



Die artefaktfreie Reproduktion spektraler Dynamik bei virtuellen Holzblasinstrumenten

Eine Bachelorarbeit im Studiengang Audiovisuelle Medien

vorgelegt von Christoph Hart am 31.07.2010

Matrikelnummer: 18311

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Prof. Uwe Schulz

Erklärung

Hiermit versichere ich, dass ich die vorliegende Bachelorarbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ort und Datum

Unterschrift

Abstract

Aufgrund der geringen Preise für Massenspeicher und der klanglichen Authentizität realer Instrumentenaufnahmen hat sich die Samplingtechnologie als wichtigstes Verfahren der virtuellen Nachbildung realer Musikinstrumente durchsetzen können. Bei der Anwendung des Verfahrens auf Blas- und Streichinstrumente treten bei der Überblendung verschiedener Dynamikstufen jedoch Kammfiltereffekte auf, die eine Echtzeitkontrolle der Instrumente erschweren. In dieser Arbeit sollen anhand der Reproduktion einer Klarinette verschiedene Lösungsansätze untersucht werden, um zwischen den unterschiedlichen Obertonstrukturen verschiedener Dynamikstufen in Echtzeit und ohne Klangartefakte überblenden zu können. Ein Ansatz ist die Verwendung des AET-Filters der Firma Native-Instruments, der seit der aktuellen Version der Sampling-Software KONTAKT verfügbar ist. Parallel dazu soll ein alternatives Modell entwickelt werden, welches Sampling mit einer speziellen Form der Wavetablesynthese kombiniert, um die verschiedenen Klangbestandteile separat nachzubilden.

Inhaltsverzeichnis

Vorbemerkungen	5
Abkürzungsverzeichnis	5
Einleitung	6
TEIL A: Vorbetrachtungen	7
A 1 – Alternative Technologien.....	7
A 1.1 – Physical Modeling.....	7
A 1.2 – Sampling.....	8
A 1.2.1 – Artefaktbildung bei der Interferenz von Schwingungen.....	10
A 2 – Erläuterungen der verwendeten Technologien.....	12
A 2.1 – Das MIDI-Protokoll.....	12
A 2.1.1 – Wichtige Controllernummern.....	14
A 2.1.2 – Breath Controller.....	14
A 2.2 – Native Instruments KONTAKT.....	15
A 2.2.1 – Skriptsprache.....	17
TEIL B: Anwendung	18
B 0 – Klangliche Analyse der Klarinette.....	18
B 1 – Erster Ansatz: Sampling + AET-Filter.....	19
B 1.1 – Experimentelle Überprüfung des Filters.....	19
B 1.2 – Anwendung auf Klarinettentöne.....	22
B 1.3 – Fazit.....	24
B 2 – Zweiter Ansatz: Wavetable Synthese + Sampling.....	24
B 2.1 – Das "Deterministic plus Stochastic"-Modell.....	24
B 2.2 – Nachbildung der deterministischen Klanganteile.....	26
B 2.2.1 – Probleme bei der exakten Frequenzbestimmung.....	27
B 2.2.2 – Erste Umsetzung des Verfahrens.....	29
B 2.2.3 – Aufhebung der Diskontinuität an Schleifenübergängen.....	30
B 2.2.4 – Superposition mehrerer Perioden.....	32
B 2.2.5 – Experiment zur Überprüfung der Effektivität.....	33
B 2.2.6 – Anwendung auf Klarinetten-Wavetables.....	35
B 2.2.7 – Dokumentation der Arbeitsschritte zur Erstellung der Wavetables.....	36
B 2.3 – Nachbildung der stochastischen Klanganteile.....	38
B 2.4 – Reproduktion des natürlichen Klangverhaltens.....	40
B 2.4.1 – Dynamik.....	42
B 2.4.2 – Tonansatz.....	43
B 2.4.3 – Tonübergänge.....	47
B 2.4.4 – Intonation.....	51
B 2.4.5 – Spieltechniken.....	54
B 2.4.6 – Modellierung der akustischen Umgebung.....	55
TEIL C: Diskussion der Ergebnisse	59
C 1 – Analyse.....	59
C 1.1 – FFT-Analyse.....	59
C 1.2 – Subjektive Beurteilung des Klangs.....	61
C 2 – Anwendbarkeit auf weitere Instrumententypen.....	62
C 3 – Fazit und Ausblick.....	63
C 4 – Ethische Bewertung der technologischen Entwicklung.....	64
Quellenangaben	68
Anhang A1 – Hörbeispiele	69
Anhang A2 – Modulatorübersicht	72

Vorbemerkungen

- Die verwendeten Notennamen in dieser Arbeit entsprechen den Tonhöhenbezeichnungen innerhalb der Software KONTAKT und Cubase: D2 entspricht d (147Hz), D3 entspricht d' (294Hz). Das deutsche h' wird zu B3, das deutsche b' zu A#3.
- Sämtliche in der Arbeit angeführten FFT-Diagramme wurden mit der kostenlosen Software Wavosaur 1.05 erstellt und für eine bessere Übersicht mit Photoshop bearbeitet (die Farbe geändert und übereinandergelegt). Das FFT-Fenster entspricht einer Blackman-Harris-Fensterfunktion mit 100dB Empfindlichkeit. Für die bessere Erkennbarkeit der Obertonreihen wurde eine lineare Frequenzdarstellung gewählt.
- Die Klarinettenaufnahmen wurden mit der Samplerate 48kHz und einer Auflösung von 24bit aufgenommen.

Abkürzungsverzeichnis

- *CC*: Control Change (MIDI-Controller)
- *dB*: Dezibel
- *f*: forte, laut
- *ff*: fortissimo, sehr laut
- *FFT*: Fast Fourier Transformation
- *LFO*: Low Frequency Oscillator
- *MIDI*: Musical Instrument Digital Interface
- *mp*: mezzopiano, mittelleise
- *mf*: mezzoforte, mittellaut
- *pp*: pianissimo, sehr leise
- *p*: piano, leise
- *AET*: Authentic Expression Technology
- *VST/VSTi*: Virtual Studio Technology
- *AHDSR*: Attack – Hold – Decay – Sustain – Release (Hüllkurventyp)

Einleitung

In der aktuellen Musikproduktion spielen virtuelle Instrumente eine bedeutende Rolle, da sie schnell und kostengünstig den Klang eines Instrumentes nachbilden können, ohne dass zusätzliche Musiker engagiert werden müssen. Das derzeit wichtigste Verfahren ist die Samplingtechnologie, bei der zuvor eingespielte Einzeltöne wiedergegeben werden. Dabei ist jedoch ein hoher Programmieraufwand in der Nachbearbeitungsphase notwendig, da zwischen verschiedenen Artikulationen gewählt werden muss, um musikalische Phrasen realistisch nachzubilden. Dies führt zu einem unintuitiven Workflow im Kompositionsprozess, der sich von der Arbeit mit realen Musikinstrumenten essentiell unterscheidet. Bei dem Versuch, eine in Echtzeit spielbare Samplelibrary von Blas- oder Streichinstrumenten zu erstellen, stösst man auf technische Schwierigkeiten, die für die Entstehung von deutlich hörbaren Klangartefakten verantwortlich sind.

In dieser Arbeit sollen zwei unterschiedliche Verfahren untersucht werden, die dieses Problem zu lösen versuchen, und somit eine in Echtzeit spielbare Reproduktion der besagten Instrumentengruppen ermöglichen. Seit der neuesten Version des wichtigsten Softwaresamplers KONTAKT der Firma Native Instruments ist ein spezieller Filtereffekt verfügbar (die sogenannte "Authentic Expression Technology"), der dieses Problem adressiert, und mithilfe einer FFT-Analyse zwischen unterschiedlichen Obertonstrukturen überblenden kann. Die unterschiedlichen Dynamikstufen können dadurch simuliert werden. Der andere Ansatz beruft sich auf das "Deterministic plus Stochastic"-Modell von X. Serra, und erzeugt die unterschiedlichen Klanganteile separat mithilfe einer Kombination aus einer speziell entwickelten Wavetable-Synthese und der Sampling-Technologie. Anhand der Reproduktion einer Klarinette soll die Effektivität dieses im Laufe der Arbeit entwickelten Verfahren getestet werden. Die Entscheidung für dieses Instrument liegt in dessen klanglicher Vielfalt, dem großen Einsatzgebiet in verschiedenen Musikstilen und der deutlichen spektralen Variation unterschiedlicher Dynamikstufen begründet, allerdings soll das Verfahren auch auf die Anwendbarkeit auf weitere Blas- oder Streichinstrumente untersucht werden.

Die Arbeit gliedert sich in drei Teile. Zur Vorbetrachtung sollen die derzeit relevanten Verfahren, namentlich Sampling und Physical Modeling vorgestellt werden. Ausserdem wird das MIDI-Protokoll und die Samplingsoftware KONTAKT erklärt, da sich die Reproduktion der Klarinette auf diesen zwei Technologien stützt. Im Hauptteil der Arbeit wird anhand der Klarinettenaufnahmen die AET-Technologie getestet, und die Entwicklung des eigenen Verfahrens dokumentiert. Im letzten Abschnitt der Arbeit wird

die virtuelle Reproduktion der Klarinette analysiert und anhand kurzer musikalischer Phrasen mit dem realen Vorbild verglichen. Nach Auswertung der Analyse werden noch Konzepte für die Weiterentwicklung des Verfahrens angeführt, um einen Ausblick auf zukünftige Aufgabenstellungen zu geben. Eine sozial-ökonomische Betrachtung dieses technologischen Fortschritts schliesst die Arbeit ab.

TEIL A: Vorbetrachtungen

A 1 – Alternative Technologien

Am derzeitigen Stand der Entwicklung spielen eigentlich nur noch zwei Technologien eine relevante Rolle: Physical Modeling und das Sampling-Verfahren. Ältere Technologien wie die FM-Synthese oder andere Syntheseformen sind aufgrund des künstlichen Klangeindrucks für die Reproduktion natürlicher Instrumentenklänge nicht mehr einsetzbar, und werden daher in dieser Arbeit nicht weiter behandelt. Bevor im Hauptteil der Arbeit an einer Weiterentwicklung der Klangqualität virtueller Instrumente gearbeitet wird, sollen erst diese beiden alternativen Ansätze erläutert werden.

A 1.1 – Physical Modeling

"Physical Modeling" steht für den Ansatz, ein physikalisches Modell realer Musikinstrumente zu implementieren, welches unter Berücksichtigung der physikalischen Bewegungsgesetze der einzelnen Instrumententeile und deren Interaktion untereinander den entstehenden Klang berechnet. Der Vorteil des Verfahrens liegt in der komplexen Reproduktion des Klangverhaltens, das auch spezielle Spieltechniken wie z.B. Überblasen ohne spezielle Modifikationen erlaubt. Durch die Manipulation der physikalischen Parameter lassen sich auch extreme Klänge erzeugen, die mit normalen Instrumenten nicht zu erreichen sind. Neben der Reproduktion des Klangverhaltens für virtuelle Instrumente wird das Verfahren auch zur Modellbildung und Überprüfung der Tonerzeugungstheorien für die Verbesserung realer Instrumente verwendet. [Hagenow, 2001, S. 24]

Es gibt mehrere Verfahren, die Physical Modeling Technologie zu implementieren. Eine detaillierte Beschreibung der einzelnen Verfahren findet sich in der Arbeit "Evaluation of Modern Sound Synthesis Methods Report 48" [Tolonen et al. 1998]. Allerdings kommt für Echtzeitanwendungen aufgrund der hohen Rechenanforderungen nur die Wellenleitersynthese in Frage. Die Beschreibung und eine Anwendung dieses Verfahrens findet sich bei [Hagenow 2001]. Da "Physical Modeling" in dieser Arbeit nur als alternative

Methode angeführt werden soll, wird an dieser Stelle nicht weiter auf die verwendeten Algorithmen eingegangen.

Im Bereich der kommerziell erhältlichen virtuellen Blasinstrumente ist Physical Modeling von mäßiger Relevanz. Der (etwas betagte) Synthesizer VL70m von Yamaha und die virtuellen Blasinstrumente von Wallander Instruments sind die wichtigsten Produkte des Segments. Die Physical-Modeling-Technologie ermöglicht eine dynamische Tonformung und tiefgreifende Soundmanipulationsmöglichkeiten sowie eine verhältnismäßig realistische Simulation spezieller Spieltechniken wie Vibrato, Flatterzunge oder Glissandi. Da der Ton allerdings komplett synthetisch erzeugt wird, ist das Klangbild im Vergleich zu gesampelten Instrumenten etwas steril.

A 1.2 – Sampling

Unter Sampling versteht man die Wiedergabe zuvor aufgenommener Audiodateien (sog. Samples). Dadurch können Instrumentenklänge reproduziert werden, indem Einzeltöne aufgezeichnet werden, und dann in variiertes Tonhöhe wiedergegeben werden können. Das erste Gerät, das diese Möglichkeiten bot, war der Mellotron, der Anfang der 50er Jahre des letzten Jahrhunderts entwickelt wurde, und auf Magnetband aufgespielten Klänge per Tastendruck abspielen konnte.

Da die Tonhöhe beim herkömmlichen Sampling durch Variation der Abspielgeschwindigkeit erzeugt wird, kommt es aufgrund der Stauchung der Frequenzspektren zu einer Artefaktbildung im Klang abhängig von der Differenz zur Originaltonhöhe. Weiterhin kann die Änderung der Obertonstruktur bei unterschiedlichen Dynamikstufen bei einer Verwendung von einem Sample für den gesamten Dynamikbereich nicht reproduziert werden.

Durch die Verbesserung der Speichertechnologien und der damit einhergehenden deflationären Preisentwicklung für digitalen Speicher seit ca. 1990 wurden die Möglichkeiten des Samplings massiv gesteigert: Der für damalige Verhältnisse umfangreich ausgestattete Sampler S1000 der Firma AKAI besitzt einen Speicher von 32MB. Heutzutage gehen die Speicheranforderungen einiger Softwareinstrumente in den dreistelligen GB-Bereich (Der virtuelle Flügel *VIENNA IMPERIAL* der Firma VSL benötigt beispielsweise 500GB Speicherplatz).

Diese technologische Entwicklung ermöglichte eine Verfeinerung des Verfahrens: da Speicherplatz (fast) keine Rolle mehr zu spielen scheint, kann durch die Erhöhung der Anzahl der verwendeten Einzeltöne die Klangqualität deutlich gesteigert werden. Dies

geschieht in mehreren Dimensionen:

Tonhöhe: mit Einzelaufnahmen von jedem Halbton (chromatisches Sampling) wird die Artefaktbildung durch Stauchung des Frequenzspektrums vermieden.

