Daniel Hofer: Diplomarbeit: "Implementierung eines planaren Mikrofonarrays für akustische Nachfeldholographie". IEM, KUG, 2008.

[26] Franz Zotter, Dissertation: "Modeling Sound and Acoustic Radiation of Selected Gamelan Instruments". IEM, KUG, 2008.

Anhang

Open Sound Control (OSC)

Open Sound Control ist eine allgemeine Art der Steuerung von Sound. Es wird hauptsächlich für realtime processing von Sound über Netze und Multimediainstallationen angewendet. Die Steuersignale können von Hard- (z. B. Midikeyboard) oder Software (SuperCollider, Pure Data, u.a.) erzeugt werden und werden via OSC an eine Schnittstelle weitergegeben, die dann eine Ausgabe steuern. Dieses können weitere Soundausgaben sein, z. B. eine Soundanwendung auf einem anderen Computer. Die Steuersignale werden über UDP als Protokoll übertragen. Open Sound Control stammt von der University of California in Berkeley. OSC ist ein transportunabhängiges Protokoll, d.h. es kann über unterschiedliche Kommunikationsschnittstellen versandt werden (z.B. USB, Firewire, Ethernet, Fast Ethernet...). Der Datendurchsatz liegt im 10+ MegaBit/Sekunde Bereich und ist im Vergleich mit Midi dadurch um einen Faktor 300 höher. Numerische Daten können in 32 oder 64 Bit Einheiten kodiert werden, wobei auch symbolische Adressierung und Zeitstempel Mitteilungen möglich sind. Der Versand von Einheiten/Daten mittels OSC erfolgt in Paketen (Datagramm). Jene Applikation (Modul) die OSC Datenpakete versendet wird OSC Client genannt, jene die Pakete empfängt OSC Server. Diese paketbasierende Sendung bietet Synchronisierungsmöglichkeiten. Sämtliche Mitteilungen die innerhalb eines Paketes versandt werden, werden auch zum gleichen Zeitpunkt aktiv. Zeitstempel Mitteilungen sog. „Time tags“ eliminieren auch Jitter, die durch den Transport verursacht werden können. Sämtliche Verbindungen werden bei der VGG-Realisierung, falls erforderlich, direkt erfolgen und somit weitere Überlegungen dazu obsolet. Allg. kann durch die Verwendung einer Wortbreite von 64Bit eine zeitliche Auflösung von bis zu 200 Pico-Sekunden erreicht werden. Open Sound Control ermöglicht eine offene, dynamische, URL-style symbolische Benennung für Mitteilungen und Adressierung.

Weitere Details finden sich unter: <http://www.cnmat.berkeley.edu/OSC>