Dynamikstufen: mit verschiedenen Aufnahmen der unterschiedlichen Dynamikstufen kann das spektrale Verhalten des Instruments beliebig genau reproduziert werden. Bei Tasten- und Schlaginstrumenten funktioniert dies ohne Probleme, soll der Ton aber nach dem Einschwingvorgang verändert werden (bei Blas- und Streichinstrumenten) kommt es ohne spezielle Gegenmaßnahmen zu starken Kammfiltereffekten (siehe unten)

Variationen desselben Tons: Wenn bei schnellen Tonwiederholungen das selbe Audiomaterial wiederholt wird, kommt es zu einem Klangeindruck, der als "Machine-Gun-Effekt" bezeichnet wird. Um dies zu verhindern, werden mehrere gleiche Töne aufgenommen und bei der Wiedergabe per Rotation variiert (das sogenannte Round-Robin Verfahren). Das Round-Robin Prinzip ist für die adäquate Nachbildung von perkussiven Instrumenten von hoher Bedeutung, sodass teilweise bis zu 10 Variationen der selben Ton- und Dynamikstufe verwendet werden.

Intervalle: Um die charakteristischen Klangmerkmale beim Übergang zwischen Tönen einzufangen, werden neben den Einzeltönen auch Intervall-Aufnahmen erstellt. Dadurch kann eine wesentlich natürlichere Phrasierung erreicht werden, da eine simple Aneinanderreihung von Einzeltönen meist nicht dem Charakter einer real gespielten Melodie entspricht.

Die Beschränkung in der Anzahl der verwendeten Samples liegt nur noch im Arbeitsaufwand des Herstellers. Für die Aufnahmen der 70 000 Einzeltöne des Vienna Imperial musste daher eine Roboter-Konstruktion, die CEUS-Technologie von Bösendorfer verwendet werden, da der Zeitaufwand für einen Pianisten nicht mehr zumutbar gewesen wäre¹.

Mithilfe dieser Verbesserungen konnte sich das Samplingverfahren gegenüber anderen konkurrierenden Technologien durchsetzen, und ist im Bereich der virtuellen Reproduktion realer Instrumente derzeit das am häufigsten eingesetzte Verfahren, was die fast unüberschaubar hohe Zahl der existierenden Firmen belegt. Bei der Reproduktion von Blas- und Streichinstrumenten kommt es jedoch zu technischen Problemen, die im nächsten Kapitel beschrieben werden.

1 <http://vsl.co.at/de/211/442/478/1701/1305.htm>

A 1.2.1 – Artefaktbildung bei der Interferenz von Schwingungen

Bei Musikinstrumenten, bei denen die Anregung (und damit die Tongestaltung) im quasistationären Zustand variiert werden kann (d.h. bei Blas- und Streichinstrumenten), muss für eine zufriedenstellende Reproduktion diese Möglichkeit ebenfalls implementiert werden. Die naheliegendste Lösung ist die gleichzeitige Wiedergabe von unterschiedlichen Dynamikstufen und die Überblendung mithilfe MIDI-Steuerdaten (vgl. *Abbildung 1*)

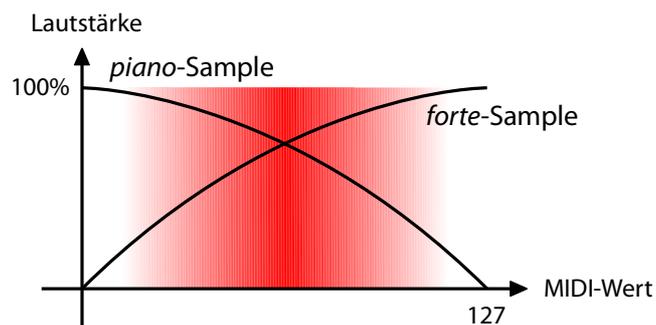


Abbildung 1: (vereinfachter) Crossfade zwischen Dynamikstufen

Allerdings kommt es bei dieser Verfahrensweise zu deutlich hörbaren Klangartefakten, deren Intensität vom jeweiligen MIDI-Wert abhängt (vgl. rote Fläche in *Abbildung 1*). Der Grund dafür sind Kammfiltereffekte, die bei der Überblendung zweier Signale mit ähnlicher Frequenz auftreten: Da die Frequenz der Schwingung bei physikalischen Systemen immer stochastischen Einflüssen unterworfen ist (z.B. Lufttemperatur, Ausdehnung usw.), weicht die Tonhöhe selbst bei zwei identisch gespielten Tönen immer minimal voneinander ab. Im Fall einer Klarinette, bei denen die Tonhöhe bei pp-Tönen bis zu 35 Cent von fortissimo-Klängen abweicht², ist dieser Effekt sogar noch stärker zu beobachten.

Die Differenzen in der Periodenlänge bewirken eine Verschiebung der Phasenlage mit der Frequenz

$$(1) \quad \Delta f = |f_{\text{forte}} - f_{\text{piano}}|$$

die zu einer periodisch wechselnden Pegeladdition und -auslöschung in unterschiedlichen Frequenzbereichen führt.

Abbildung 2 zeigt exemplarisch die Auswirkung dieses Effekts auf eine Überlagerung zweier Schwingungen, in diesem Beispiel zwei an sich identische Sinustöne mit Grundton

² empirischer Durchschnittswert der Klarinettenaufnahmen

200 Hz, wobei einer um 2 Cent nach unten verschoben wurde. Die Zeit zwischen zwei Auslöschungen (in diesem Fall 4,35 Sekunden) ist der Kehrwert der Frequenzdifferenz, d.h. die Töne weichen um 0,22 Hz voneinander ab.

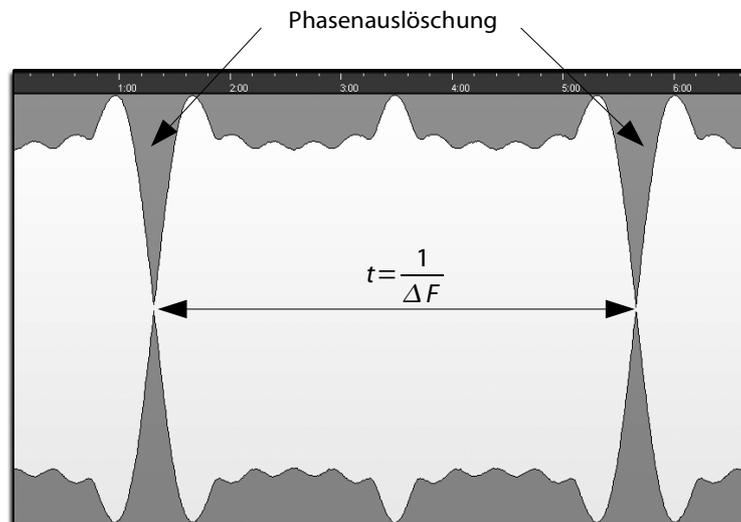


Abbildung 2: Phasenauslöschung bei Überlagerung ähnlicher Frequenzen

Eine vollständige Phasenauslöschung des gesamten Pegels wie im Beispiel kommt natürlich nur bei identischem Spektrum der überlagerten Töne vor: Da unterschiedliche Dynamikstufen natürlich unterschiedliche Spektralverläufe besitzen, kommt es nicht zu einer vollständigen Auslöschung. Stattdessen tritt der Effekt für jeden Oberton getrennt auf. Es kommt also zu einer unterschiedlichen, aber konstant periodischen Auslöschung und Verstärkung der einzelnen Obertöne. Der entstehende Klangeffekt ist von Modulationseffekten (Flanger / Phaser) bekannt, die diesen Effekt bewusst zur Klangverfremdung nutzen. Für eine überzeugende Reproduktion natürlicher spektraler Dynamik kommt das Verfahren der simplen Überblendung der Dynamikstufen daher nicht in Betracht.

Der vielversprechendste Lösungsansatz kommt von der Firma Sample-Modeling, die mithilfe ihrer patentierten "Harmonic Alignment"-Technologie eine bemerkenswert realistische Nachbildung von Blechblasinstrumenten erreicht hat. Ein weiterer Ansatz ist die "Authentic Expression Technologie" von Native-Instruments, die auf Überblendung verschiedener Töne verzichtet und mithilfe FFT-Analyse und -Filterung die unterschiedlichen Dynamikstufen nachzubilden versucht. Die Analyse der Effektivität dieser (für KONTAKT-Besitzer offen nutzbaren) Technologie und die Entwicklung eines eigenen Verfahrens zur Reproduktion spektraler Dynamik stellt die Hauptaufgabe dieser Arbeit dar.

A 2 – Erläuterungen der verwendeten Technologien

A 2.1 – Das MIDI-Protokoll

Das MIDI-Protokoll wurde 1983 von der MIDI-Manufacturers Association vorgestellt, und ist seitdem das wichtigste Steuerprotokoll für elektronische Musikinstrumente. Da die Steuerung der geplanten Klarinette mithilfe dieses Protokolls geschehen wird, soll hier ein kurzer Überblick über die Funktionen aber auch über die technischen Begrenzungen des mittlerweile fast dreißig Jahre alten Standards gegeben werden. [Wikipedia, Artikel "MIDI"]

Die Übertragung im MIDI-Protokoll geschieht über eine unidirektionale Datenübertragung mit einer Geschwindigkeit von 31,25 kBaud/sec. Eine MIDI-Botschaft besteht aus einem Statusbyte, das die Art des MIDI-Befehls angibt, sowie zwei Datenbytes, die die Nutzdaten enthalten. Da das erste Bit jeweils für die Unterscheidung zwischen Status- und Datenbyte verwendet wird, können pro Datenbyte nur $2^7 = 128$ verschiedene Werte übertragen werden. Es gibt eine Vielzahl von verschiedenen Befehlstypen, allerdings soll sich hier auf die Befehle beschränkt werden, die im Laufe dieser Arbeit relevant werden:

Note-On / Note-Off Befehle:

Die Note-On / Note Off-Befehle stellen den Kern des MIDI-Protokolls dar, und geben an, wie lange und mit welcher Tonhöhe und Lautstärke ein Ton gespielt werden soll. Der Note-On Befehl besteht aus dem Statusbyte $0x9n$, wobei n den MIDI-Kanal von 0-15 darstellt und den Datenbytes für Tonhöhe (von 0 bis 127) und der Anschlagsstärke (ebenfalls in 7bit). Der Note-Off Befehl hat das Statusbyte $0x8n$ und die Datenbytes Tonhöhe und Release-Velocity (analog zur Anschlagsdynamik, allerdings gibt es fast kein Eingabegerät, das diese Größe überträgt.)

Pitch-Bend:

Der Pitchbend-Befehl erlaubt eine Modifikation der Tonhöhe des jeweiligen MIDI-Kanals. Das Statusbyte lautet $0xE_n$ (n = Kanalnummer), und die beiden Datenbytes geben die Tonhöhenmodulation (mit einer Auflösung von 14bit) an. Den Umfang der Tonmodulation legt allerdings der Klangerzeuger alleine fest (meistens zwischen einem Halbton und einer Oktave).

Controller Befehle (MIDI-CCs)

Die Controller-Befehle dienen der Übertragung von Echtzeitparametern, die für die Klanggestaltung innerhalb des Klangerzeugers verwendet werden können. Das Statusbyte lautet $0xBn$. Das erste Datenbyte überträgt die Controllernummer (0-127) und das zweite Datenbyte den Controllerwert (ebenfalls 0-127).

Da Controllerbefehle im Gegensatz zur Anschlagsstärke nicht mit einem Note-On Befehl zusammenhängen, bietet sich eine Verwendung für die Dynamikkontrolle von Blasinstrumenten an, da der Ton zu jeder Zeit verändert werden kann.

Die Beschränkung auf den 7bit-Wertebereich ist für die Reproduktion von Blasinstrumenten die größte technische Einschränkung: der damit maximal darstellbare Dynamikumfang von ca. 42dB^3 reicht nicht für eine ausreichend abgestufte Nachbildung der Instrumentendynamik.

Ebenso kritisch ist die zeitliche Auflösung, die selbst im Idealfall (d.h. wenn die gesamte Bandbreite der MIDI-Schnittstelle für die Übertragung der Controllerdaten zur Verfügung stünde) nur bei $0,8\text{ms}$ liegt. Der reale Wert liegt deutlich darunter: das Abtastintervall bei Tastaturen der Firma CME beträgt beispielsweise ganze 20ms . Vor allem bei den schnellen Änderungen in der Ein- und Ausschwingphase käme es zu hörbaren Artefakten, die allerdings durch eine software-interne Glättung etwas abgeschwächt werden können.

Vergleicht man diese Daten mit der Abtastauflösung des Red-Book CD-Standard⁴, kommt man auf 3% der zeitlichen Auflösung und 0,2% der Dynamikauflösung. Obwohl sich diese Einschränkungen aus den Spezifikationen des MIDI-Protokolls ergeben, wäre ein Workaround in den Grenzen des MIDI-Standards denkbar:

- Indem der Befehlstyp des Pitchbends (siehe oben) mit der erhöhten Auflösung von 14bit (16384 Stufen) für die Dynamiksteuerung verwendet würde, könnte die vertikale Auflösung auf 84dB (25% des Redbook-Standards) erhöht werden.
- Mit der Übertragung der MIDI-Daten per USB-Schnittstelle kann die Übertragungsrates beinahe beliebig gesteigert werden. Die Grenzen liegen hierbei in der Verarbeitungsgeschwindigkeit der Klangerzeuger.

Da in dieser Arbeit die technischen Rahmenbedingungen allerdings nur angedeutet werden sollen, würde eine genauere Untersuchung den Umfang der Arbeit sprengen.

³ nach der Überschlagsformel $\text{Gesamtdynamik} = \text{Bitzahl} * 6\text{dB}$

⁴ $44,1\text{ kHz}$ Abtastfrequenz und 16bit Auflösung

A 2.1.1 – Wichtige Controllernummern

Das erste Datenbyte bei Controllerbefehlen gibt die Nummer des Controllers an. In den General-MIDI Spezifikationen ist jeder Controllernummer eine spezielle Funktion zugeordnet. Im folgenden werden die Controllernummern aufgezählt und erläutert, die im Verlauf der Arbeit verwendet werden:

- CC Nr. 1: Modulation Wheel – Das Modulationwheel ist ein Steuerrad, das bei den meisten Keyboards links neben der Klaviertastatur angeordnet ist, und mit Steuerfunktionen belegt werden kann.
- CC Nr. 2: Breath Controller – Der Breath-Controller ist ein Gerät, das den Blasdruck per Drucksensor erfasst, und als MIDI-CC überträgt (siehe unten)
- CC Nr. 64: Sustain Pedal - Das Sustainpedal ist vor allem bei der Steuerung von Tasteninstrumenten von hoher Bedeutung, da hier die Funktion des Haltepedals bei Klavieren simuliert werden kann.
- Aftertouch: Bei vielen Keyboards (schätzungsweise 60%) ist unter der Tastatur ein Drucksensor angebracht, der nachdem der Ton gedrückt wurde, Werte übertragen kann. Obwohl Aftertouch eigentlich kein Controller-Befehl, sondern ein eigener Befehlstyp mit Statusbyte 0xDn ist, kann man ihn aufgrund der ähnlichen Struktur (ebenfalls 7bit Wertebereich) als Controller-Befehl auffassen (in der Skriptsprache von KONTAKT geschieht der Zugriff auf Aftertouch-Werte über den virtuellen CC-Wert 129)

A 2.1.2 – Breath Controller

Der Breath Controller (*Abbildung 3*) ist ein von Yamaha entwickelter MIDI-Controller. Die Funktionsweise des Controllers ist relativ simpel: Mithilfe eines Drucksensors wird der Atemdruck des Musikers in ein MIDI-Controllerbefehl umgewandelt.



Abbildung 3: Yamaha BC3

Diese Größe steht nun zur Steuerung beliebiger Parameter zur Verfügung. Natürlich bietet sich eine Steuerung der Dynamik an, es sind theoretisch aber auch andere Parameterkonfigurationen möglich. Im Vergleich zu anderen Eingabegeräten (z.B. Modulationsrad) führt die Verwendung des Breathcontrollers zu wesentlich natürlicheren Controllerverläufen. Zudem werden virtuelle Blasinstrumente intuitiv spielbar, was den Spielspass deutlich erhöht.

A 2.2 – Native Instruments KONTAKT

Native Instruments KONTAKT ist ein Softwaresampler, der über die VSTi-Schnittstelle in Host-Sequencer eingebunden werden kann. Der Sampler lädt Audiodateien (Samples) in den Hauptspeicher, die über MIDI-Befehle angesteuert werden können. Dadurch kann der Klang realer Instrumente reproduziert werden. Zusätzlich bietet er die Möglichkeit, über interne Effekte und Modulatoren wie LFOs oder Hüllkurven das Klangbild der virtuellen Instrumente zu modifizieren. Die Organisation der Samples findet dabei auf verschiedenen Ebenen statt [Morgenstern, 2009]:

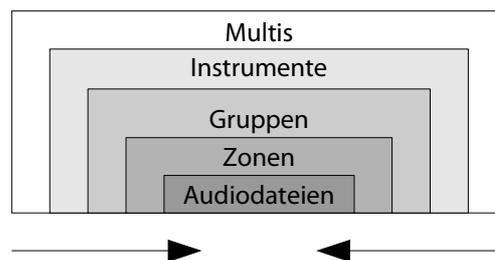


Abbildung 4: Organisation der Ebenen

Sample-Mapping:

Die Verteilung der Audiodateien auf unterschiedliche Bereiche des MIDI-Keyboards nennt man Samplemapping. Man unterscheidet dabei zwischen horizontalem Mapping (Verteilung auf unterschiedliche Tonhöhen) und vertikalem Mapping (Verteilung auf unterschiedliche Anschlagstärken)

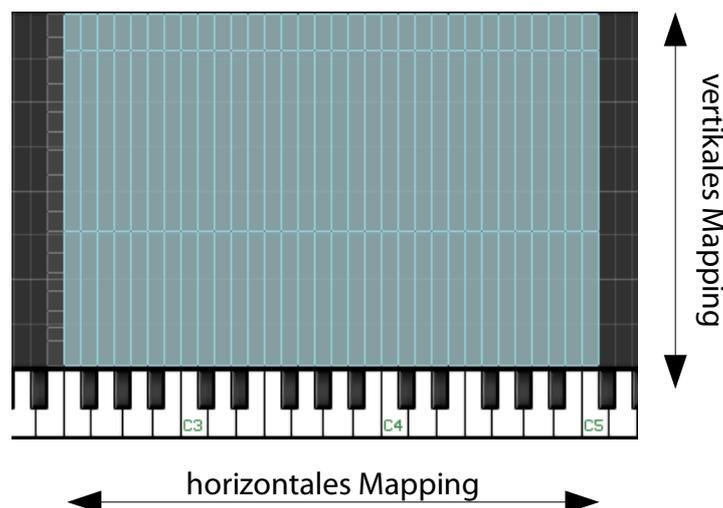


Abbildung 5: Beispiel eines Sample-Mapping

Abbildung 5 zeigt ein Beispiel für ein Samplemapping: die einzelnen blauen Rechtecke

(sogenannte Zonen) beinhalten einzelne Audiodateien, die mit einem Tastendruck auf die jeweilige Keyboardtaste gestartet werden können. Für jede Zone kann die Tonhöhe, die Lautstärke und die Panoramaverteilung separat eingestellt werden, zudem können die Startzeit, Looppunkte (und Crossfades) gesetzt werden.

Gruppen:

Die nächsthöhere Ebene sind sogenannte Gruppen. Man kann mehrere Zonen in Gruppen zusammenfassen, um sie gemeinsam zu modifizieren. Die Modulatoren (Hüllkurven, LFOs, MIDI-CCs) in KONTAKT arbeiten auf Gruppenebene. Ausserdem verfügt jede Gruppe über spezielle Effekt-Slots, die mit internen Insert-Effekten gefüllt und moduliert werden können.

Es gibt die Möglichkeit, Bedingungen für die Aktivierung von einzelnen Gruppen anzugeben. Dadurch kann beispielsweise das Round-Robin Verfahren implementiert werden, indem die verschiedenen Gruppen mit jeder gespielten Note zyklisch durchlaufen werden, um den sogenannten "Maschine-Gun-Effekt" zu verhindern.

Instrument:

Ein Instrument besteht aus einer (oder mehreren) Gruppen, die gleichzeitig angesteuert werden.

Zusätzlich zu den Gruppen-Effektslots gibt es noch Instrument-Insert- und -Send-Effektslots, mit denen der Klang modifiziert werden kann. Allerdings haben die internen Modulatoren keinen Zugriff auf die Parameter dieser Effekte, welche nur durch Skriptbefehle angesprochen werden können.

Multis:

Multis sind Konfigurationen aus mehreren Instrumenten. Normalerweise werden Multis für Ensemblesounds eingesetzt. In dem hier entwickelten Verfahren werden jedoch die einzelnen Klangbausteine (das heisst die Wavetables und die verschiedenen Nebengeräusche) auf Instrumentenebene erstellt, um die Skriptprogrammierung übersichtlicher zu gestalten und um eine gegenseitige Beeinflussung der einzelnen Skriptmodule zu verhindern.

A 2.2.1 – Skriptsprache

Die Software KONTAKT besitzt als einziger Softwaresampler die Möglichkeit, über eine eingebaute Programmiersprache weitere Funktionen zu implementieren. Von der Erstellung individueller Benutzeroberflächen über Modifikation der eingehenden MIDI-Daten bis zu komplexer Klangmodulation mithilfe Modifikation der internen Effekte (auch auf Instrumentenebene) bieten sich tiefgreifende Möglichkeiten, den Klang und das Verhalten der Instrumente zu verändern, ohne sich mit Low-Level Funktionalitäten wie Disk-Streaming und Portabilität / Stabilität auseinandersetzen zu müssen.

Die Syntax der Programmiersprache ist an PASCAL angelehnt. Die Grundbausteine der Skriptsprache sind sogenannte "Callbacks" (Interrupts), d.h. Programmanweisungen, die jedesmal aufgerufen werden, wenn ein bestimmtes Ereignis eintritt: sobald beispielsweise eine Taste gedrückt wird, wird der "on note"-callback gestartet und die darin enthaltenen Anweisungen prozedural ausgeführt. Die Initialisierung der Variablen und Erstellung der Benutzeroberfläche geschieht normalerweise im "on init"-Callback, der nach dem erfolgreichem Ladevorgang des Instruments ausgeführt wird.

```
1  on init
2      declare const $kleine_tertz := 3
3  end on
4  on note
5      ignore_event($EVENT_ID)
6      play_note($EVENT_NOTE + $kleine_tertz, $EVENT_VELOCITY, 0, -1)
7  end on
```

Codebeispiel 1: einfache MIDI-Transposition (kleine Tertz nach oben)

Codebeispiel 1 zeigt exemplarisch ein kleines Skript, das eine Transponierung um eine kleine Tertz (drei Halbtonschritte) nach oben bewirkt: Jede Note löst den "on note"-Callback aus. Diese wird zunächst ignoriert (Zeile 5). In Zeile 6 wird eine neue Note erzeugt, die eine beim Programmstart festgelegte Integer-Konstante *\$kleine_tertz* zu der alten Tonhöhe dazu addiert.

Für die Erstellung der Skripte wurde der kostenlose KScriptEditor von Nils Liberg verwendet, der zusätzliche Funktionen wie Familienorganisation von Variablen und benutzerdefinierte Funktionen / Makros zur Verfügung stellt. Daher sind Teile der im Laufe der Arbeit angeführten Codebeispiele nicht direkt in KONTAKT einsetzbar, sondern müssen erst mit besagtem Editor vorkompiliert werden.

TEIL B: Anwendung

B 0 – Klangliche Analyse der Klarinette

Die Klarinette gehört zu den transponierenden Holzblasinstrumenten mit einfachem Rohrblattmundstück. Die Tonerzeugung findet im Mundstück statt: Der Luftstrom, der vom Spieler in die Klarinette geblasen wird, versetzt das Rohrblatt in Schwingung. Diese Schwingung regt die Luftsäule innerhalb der Klarinette an, die eine stehende Welle ausbildet. Durch Verkürzung der Luftsäule mithilfe der Klappen kann die Tonhöhe verändert werden.

Aufgrund der mechanischen Konstruktion als einseitig geschlossener Zylinder entstehen nur die ungeradzahigen Vielfachen der Grundschwingung. Die Klarinette überbläst daher in die Duodezime (den dritten Naturton). Dadurch entsteht ein vergleichsweise großer Tonumfang von ca. dreieinhalb Oktaven (klingend D2 bis ca. A5)

Durch den Teiltonaufbau entsteht der charakteristische, hohle Klang der Klarinette, der je nach Tonlage unterschiedlich ausgeprägt ist. Aufgrund der unterschiedlichen Klangfarben in den einzelnen Lagen teilt man den Tonumfang der Klarinette daher in drei Register ein: das "Chalmeau"-Register von D2 bis D3, bei dem der beschriebene Teiltonaufbau besonders ausgeprägt ist, das "Clarin"-Register von D#3 bis G4, bei dem nur noch der zweite Oberton schwach ist, und das "Flageolett-Register" von G#4 bis ca. A5, bei dem eine verhältnismäßig gleichmäßige Verteilung der Obertöne und eine Formantbildung bei ca. 3 – 4 kHz auftritt. [Dickreiter, 1987; Wikipedia-Artikel "Klarinette"]

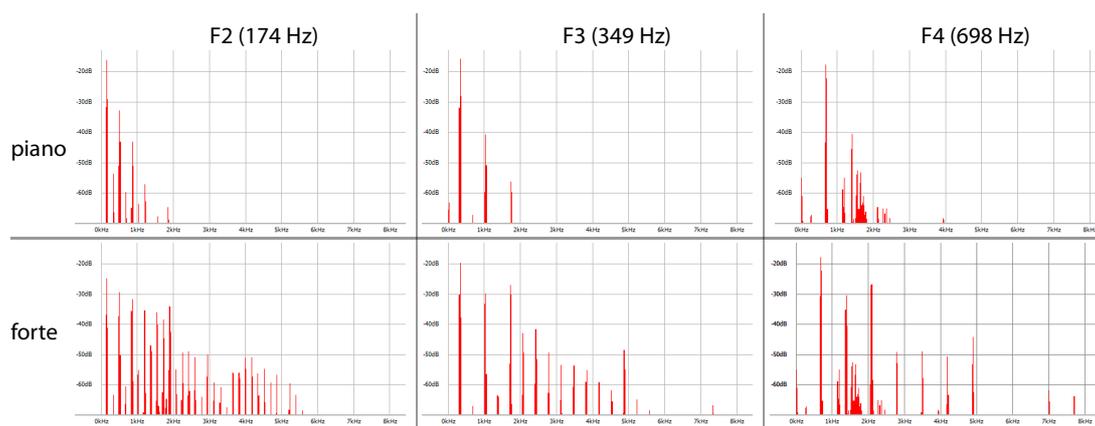


Abbildung 6: FFT-Analyse verschiedener Klarinettenöne

Abbildung 6 beschreibt den spektralen Verlauf verschiedener Dynamikstufen und verschiedener Tonhöhen der Klarinette. Während die Obertöne bei dem piano-Ton deutlich leiser als der Grundton sind, kommen mit steigender Dynamik die

ungeradzahligem Teiltönen hinzu. Im Vergleich zu den tiefen Tonlagen ist die Veränderung der Obertonstruktur in den höheren Registern weniger ausgeprägt.

B 1 – Erster Ansatz: Sampling + AET-Filter

Hinter der "Authentic Expression Technology" (AET) verbirgt sich ein spezieller FFT-Filter, der seit der neusten Version KONTAKT 4 verfügbar ist. Bevor die Anwendbarkeit des Filters auf die Klarinettenklänge untersucht wird, soll anhand mithilfe synthetischer Testsignale die Arbeits- und Wirkungsweise experimentell überprüft werden.

Auf der Produkt-Website verspricht der Hersteller eine artefaktfreie Überblendung (Morphing genannt) zwischen verschiedenen Obertonspektren. Dadurch sollen die verschiedenen Dynamikstufen eines Musikinstrumentes in Echtzeit durchlaufen werden können, ohne mit den anfangs erwähnten Kammfiltereffekten konfrontiert zu sein. Dabei werden die unterschiedlichen Spektren analysiert, und auf ein Sample aufgeprägt. Abbildung 7 zeigt eine mögliche Anwendung des Filters, bei dem das Spektrum des piano-Samples auf das Spektrum des forte-Samples angewendet wird.



Abbildung 7: exemplarische Anwendung des AET-Filters

Für die FFT-Analyse kann ein Glättungsparameter (Smoothing) angegeben werden. Dies führt zu starken Veränderungen im Ergebnis, die teilweise unvorhersehbar sind. Durch systematische Änderung der beiden Smoothingparameter hat sich jedoch ergeben, dass die besten Resultate mit einem hohen Glättungsparameter für das Ausgangssample und einem kleinen Parameter für das Analysesample erreicht werden.

B 1.1 – Experimentelle Überprüfung des Filters

Für eine modellhafte Anwendung wird ein Sinuston (keine Obertöne), eine Dreiecksschwingung (nur ungeradzahligem Obertöne) und ein Sägezahnklang (alle harmonischen Obertöne) mit derselben Frequenz erzeugt. Nun wird das Filter auf die verschiedenen Klänge angewandt. Abbildung 8 zeigt den Effekt des Filters bei einer Anwendung des Sinus-Spektrums auf den Sägezahnklang:

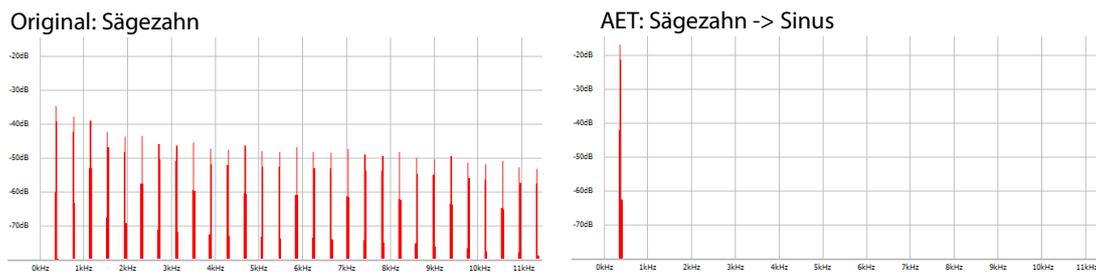


Abbildung 8: Sägezahn / Sinus

Die Obertöne des Sägezahn werden vollständig gedämpft, und es bleibt nur noch die Grundschwingung übrig. Der Hörtest [Hörbeispiel Nr. 1] bestätigt die Effektivität. Die Pegeldifferenz zwischen ersten und zweiten Oberton gibt eine Vorstellung über die Flankensteilheit. Vergleicht man die Effektivität mit einem herkömmlichen Tiefpassfilter, entspricht die Dämpfung einer Flankensteilheit von ca 80dB / Oktave und damit einem Tiefpassfilter 13ter-Ordnung (!).

Als nächstes wird das Dreiecksspektrum auf das Sägezahn-Sample aufgeprägt (Abbildung 9). Das linke Diagramm zeigt das Spektrum der Dreiecksschwingung mit ihren ungeraden Teiltönen. Das rechte Diagramm zeigt die FFT-Analyse des bearbeiteten Sägezahntons.

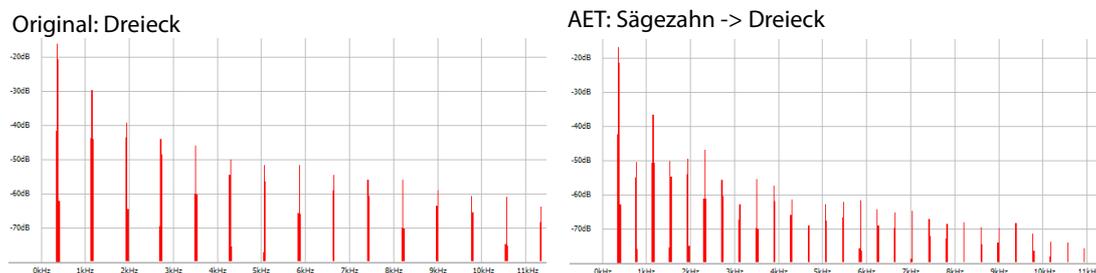


Abbildung 9: Sägezahn / Dreieck

Die geradzahligen Obertöne werden im Vergleich zum Sägezahnspektrum um ca. 11,63 dB gedämpft (Durchschnittswert der ersten 10 geradzahligen Teiltöne). Die Differenz der Pegel der ungeradzahligen Obertöne zur Originalwellenform beträgt durchschnittlich 10dB. Im Hörtest [HB Nr. 2] lässt sich eine leichte Ähnlichkeit zur Original-Dreiecksschwingung erkennen.

Bisher wurde der Filter nur subtraktiv (da die Obertöne gedämpft wurden) eingesetzt. Daher soll nun die Auswirkungen des Filters bei einer Vertauschung des Ausgangs- und Analysesamples untersucht werden (Abbildung 10): Mithilfe der Analyse des Sägezahnspektrums (links) und dessen Anwendung auf die Dreiecksschwingung (rechts) kann überprüft werden, ob das Filter die nicht vorhandenen geraden Teiltöne erzeugen kann [HB Nr. 3].

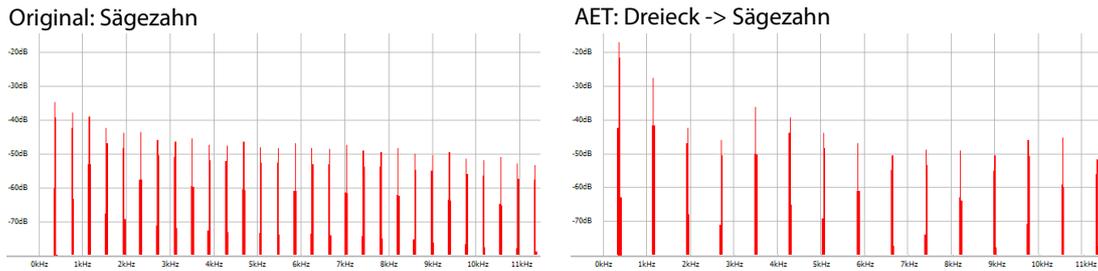


Abbildung 10: additive Anwendung Dreieck -> Sägezahn

Wie zu erwarten ist, entstehen keine neuen Teiltöne: ein FFT-Filter ist ein lineares System, das keine nichtlinearen Verzerrungen erzeugen kann [vgl. Rudow 2008, S. 25]. Daher können nur die bereits im Signal vorhandenen Teiltöne verstärkt / abgeschwächt werden. Die Auswirkungen auf die (in der Dreiecksform vorhandenen) ungeraden Obertöne soll nun genauer untersucht werden. *Abbildung 11* zeigt die Pegeldifferenzen der ungeraden Teiltöne des bearbeiteten Tons zu den beiden Ausgangswellenformen Sägezahn und Dreieck.

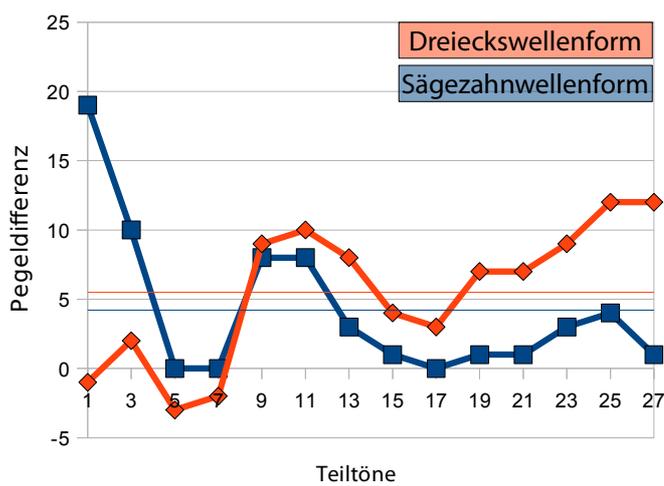
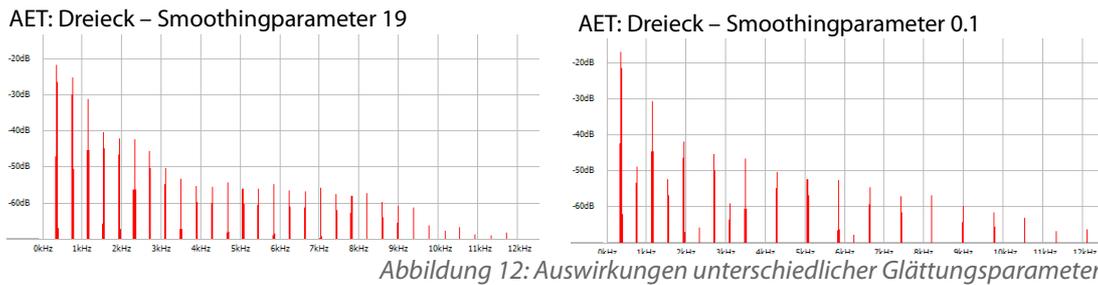


Abbildung 11: Pegeldifferenzen der einzelnen Teiltöne

Die blaue Kurve beschreibt die Differenz zur Sägezahnwellenform, während die orangene Kurve die Differenz zur Dreieckschwingung darstellt. Im tieferen Frequenzspektrum (bis zum 3. Teilton) ähnelt der Ton eher der Dreieckswellenform. Ab dem 5. Teilton ist jedoch die Ähnlichkeit zur Sägezahnschwingung größer.

Insgesamt überwiegt die Ähnlichkeit zur Sägezahnwellenform (die waagrechten Geraden sind die Mittelwerte der Differenzen).

Zuletzt soll noch die Auswirkung des oben erwähnten Smoothing-Parameters untersucht werden. *Abbildung 12* zeigt den Unterschied zwischen zwei gefilterten Sägezahnwellen, die mit dem Dreiecksspektrum bearbeitet werden. Bei dem linken Diagramm wurde für das Analysesample der höchste Smoothingparameter 19 gewählt, für das rechte Diagramm der niedrigste (0.1).



Die Dämpfung der geradzahigen Teiltöne ist bei dem geringen Smoothingwert deutlich stärker. Durch die Glättung des Spektralverlaufs verliert das Filter also an Effektivität (entspricht also einer Verringerung der FFT-Auflösung). In der praktischen Anwendung stellt man allerdings fest, dass mit kleineren Parametern die Wirkung immer unberechenbarer wird, und es zu extremen Verstärkungen einzelner Frequenzen (bis zum Clipping) kommt. Eine pauschale Bewertung ist daher nicht möglich, und muss für jeden Anwendungszweck individuell vorgenommen werden.

Man kann die Ergebnisse des Experiments folgendermaßen zusammenfassen:

1. Der Filter arbeitet sowohl subtraktiv als auch additiv etwa gleich effektiv.
2. Obertöne, die im Ausgangsmaterial nicht enthalten sind, können nicht erzeugt werden
3. Der Smoothing-Parameter erhöht die Effektivität, führt allerdings unter Umständen zu Verzerrungen

B 1.2 – Anwendung auf Klarinettentöne

Nun soll die Wirkungsweise des Filters für die Anwendung auf Klarinettentöne getestet werden. Mit Testaufnahmen einzelner Töne in leiser, mittellauter und lauter Dynamik soll überprüft werden, ob die spektralen Veränderungen mithilfe des AET-Filters reproduziert werden können.

Abbildung 13 zeigt die FFT-Analysen dreier Töne: links der originale pianissimo-Ton, in der Mitte der bearbeitete mezzoforte-Ton und rechts der bearbeitete fortissimo-Ton [HB Nr. 4].

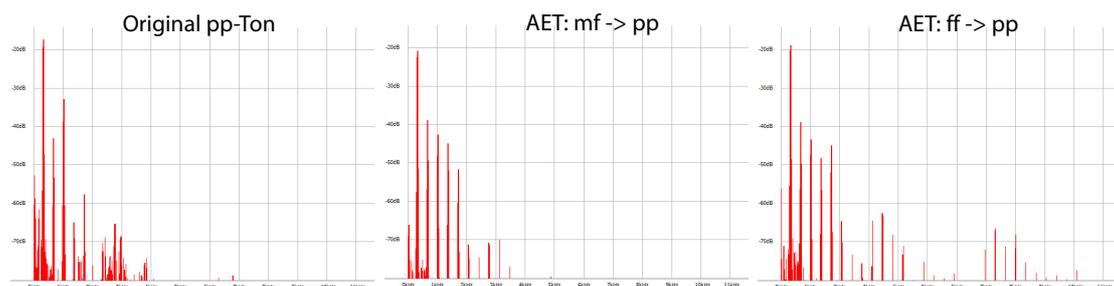


Abbildung 13: AET-Nachbildung des pianissimo-Tons

Auf den ersten Blick sieht das Spektrum des mf-Tons ähnlicher aus als das Spektrum des ff-Tons: Oberhalb 3 kHz sind kaum noch Obertöne vorhanden (im Gegensatz zum ff-Ton). Allerdings ist die charakteristische Obertonverteilung der Klarinette bei beiden Tönen nicht erkennbar: der zweite Oberton ist bei beiden Spektren stärker als der dritte, während im Original der zweite und vierte Oberton ca. 10dB leiser als ihre Nachfolger sind.

Abbildung 14 zeigt die Auswirkungen einer additiven Verwendung des Filters. Links sieht man das Spektrum des Originalen ff-Tons, und links das Spektrum des damit bearbeiteten mf-Tons.

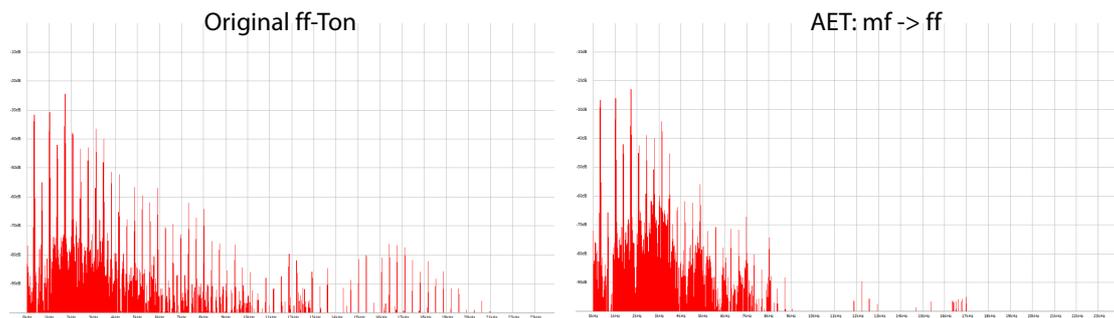


Abbildung 14: AET-Nachbildung des fortissimo-Tons

Auf eine Darstellung des bearbeiteten pp-Tons wurde verzichtet, da man schon am mf-Ton erkennen kann, dass die Obertöne des ff-Tons nicht reproduziert werden können: Während die Obertöne beim Originalton bis an die Grenzen des hörbaren Bereichs gehen, treten bereits bei 8kHz nur noch vereinzelt Obertöne im gefilterten Klang auf. Eine detaillierte Analyse der Pegeldifferenzen lässt sich in *Abbildung 15* erkennen.

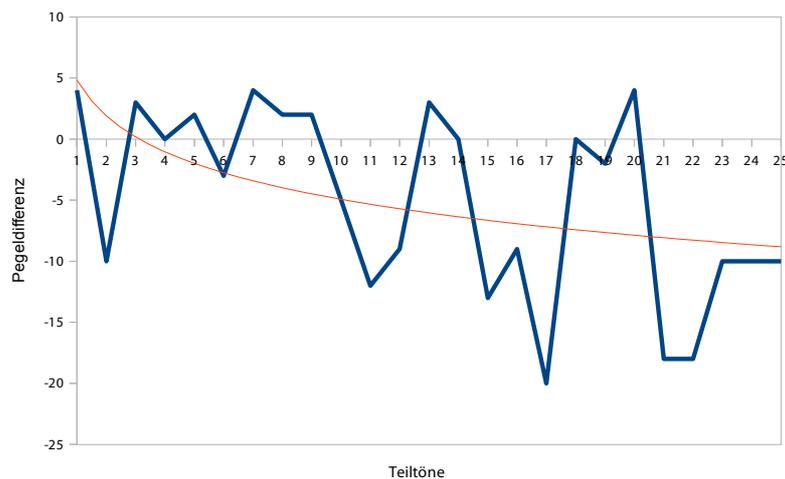


Abbildung 15: Pegeldifferenzen bei additiver Anwendung

Die Regressionskurve (rot) zeigt, dass die Differenz mit steigender Obertonanzahl zunimmt. Die charakteristischen Obertöne der lautesten Dynamikstufe werden also zunehmend ungenauer nachgebildet.

B 1.3 – Fazit

Für die Wahl der Dynamik des Ausgangssamples stößt man nach Auswertung der Testaufnahmen auf folgendes Problem: Während der Filter die Obertöne des ff-Tons für einen pp-Ton nicht effektiv genug dämpft, können die Obertöne des ff-Tons nicht ausreichend mithilfe des mf-Samples nachgebildet werden. Da der Filter zudem auf das gesamte Sample angewendet wird, werden die im Einschwingvorgang prominenteren Luftgeräusche verfälscht.

Aufgrund dieser Einschränkungen wird dieser Ansatz nicht weiter verfolgt. Stattdessen soll ein eigenes Verfahren entwickelt werden, das die geforderte spektrale Dynamik besser nachbilden kann.

B 2 – Zweiter Ansatz: Wavetable Synthese + Sampling

Das Kammfilterproblem der Sampling-Technologie kann man darauf zurückführen, dass zwischen zwei Samples mit gleichem Grundton die Phasenlage nicht überprüf- bzw. korrigierbar ist, da minimale Tonhöenschwankungen auftreten. Es gibt allerdings eine Möglichkeit, die Frequenz zu stabilisieren: wenn anstelle der gesamten Tonlänge nur einzelne Periodenausschnitte in einer Schleife wiedergegeben werden, bleibt die Frequenz konstant. Dieses Konzept ist unter dem Namen "Wavetable-Synthese" bekannt. Eine detaillierte Beschreibung des Verfahrens findet sich in der Arbeit "Wavetable Synthesis 101 - A Fundamental Perspective" von Robert Bristow-Johnson. Diese recht alte Technologie wird heutzutage eigentlich nur noch in billigen Klangerzeugern⁵ eingesetzt um Speicherplatz zu sparen, da nicht mehr ganze Töne, sondern nur noch einzelne Wellenformen gespeichert werden müssen. Im Falle des in dieser Arbeit entwickelten Verfahrens spielen die geringe Ressourcenanforderungen jedoch nur eine untergeordnete Rolle. Viel interessanter ist die Möglichkeit aufgrund der festen Frequenz eine stabile Phasenlage zwischen verschiedenen Dynamikstufen herzustellen. Bevor die Vorgehensweise im Detail beschrieben wird, soll zunächst ein allgemeines Klangmodell erläutert werden, das für das Verständnis der folgenden Schritte von essentieller Bedeutung ist.

B 2.1 – Das "Deterministic plus Stochastic"-Modell

Das "Deterministic plus Stochastic Model" geht davon aus, dass jede akustische Schwingung, die von einem physikalischen System (insbesondere von Musikinstrumenten) erzeugt wurde, in einen deterministischen Anteil und einen

⁵ z.B. GM-Soundchips heutiger Onboard-Soundkarten oder Mobiltelefone

stochastischen Anteil unterteilt werden kann. Der deterministische Teil ist die sich periodisch wiederholende Wellenform im quasistationären Zustand, die durch die Ausbildung stehender Wellen innerhalb des physikalischen Systems entsteht. Der stochastische Teil beschreibt die Energie, die nicht in periodische Schwingungen umgesetzt wird, und kann als zeitvariantes Leistungsdichtespektrum vollständig beschrieben werden. [vgl. Serra, 1997].

Im Falle einer Klarinette ist der deterministische Teil die periodische Schwingung der Luftsäule in der Klarinette, während der stochastische Anteil (das Luftgeräusch) durch den turbulenten Luftstrom im Mundstück verursacht wird.

Mathematisch lässt sich das Modell folgendermaßen ausdrücken:

$$(2) \quad s(t) = \sum_{r=1}^R A_r(t) \cdot \cos[\theta_r(t)] + e(t)$$

wobei sich $s(t)$, der erzeugte Klang, aus der Summe von R Obertönen mit der jeweiligen Amplituden- und Phasenfunktion $A_r(t)$ bzw. $\theta_r(t)$ und der Geräuschkomponente $e(t)$ zusammensetzt. [vgl. ebd.]

Im Zuge der Arbeit "Musical Sound Modeling with Sinusoids plus Noise" beschreibt Xavier Serra ein Verfahren, mit dem eine synthetische Nachbildung von Schwingungen physikalischer Systeme erreicht werden kann: Durch FFT-Analyse des Obertonspektrums und dessen anschließender Modellierung mithilfe additiver Synthese (d.h. durch Kombination verschiedener Sinus-Schwingungen) kann der deterministische Teil nachgebildet werden. Der stochastische (oder residuale) Teil wird erzeugt durch Subtraktion dieser additiv erzeugten Sinusschwingungen vom Originalklang und der Erstellung einer Leistungsdichtenüllkurve der Differenz zwischen dem Originalklang und den synthetisch erzeugten Sinusoiden [vgl. ebd.].

Die Schwierigkeiten bei diesem Ansatz für die Reproduktion spielbarer virtueller Instrumente liegt in der Komplexität der Obertonstruktur der meisten Instrumente, die teilweise bis zu 20 hörbare Obertöne besitzen, deren Verhältnis sich nach Tonlage, Dynamik usw. drastisch verändern kann (vgl. Kapitel B.0 – Klanganalyse der Klarinette). Ausserdem ist die adäquate Nachbildung des stochastischen Anteils durch gefiltertes Rauschen eher schwierig zu realisieren, zumal das Verfahren eine einwandfreie Analyse und Subtraktion der Obertöne voraussetzt, deren Lautstärkeverhältnis selbst im quasistationären Zustand gewissen Änderungen unterworfen sind.

Das in dieser Arbeit entwickelte Verfahren basiert ebenfalls auf diesem Modell, allerdings

wird für die Erzeugung des deterministischen Anteils eine spezielle Wavetablesynthese verwendet, während für die stochastischen Klanganteile auf das Samplingverfahren zurückgegriffen wird.

B 2.2 – Nachbildung der deterministischen Klanganteile

Um das Obertonspektrum der verschiedenen Dynamikstufen abzubilden, wird auf das oben beschriebene Verfahren der Wavetablesynthese zurückgegriffen: dabei wird aus den Crescendotönen der Klarinette pro Dynamikstufe jeweils ein kurzes Stück mit der (ideal berechneten) Periodenlänge T_G des Grundtons herausgeschnitten und in einer Schleife wiedergegeben. Da eine Grundtonperiode genau r Perioden des r -ten Obertons enthält, kann somit das gesamte (harmonische) Obertonspektrum nachgebildet werden. Die entstehende Schwingung wird durch die Formel (3) vollständig beschrieben, wobei A_r die Amplitude des r -ten Teiltons ist. Die Phase θ ist für jeden Oberton unterschiedlich, bleibt aber aufgrund des ganzzahligen Faktors zwischen T_r (der Periodenlänge des r -ten Obertons) und T_G konstant.

$$(3) \quad s(t) = \sum_{r=1}^R A_r \cdot \cos(\omega_r t + \theta_r), \quad \omega_r = \frac{2\pi \cdot r}{T_G}$$

$$(4) \quad s(t) = \sum_{r=1}^R A_r(t) \cdot \cos(\theta_r(t)) + e(t)$$

Vergleicht man (3) mit der Formel des allgemeinen Modells (4) aus Kapitel B 2.1, stellt man folgende Unterschiede fest:

1. Der stochastische Summand $e(t)$ ist nicht vorhanden.
2. Die zeitabhängige Amplitudenfunktion $A_r(t)$ ist eine Konstante
3. Die im Modell unspezifizierte Phasenfunktion $\theta_r(t)$ ist eine lineare Funktion.

Für die Reproduktion der stochastischen Klanganteile muss man daher die Wavetablesynthese mit einem weiteren Verfahren kombinieren, das diese Anteile erzeugen kann. Ausserdem müssen die zeitvarianten Amplituden der einzelnen Teiltöne durch Überblendung und Hüllkurvenmodulation der Wavetables nachgebildet werden. Zunächst soll sich jedoch auf die ausreichend genaue Reproduktion der deterministischen Klanganteile konzentriert werden.

Die Anzahl der Obertöne R beschränkt sich aufgrund des Shannonschen Abtasttheorems

auf

$$(5) \quad R_{max} < \frac{f_s}{2 \cdot f_G}$$

wobei f_s für die Abtastfrequenz und f_G für die Grundfrequenz ($r = 1$) steht. Bei dem höchsten aufgenommenen Ton der Klarinette (G#5, 1661 Hz) und einer Abtastfrequenz von 48 kHz kommt man auf eine maximale Anzahl von 14 Obertönen. Da die weiteren Obertöne ausserhalb des menschlichen Hörbereichs liegen, stellt dies allerdings keine relevante Einschränkung dar.

B 2.2.1 – Probleme bei der exakten Frequenzbestimmung

Wenn bei der Erstellung der Wavetables auf eine samplegenaue Abstimmung der Schleifenlängen der verschiedenen Dynamikstufen geachtet wird, treten die anfangs beschriebenen Kammfiltereffekte nicht auf, da die Phasenlage konstant bleibt ($\Delta f = 0$) Die Schleifenlänge berechnet sich nach der Formel

$$(6) \quad T_s = \frac{f_s}{f_G}, \quad T_s \in \mathbb{N}$$

wobei (das gerundete) T_s für die Anzahl der Samples steht. Im Falle des Tons F#4 kommt man z.B. auf eine Schleifenlänge von 130 Samples.

Aufgrund der diskreten Abtastung sind nur ganzzahlige Schleifenlängen möglich. Dadurch entsteht ein statistischer Fehler im Vergleich zur reellen Originalperiodenlänge $T_G \in \mathbb{R}$ in Höhe von maximal $\pm 0,5$ Samples. Die Frequenzabweichung in Cent lässt sich nach der Formel (3) berechnen.

$$(7) \quad \Delta f = 1200 \cdot \log_2\left(\frac{T_s}{T_G}\right)$$

Beim Ton F#4 beispielsweise entsteht ein Frequenzfehler von +3,5 Cent. Die größte Abweichung von -23,3 Cent bei der Samplerate von 48 kHz weist der Ton F#5 auf. *Abbildung 16* zeigt die Verteilung des Tonhöhenfehlers der verschiedenen Tonhöhen (unter Verwendung der MIDI-Nomenklatur für die Tonhöhenbezeichnung)

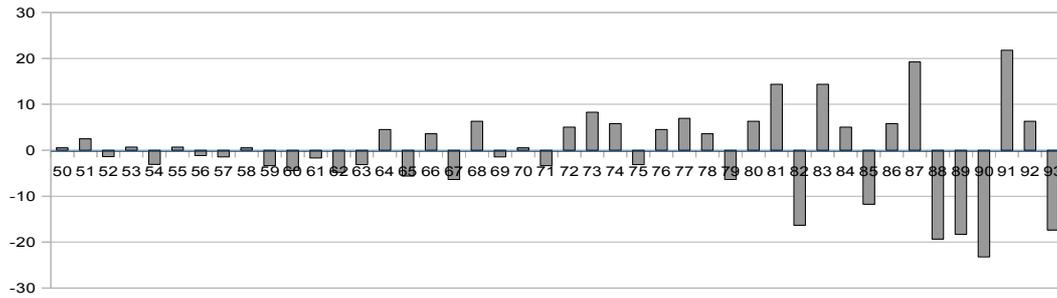


Abbildung 16: statistische Verteilung des Tonhöhenfehlers

Es gibt zwei Möglichkeiten, das Problem zu lösen:

1. Durch eine Erhöhung der Samplerate wird der Fehler minimiert. Bei der maximalen gängigen Samplerate von 192 kHz läge die größte Abweichung nur noch bei 6,3 Cent (beim Ton G4).
2. Durch eine "falsche" Tonhöhenkorrektur der einzelnen Wavetables unter Berücksichtigung des statistischen Fehlers und einer Korrektur des Fehlers bei der Wiedergabe (Abbildung 17) kann der Fehler theoretisch vollständig beseitigt werden, da die tatsächliche Periodenlänge nun exakt mit der ganzzahligen Samplelänge übereinstimmt.

Schritt 1: Wavetableerstellung

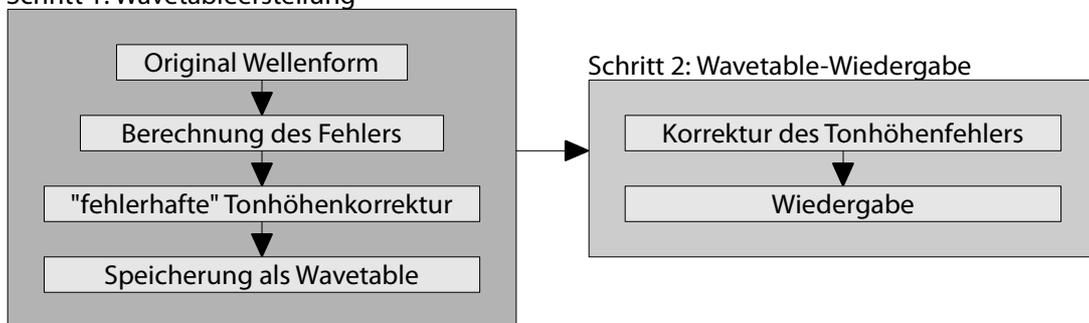


Abbildung 17: schematische Darstellung des zweiten Lösungsmodells

Da Lösung 1 nur eine graduelle Verbesserung bietet und im Normalfall die Tonhöhe der Aufnahmen von den Idealfrequenzen abweichen, und daher sowieso korrigiert werden müssen, wird das zweite Verfahren angewendet. Wenn beispielsweise die Tonhöhe der mezzopiano-Wellenform des F#4-Tons um +20 Cent abweicht, wird der Ton nur um -16,5 Cent nach unten korrigiert ($20 - 3,5$), damit die neue Wellenlänge genau in das ganzzahlige Sampleraster passt. Innerhalb Kontakt wird dann die Tonhöhe der Zone, welche die Wellenform enthält, um weitere 3,5 Cent nach unten korrigiert, um bei der Wiedergabe die Originaltonhöhe zu erreichen. Für die Tonhöhenkorrektur bei der Wellenformerstellung wird auf den Pitchshifting-Algorithmus von Cubase zurückgegriffen.

B 2.2.2 – Erste Umsetzung des Verfahrens

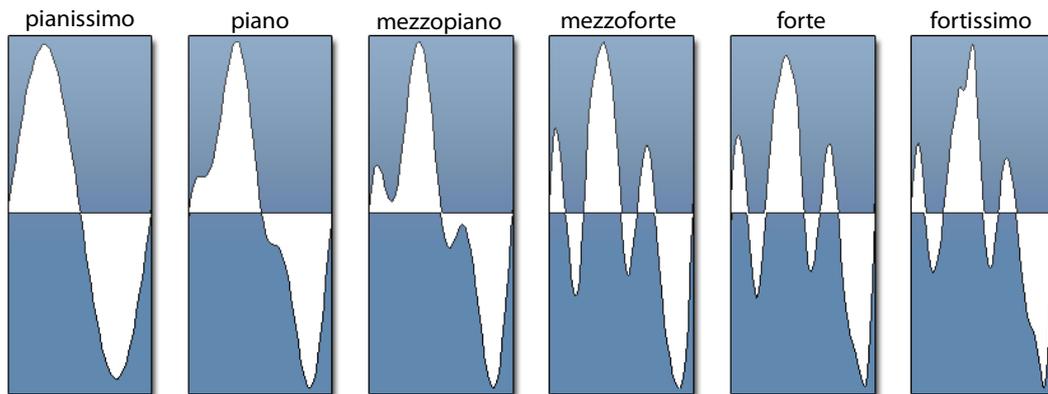


Abbildung 18: visuelle Repräsentation der unterschiedlichen Wellenformen

Abbildung 18 zeigt die verschiedenen Wavetables unterschiedlicher Dynamikstufen des Tons F#3. Die willkürliche Festlegung auf sechs Dynamikstufen, die sich auch durch die weitere Arbeit ziehen wird, reicht aus, um die unterschiedlichen Stadien zufriedenstellend nachzubilden, zumal durch deren Überblendung die tatsächliche Anzahl der Dynamikstufen theoretisch unendlich erhöht wird (praktisch ist man auf den 7bit-Wertebereich des MIDI-Protokolls beschränkt).

Man sieht deutlich die Zunahme der harmonischen Obertöne. Gleichzeitig kann man die konstante Phasenlage erkennen, da die pianissimo-Wellenform aufgrund ihrer Ähnlichkeit zur Sinuswelle als Referenz dienen kann.

Bei der Schleifenwiedergabe der Wavetables stellt man allerdings fest, dass das resultierende Obertonspektrum nicht mit dem zu erwartenden (dem originalen) Spektrum übereinstimmt. *Abbildung 19* zeigt die Spektralcharakteristiken eines realen pianissimo-Tons und der per Wavetablesynthese erzeugten Nachbildung.

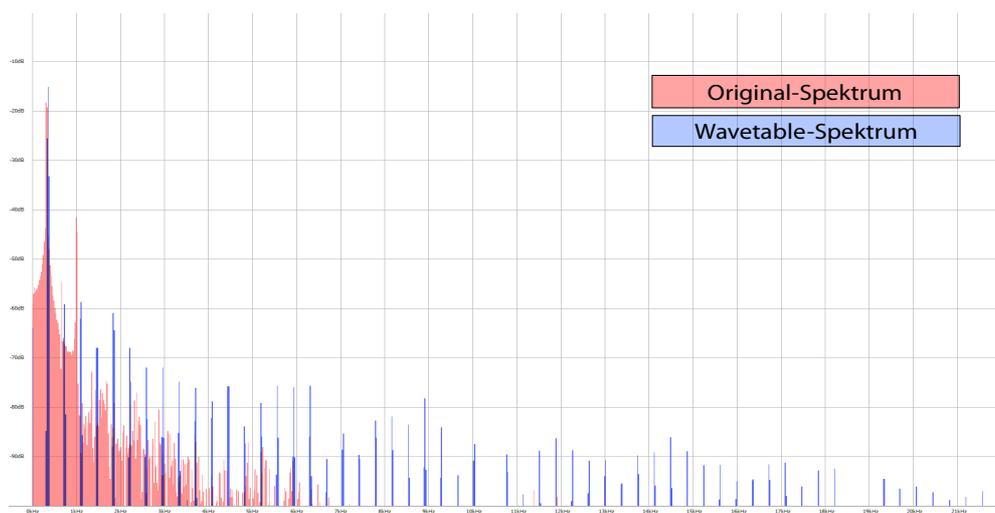


Abbildung 19: Vergleich zwischen Original-Spektrum und Wavetable-Spektrum

Es lassen sich zwei deutliche Unterschiede feststellen:

1. Das Originalspektrum besitzt eine höhere Spektraldichte im Bereich bis 4 kHz.
2. Die Obertöne des Wavetable-Spektrums sind deutlich höher als im Original.

Die höhere Spektraldichte lässt sich damit erklären, dass mit der Wavetablesynthese nur der deterministische Teil des Klanges (d.h. das harmonische Obertonspektrum) nachgebildet werden kann. Frequenzen, die kein ganzzahliges Vielfaches des Grundtons sind, können daher nicht erzeugt werden. Da die Wavetablesynthese jedoch nur zur Reproduktion des deterministischen Klanganteils dient, stellt diese Beobachtung allerdings kein Problem dar, sondern bestätigt nur die Notwendigkeit der getrennten Reproduktion deterministischer und stochastischer Klangelemente.

Das verstärkte Auftreten von Obertönen ist jedoch problematisch, da sie deutlich überhalb des Geräuschniveaus liegen, und daher nicht vom stochastischen Klanganteil maskiert werden können. Der subjektive Klangeindruck bestätigt die Messergebnisse: der Klang ähnelt eher einer Sägezahn-Klang als einem Sinuston.

Die Obertöne entstehen aufgrund den stochastischen Anteilen im Ausgangsmaterial der Wavetables. Die Luftgeräusche (= Rauschen) bewirken eine zufällige Abweichung der realen Wellenform von der idealen (deterministischen) Wellenform. Wird nun eine Periodenlänge in einer Schleife wiedergegeben, verwandeln sich diese vormals stochastisch auftretenden Abweichungen in einen Teil der sich periodisch wiederholenden Wellenform, wodurch die ungewünschten Obertöne entstehen. Um eine zufriedenstellende Anwendung des Verfahrens zu ermöglichen, muss daher nach einer Lösung dieses Problems gesucht werden.

B 2.2.3 – Aufhebung der Diskontinuität an Schleifenübergängen

Zunächst wird die Fensterfunktion der Wavetables geändert. Die Schleifenwiedergabe einer einzelnen Periode entspricht einer rechteckigen Fensterfunktion (8) (Abb. 20 links).

$$(8) \quad f(t) = \begin{cases} 1 & , \quad 0 < t < T_s \\ 0 & , \quad t < 0 \vee t > T_s \end{cases}$$

Da man aufgrund der stochastischen Klanganteile nicht von einer perfekten periodischen Schwingung ausgehen kann, kommt es beim Übergang der Schleifenwiedergabe zu einer Diskontinuität im Wellenformverlauf, die zur Entstehung harmonischer Obertöne führt [vgl. Bristow-Johnson, S. 6f]. Die Fensterfunktion (9), Abbildung 20 rechts, entspricht einem

Equal-Power Crossfade, bei dem die Leistung konstant bleibt.

$$(9) \quad f(t) = \begin{cases} \cos\left(\frac{\pi}{2T_s} \cdot t\right) & , \quad t < |T_s| \\ 0 & , \quad t > |T_s| \end{cases}$$

Mithilfe dieser Fensterfunktion und der dafür benötigten Verdoppelung der verwendeten Perioden kann die Diskontinuität am Schleifenübergang abgefangen werden.

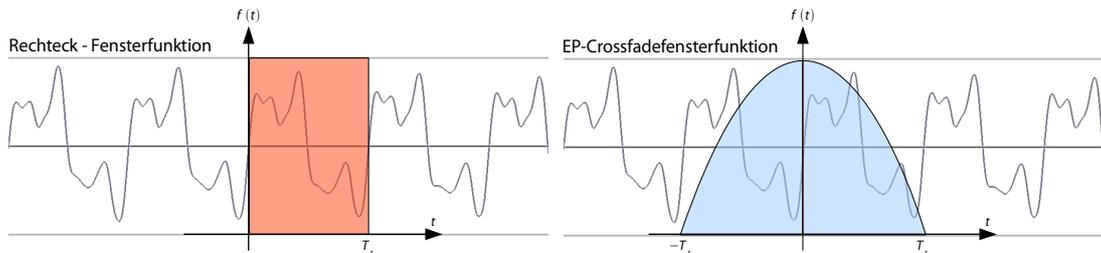


Abbildung 20: Unterschiedliche Fensterfunktionen

Die Implementierung dieser Fensterfunktion in KONTAKT ist vergleichsweise einfach:

Es wird die Anzahl der verwendeten Periodenlängen verdoppelt, und ein Crossfade (mit der Länge einer Periode zwischen beiden Perioden erstellt (Abbildung 21). Da die Tonhöhe innerhalb zwei Perioden nahezu konstant ist, bleibt die Phasenlage stabil, und es treten keine unerwünschten Nebeneffekte auf. Die dritte Wellenform in der Abbildung dient der manuellen Schleifen / Phasenkorrektur, falls die Periodenlängen nicht samplegenau übereinstimmen. Über den Verlauf der Crossfadefunktion wird keine Auskunft gegeben, allerdings deutet das Ergebnis darauf hin, dass ein "Equal-Power"-Crossfade verwendet wird.

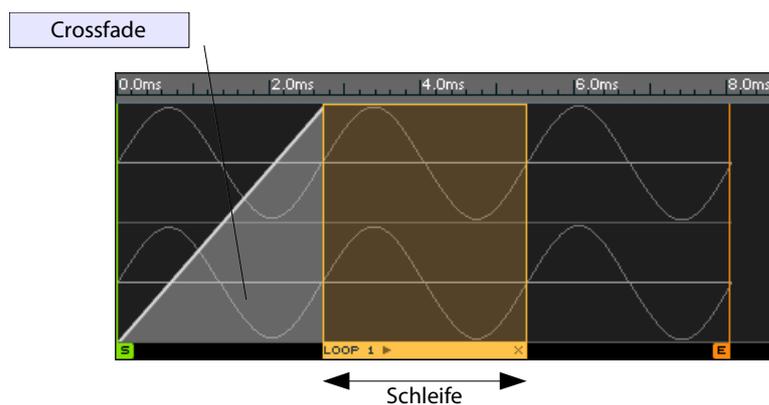


Abbildung 21: Implementierung der Fensterfunktion innerhalb KONTAKT

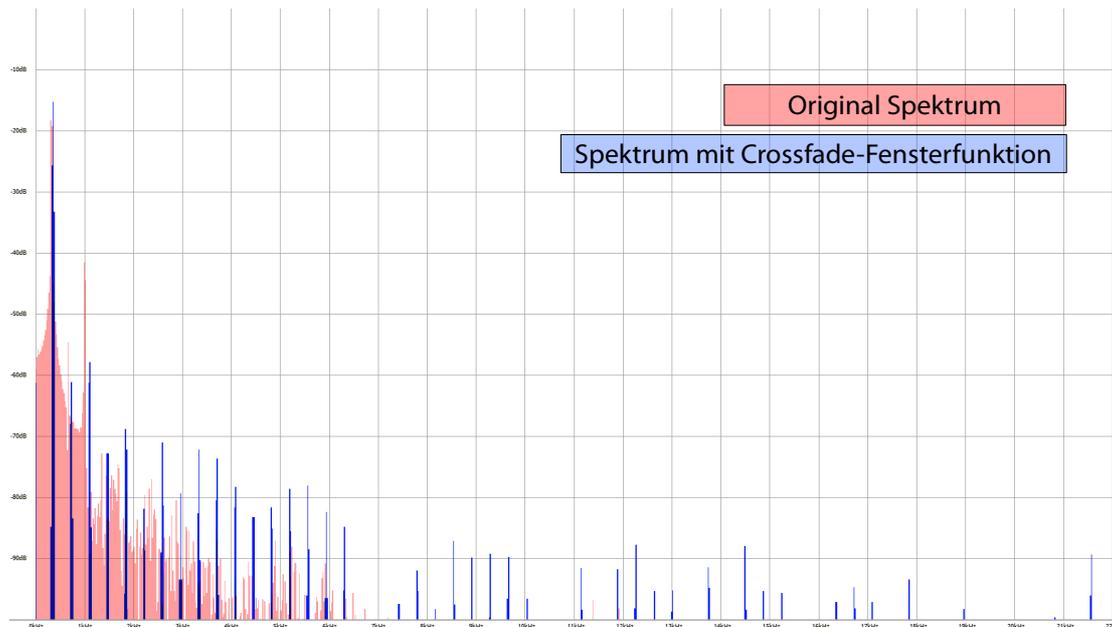


Abbildung 22: Spektralvergleich mit / ohne Crossfade

Abbildung 22 zeigt den Vergleich des Spektrums mit und ohne Crossfade. Im Vergleich zu *Abbildung 19* ist schon ein deutlicher Abfall der Obertöne zu beobachten. Dennoch ist das Klangergebnis noch nicht zufriedenstellend, da die verbleibenden Obertöne immer noch einen synthetischen Klangcharakter erzeugen. Es bedarf daher eines weiteren Verfahrensschrittes, um das natürliche Obertonspektrum zu erhalten.

B 2.2.4 – Superposition mehrerer Perioden

Das Problem bei dem angeführten Verfahren besteht in den stochastischen Anteilen des Klanges, die durch das Verfahren in deterministischen Anteile umgewandelt werden (und neue Obertöne erzeugen).

Das Ziel ist demnach eine Reduktion der stochastischen Anteile. Dies wird ermöglicht durch die Überlagerung mehrerer aneinander angrenzenden Perioden.

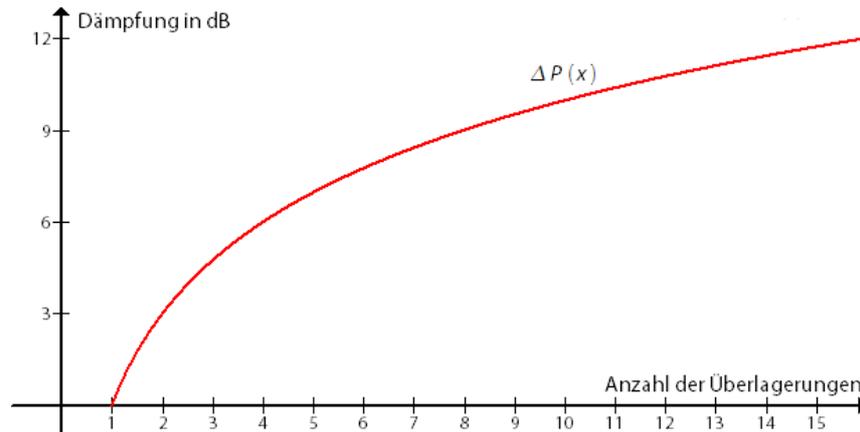
Wie bereits erwähnt, bleibt die Tonhöhe nahezu konstant, wenn die Periodensegmente aufeinanderfolgend ausgewählt werden. Daher können ohne Phasenauslöschungen mehrere Perioden überlagert werden. Man kann davon ausgehen, dass die deterministischen Anteile kohärent sind (sofern die Phase übereinstimmt), während die stochastischen Anteile per Definition inkohärent sind. Die Gesetze der Pegeladdition besagen, dass sich bei einer Addition kohärente Pegel um 6dB verstärken, während inkohärente Pegel nur um 3dB zunehmen [vgl. Runow 2008, S.19]. Diesen Umstand kann man sich zu Nutze machen, um das Verhältnis der stochastischen Elemente zu reduzieren: Mit einer Überlagerung zweier Wellenformen nimmt der deterministische Klanganteil um

6dB zu, während die stochastischen Elemente nur mit 3dB verstärkt werden. Dadurch wird das Verhältnis der stochastischen Klanganteile um 3dB verringert. Um die nächste Verringerung um 3dB zu erreichen, müssen allerdings schon 4 Wellenformen überlagert werden.

Mathematisch ausgedrückt verhält sich der Pegelabstand (11) nach der Formel (10)

$$(10) \quad \Delta P(x) = \log_2(x) \cdot 3\text{dB}, \quad x \in \mathbb{N}$$

$$(11) \quad \Delta P = P_{det} - P_{stoch}$$



wobei x die Anzahl der überlagerten Wellenformen und $\Delta P(x)$ die Dämpfung der stochastischen Anteile beschreibt. Für die Produktion der Klarinette werde ich mich auf eine vierfache Abtastung beschränken, da 6dB Dämpfung im praktischen Anwendungsfall ausreichend sind.

B 2.2.5 – Experiment zur Überprüfung der Effektivität

Bevor das Verfahren auf die Klarinettenwavetables angewendet wird, bietet sich eine empirische Überprüfung der aufgestellten Formel anhand eines Testsignals an. Mithilfe eines Tongenerators wurde eine Sinuswelle mit 100 Hz und -6dB Pegel erzeugt und mit -36dB weißem Rauschen überlagert. Als nächster Schritt wurden 8 Stücke mit jeweils 3 Periodenlängen extrahiert, und phasengenau übereinandergelegt (nach einer Normalisierung auf -18 dB, um Clipping zu vermeiden). *Abbildung 5* zeigt die Auswirkungen dieses Schrittes anhand der Originalwellenform und der überlagerten und normalisierten Wellenform (Wellenformdarstellung aus Cubase). Der Verlauf der Sinuswelle erscheint deutlich geglättet.

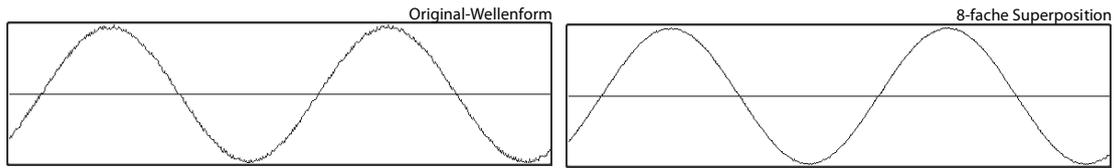


Abbildung 23: Auswirkungen der Mehrfachabtastung auf Wellenformbild

Nach der Formel (1) müsste bei einer achtfachen Überlagerung mit einer Dämpfung von 9dB zu rechnen sein. Um dies zu überprüfen, bietet sich eine FFT-Analyse beider Wellenformen an (Abbildung 6)

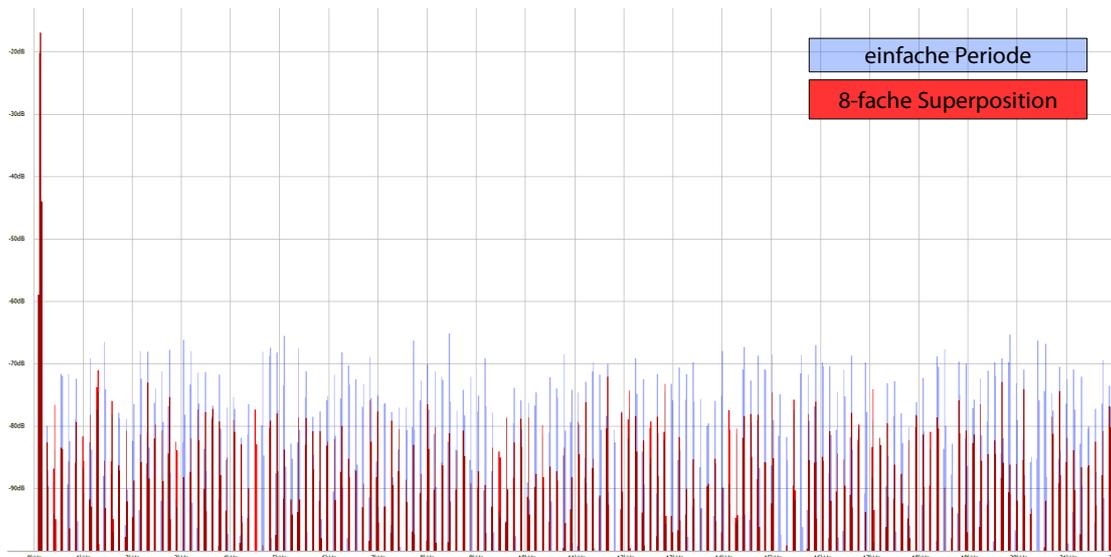


Abbildung 24: Vergleich der Spektralverläufe von Original- und überlagertem Sinusschwingung

Auf den ersten Blick lässt sich eine deutliche Dämpfung der Obertöne vernehmen. Eine genaue Angabe über den Dämpfungswert lässt sich aber nicht ablesen. Abbildung 7 zeigt die Pegeldifferenzfunktion der einzelnen Teiltöne $\Delta A(r) = A_r - A'_r$, die durch visuelle Auswertung der Spektren aus Abbildung 6 entstanden sind. (A_r steht für die Amplitude des r-ten Teiltons ohne, und A'_r für die Amplitude mit Superposition)

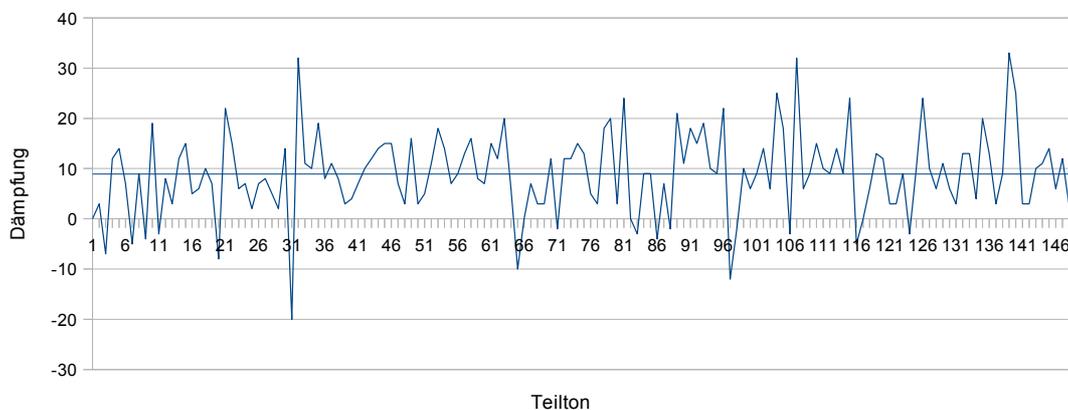


Abbildung 25: Pegeldifferenz zwischen einfacher und achtfach überlagertem Periode

Die teiltonabhängige Dämpfung besitzt einen klar erkennbaren stochastischen Charakter, was darin begründet ist, dass die neu entstandenen Obertöne der nicht überlagerten Wellenform ebenfalls von stochastischen Einflüssen stammen. Eine Mittelwertberechnung der Funktion $\Delta A(t)$, und damit das Äquivalent zu einem Rauschpegel ergibt eine durchschnittliche Dämpfung von $8,92 \text{ dB}$, was einer $0,8\%$ Abweichung von dem erwarteten Wert entspricht. Dieses Ergebnis kann somit als Bestätigung der Überlegungen gedeutet werden.

Während die stochastisch begründeten Obertöne wirkungsvoll gedämpft werden, ist die Auswirkung auf deterministische Obertöne (d.h. Im Originalklang tatsächlich vorkommende Obertöne) lediglich eine Mittelwertbildung \bar{A}_r der Obertonamplituden (12) bei einer Anzahl von x Interpolationen (unter der Voraussetzung, dass die Phaselage konstant ist)

$$(12) \quad \bar{A}_r = \frac{\sum_{i=1}^x A_r(x)}{x}$$

B 2.2.6 – Anwendung auf Klarinetten-Wavetables

Das Verfahren soll nun auf den anfangs verwendeten pianissimo-Ton der Klarinette angewandt werden, um die Effektivität der Superposition zu überprüfen. *Abbildung 7* zeigt das bekannte Originalspektrum, diesmal jedoch im Vergleich mit der vierfach überlagerten und mithilfe der Fensterfunktion (9) geglätteten Wellenform.

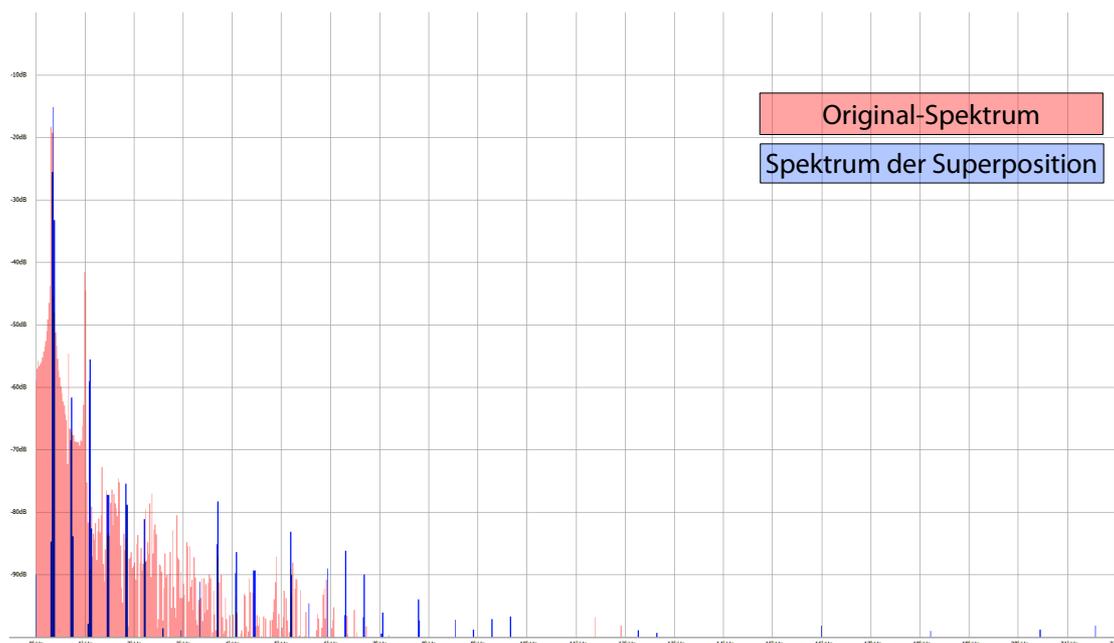


Abbildung 26: Vergleich zwischen Original- und überlagertem Spektrum

Nun kann von einer zufriedenstellender Reproduktion der Obertonstruktur gesprochen werden. Die übrig gebliebenen Obertöne können im Verlauf der Klarinettenproduktion von den stochastischen Klanganteilen ausreichend maskiert werden. Es sollte noch erwähnt werden, dass die Auswirkungen der Interpolation auf pianissimo-Töne am stärksten ist, da kaum natürliche Obertöne vorhanden sind, und daher die künstlich entstehenden Teiltöne deutlicher hörbar werden als bei lauten Tönen, die schon von sich aus eine ausgeprägte Obertonstruktur besitzen. Dennoch profitieren sämtliche Dynamikstufen von der Dämpfung der stochastischen Anteile. Ob die natürlichen Obertöne durch das Verfahren tatsächlich nicht beeinträchtigt werden, soll im letzten Teil der Arbeit mithilfe einer Analyse des fertigen Klarinettenklanges untersucht werden.

Obwohl mit dieser speziellen Form der Wavetablesynthese der Grundstein für eine adäquate Nachbildung realer Musikinstrumente gelegt wurde, müssen die anfangs erwähnten stochastischen Klanganteile separat erzeugt werden. Ausserdem muss das aufgrund der konstanten Obertonamplituden etwas statische Klangbild organischer gestaltet werden. Die Möglichkeiten der KONTAKT-Software bieten für beide Aufgabenstellungen das passende Werkzeug: Der stochastische Klanganteil wird durch Sampling originaler Nebengeräusche bereitgestellt. Die Dynamik im Obertonspektrum sowie das Klangverhalten der Klarinette werden mithilfe komplexer Modulationsstrukturen unter Verwendung von Hüllkurven, MIDI-Daten, und LFOs nachgebildet (siehe Kapitel B 2.4).

B 2.2.7 – Dokumentation der Arbeitsschritte zur Erstellung der Wavetables

Für eine angemessene Reproduktion der spektralen Dynamik ergaben sich sechs Dynamikstufen als ausreichend. Aufgrund der Formantbildungen in den höheren Registern und der spektralen Variation benachbarter Halbtöne werden von jedem Halbton separate Wavetables erstellt, um dem Klangbild der Klarinette gerecht zu werden. Bei einem Tonumfang von 3,5 Oktaven (D2 - G#5) kommt man somit auf 258 Wavetables, die manuell erstellt werden müssen.

Zunächst wurden ca. 10 Sekunden lange Crescendo-Töne auf jeder Halbtonstufe aufgezeichnet. Diese Aufnahmen dienen als Ausgangsmaterial für die Erstellung der Wavetables. Um den hohen Zeitaufwand zu minimieren, müssen die Arbeitsschritte systematisiert werden:

Manuelles Festlegen der unterschiedlichen Dynamikstufen

Nach subjektivem Höreindruck werden aus dem kontinuierlichen Crescendo sechs kurze Stücke herausgeschnitten, die den jeweiligen Dynamikstufen entsprechen. Die dabei entstehenden Unregelmäßigkeiten (da für jeden Ton unterschiedliche Ausschnitte verwendet werden) sind dabei nicht störend, sondern fördern eher den natürlichen Klangcharakter der Klarinette.

Manuelles Korrigieren der Tonhöhe

Die Tonhöhe der pianissimo-Dynamikstufe weist bis zu 35 Cent von der lautesten Stufe ab. Die einzelnen Ausschnitte müssen daher einzeln in der Tonhöhe korrigiert werden. Durch Berechnung des Fehlers als Quotient zwischen realer und idealer Schleifenlänge (unter Berücksichtigung des im vorherigen Kapitel erwähnten statistischen Fehler) und der Tonhöhenkorrektur-Funktion kann die Tonhöhe auf ± 1 Cent genau eingestellt werden.

Erstellung der Wavetables

Nun werden pro Dynamikstufe 4 Stücke mit jeweils 3 aufeinanderfolgenden Periodenlängen erstellt. Bei höheren Registern werden sechs Periodenlängen verwendet, da bei zu kleinen Dateien Fehler bei der Exportierung entstehen. Die Ausschnitte werden überlagert und auf -12dB normalisiert (sodass sich die deterministischen Anteile auf genau 0dB addieren). Die systematische Dateibenennung beim Export lautet "Tonhöhe_Dynamikstufe.wav", sodass bei einer alphabetischen Ordnung der Dateien die einzelnen Tonhöhen zusammen sortiert werden.

Import in KONTAKT

Die Samples werden nun in KONTAKT importiert. Die einzelnen Dynamikstufen werden in sechs Gruppen (pp, p, ...) organisiert, um durch Crossfade-Modulationen die Pegel der einzelnen Stufen steuern zu können. Die kurzen Ausschnitte werden in einer überblendeten Schleife wiedergegeben (siehe *Abbildung 21*). Als letzter Schritt werden mithilfe eines Stimmgeräts und einer konstante Tonhöhenänderung der sechs Dynamikstufen auf Zonenebene der Tonhöhenfehler jedes Halbtons ausgeglichen.

B 2.3 – Nachbildung der stochastischen Klanganteile

Nach den deterministischen Klanganteilen, die im letzten Kapitel behandelt wurden, soll nun versucht werden, die noch fehlende stochastische Komponente $e(t)$ nachzubilden. Das von Serra vorgeschlagene Verfahren der Subtraktion der per additiver Synthese erzeugten Teiltöne vom Gesamtklang und der Analyse des residualen Klangmaterials liefert bei der Resynthese realer Melodiephrasen gute Ergebnisse. Für eine Anwendung für die virtuelle Instrumentenproduktion ist dieses Konzept allerdings nicht gewinnbringend, da sich die zeitliche Veränderung des Leistungsdichtespektrums des residualen Anteils als zu komplex für eine Echtzeitkalkulation darstellt. Deswegen wird die Technik des Samplings angewendet, um die Nebengeräusche zu reproduzieren. Dabei werden die Geräusche getrennt von den Tönen aufgenommen. Dies setzt die Möglichkeit voraus, die Geräusche zu erzeugen, ohne das physikalische System (in diesem Fall die Klarinette) zur Schwingung anzuregen: indem das Rohrblatt durch Druckanwendung an der freien Schwingung gehindert wird, können beispielsweise die Luftgeräusche erzeugt werden, ohne dass ein Ton gespielt wird. Nach eingehender Analyse des Klangbildes einiger Melodiephrasen wurden die wichtigsten Geräusche einzeln aufgenommen. *Tabelle 1* zeigt einen Überblick über die Aufnahmen der verschiedenen Geräusche:

Erläuterung der einzelnen Geräusche

Kategorie	Art	Anzahl
<i>Luftgeräusche</i>	Luftgeräusch	9
	Nebenluftgeräusche	5
	kurze Luftimpulse	18
<i>mechanische Geräusche</i>	Klappengeräusche	42
	tonale Klappengeräusche	12
<i>Nebengeräusche</i>	Atmungsgeräusche	50
	Mundgeräusche (Schmatzen)	9
	Bewegungsgeräusche	2
<i>sonstige Geräusche</i>	Multiphonics	11

Tabelle 1: Überblick der Geräuschaufnahmen

Luftgeräusche: Das Geräusch der turbulenten Luftströmung durch den Korpus der Klarinette. Da unterschiedliche Klappenstellungen Einfluss auf das Geräusch haben, wurden mehrere Geräusche aufgenommen. Da der Anteil der Luftgeräusche bei leisen Tönen lauter ist, wurden die Geräusche relativ leise aufgenommen.

Nebenluft: Bei lauten Tönen kann der hohe Druck im Mundraum des Spielers zu Luftentweichungen ausserhalb der Klarinette führen. Das entstehende Luftgeräusch weicht von dem obigen Geräusch ab, und wird daher separat aufgenommen. Für eine ausreichende klangliche Variation wurde eine Anzahl von 5 Samples aufgenommen.

Kurze Luftimpulse: Für einen realistischen Tonansatz bei schnellen Staccatopassagen muss die Einschwingphase des Luftgeräuschs separat modelliert werden. Dazu werden kurze Luftimpulse (impulsive "T"-Laute) aufgenommen, um später in sechsfacher Round-Robin Variation abgespielt zu werden.

Klappengeräusche: Bei der Analyse von Tonübergängen bei Melodiephrasen stellt man fest, dass das Geräusch der sich schliessenden Klappe einen hörbaren Anteil am Klang des Intervallübergangs besitzt. Diese Geräusche werden daher in hoher Zahl aufgenommen, da die Prominenz des Klangs eine hohe Variation erfordert.

Tonale Klappengeräusche: Zusätzlich zu den Klickgeräuschen kommt es bei manchen Intervallübergängen zu einem tonalen Klappengeräusch. Der Grund dafür liegt in dem Luftimpuls der durch das Schliessen der Klappe in den Korpus gelangt, und die Klarinette für kurze Zeit in Schwingung versetzt. Um dieses Geräusch zu imitieren, wurde durch extreme Filterung der Klappengeräusche, namentlich durch eine schmalbandige 48dB-Verstärkung des jeweiligen Grundtons, der tonale Charakter verstärkt.

Multiphonics: Multiphonics ist eine Spieltechnik, bei der die Klappen der Klarinette nur halb geschlossen werden. Dadurch entstehen mehrere Luftsäulen innerhalb der Klarinette, was zu einer gleichzeitigen Ausbildung unterschiedlicher stehender Wellen führt. Multiphonics werden zum Großteil in der Neuen Musik verwendet. Bei der Analyse der Testaufnahmen stellte sich jedoch heraus, dass bei etwas unsauberer Spielweise in den Intervallen und bei manchen Tonansätzen kurzzeitig Multiphonics-ähnliche Klänge entstehen. Um dies bei der virtuellen Reproduktion zu berücksichtigen, wurden einige beliebige Multiphonicstöne aufgenommen, die nun für die Programmierung der Intervalle und des Tonansatzes zur Verfügung stehen.

B 2.4 – Reproduktion des natürlichen Klangverhaltens

Mithilfe des im letzten Kapitel entwickelten Verfahren lassen sich die spektralen Unterschiede verschiedener Dynamikstufen ausreichend genau und ohne Kammfiltereffekte nachbilden. Das elementare Problem der Sampling-Technologie kann somit umgangen werden. Im Vergleich zur herkömmlichen Wavetable-Synthese können auch stochastische Klanganteile nachgebildet werden. Allerdings ist damit nur der Grundstein für eine natürliche Reproduktion realer Instrumente gelegt: Das Klangverhalten in der Einschwingphase (Tonansatz), bei Tonübergängen (Intervalle), sowie die Implementierung verschiedener Spieltechniken wie Vibrato oder Growl müssen manuell nachmodelliert werden. Das Ergebnis und die Klangqualität hängen dabei in erster Linie vom Geschick und der ästhetischen Vorstellung des Programmierers ab, was interessanterweise eine Parallele zur Fertigung realer Musikinstrumente darstellt.

Bevor die einzelnen Klangkomplexe detailliert beschrieben werden, soll ein kurzer Überblick über die verwendeten Modulatortypen gegeben werden. Dabei soll sich auf die Auflistung der wichtigsten Modulatoren beschränkt werden.

MIDI – Daten

Die MIDI – Daten stellen als Eingabeprotokoll die Interaktion zwischen dem Musiker und der Software her. *Abb. 27* zeigt schematisch das Konzept der Instrumentenbedienung.

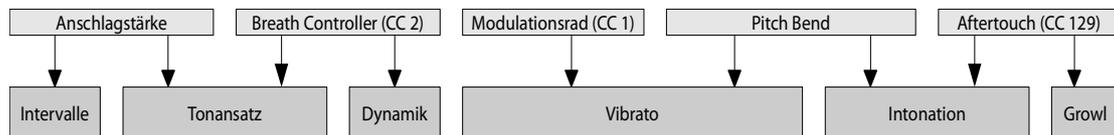


Abbildung 27: Verarbeitung der MIDI-Daten

Während die Anschlagstärke bei Tasten und Schlaginstrumenten die wichtigste Größe zur Steuerung der Dynamik ist, wird sie im Falle der Klarinette nur für die Kontrolle des Tonansatzes und für die Steuerung der Intervallübergänge eingesetzt.

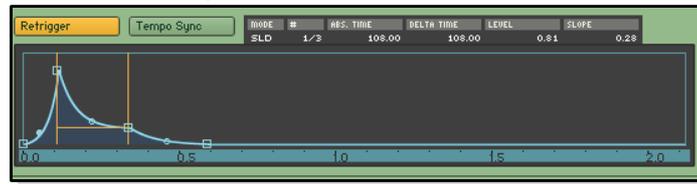
Die Dynamik (und damit hauptsächlich die Überblendung zwischen den einzelnen Wavetables und die Lautstärke der Luftgeräusche) soll nur vom Breathcontroller gesteuert werden. Tatsächlich wurde noch eine Vielzahl anderer Modulatoren eingesetzt, um das Klangbild natürlicher zu gestalten.

Die Kontrolle des Vibrato (einer relativ langsamen, periodischen Modulation von Tonhöhe und Lautstärke) soll mithilfe der beiden Steuerrädern, die sich links neben der Tastatur befinden, vollzogen werden, wobei das Modulationsrad die Intensität eines LFOs (siehe unten) steuert, während das Pitchwheel eine manuelle Kontrolle des Vibratos erlaubt.

Durch Druck auf die Tastatur (Aftertouch) soll die Spieltechnik des Growlings zusammen mit einer Änderung der Intonation (ca. max -10 Cent) gesteuert werden. Dadurch kann die Expressivität gesteigert werden.

Hüllkurven

Flexible Envelope



AHDSR - Envelope



Abbildung 28: Hüllkurventypen in KONTAKT

Hüllkurven ermöglichen die zeitliche Modulation eines Parameters nach vorgeschriebenem Verlauf. In KONTAKT gibt es zwei verschiedene Hüllkurventypen (siehe Abb. 28), die für unterschiedliche Aufgaben eingesetzt werden können. Die herkömmliche AHDSR-Hüllkurve ist für die meisten Aufgaben ausreichend. Für die Reproduktion des Einschwingverhaltens beispielsweise wurde jedoch ein "Flexible Envelope" eingesetzt, um die Intensität des Tonansatzes unabhängig vom Sustain-Pegel zu steuern (siehe B 2.4.2)

LFOs

LFO ("Low Frequency Oscillators"), sind Schwingungsgeneratoren, die für akustische Verhältnisse langsame Schwingungen erzeugen⁶. Diese Schwingung kann zur Modulation beliebiger Parameter verwendet werden. Das prominenteste Beispiel ist die Tonhöhenmodulation per LFO zur Reproduktion der Vibrato-Spieltechnik. Es stehen verschiedene Wellenformen (Sinus, Dreieck, Sägezahn) aber auch Rauschen⁷ zur Verfügung.

Weitere Modulatoren

Neben den erwähnten Modulatoren gibt es noch eine Vielzahl anderer Möglichkeiten, den Klang zu beeinflussen, wie z.B. Zufallsgeneratoren (nicht zu verwechseln mit einem rauschenden LFO, da nur pro angeschlagener Taste ein neuer Wert erzeugt wird), die

⁶ Bei KONTAKT ist der Frequenzbereich 0.01 Hz – 213 Hz)

⁷ In diesem Fall steht die bei Rauschen ansonsten sinnlose Angabe der Frequenz für die Geschwindigkeit, mit der neue Zufallswerte erzeugt werden, d.h. bei 5 Hz werden alle 200ms neue Zufallswerte erzeugt.

Tonhöhe (sogenannte "Key Position") und Konstanten.

Für die Modellierung des Klangverhaltens der Klarinette wurden insgesamt 176 solcher Modulatoren eingesetzt⁸. Anhang A2 enthält die vollständige Modulatorkonfiguration der verschiedenen Gruppen. Im Verlauf des Kapitels wird die Reproduktion der einzelnen Klangparameter auch unter Berücksichtigung etwaiger Skripte dokumentiert.

B 2.4.1 – Dynamik

Für die Reproduktion der Dynamik werden die Breathcontrollerdaten zur Überblendung der Wavetables verwendet (Abbildung 29), die in sechs Gruppen geordnet sind.

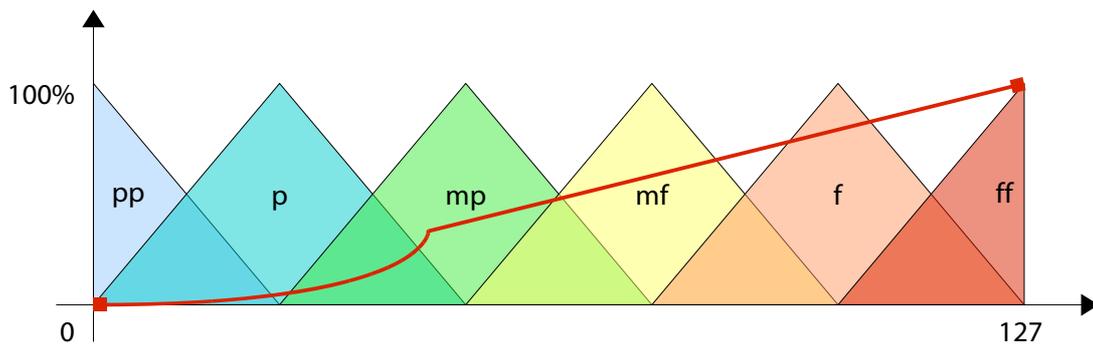


Abbildung 29: Überblendung der mit Breath-Controllerdaten

Die farbigen Dreiecke stehen für die unterschiedlichen Dynamikzonen. Die rote Linie beschreibt einen zusätzlichen Modulator, der auf allen sechs Dynamikgruppen gleich verläuft, und eine Lautstärkeanpassung der (normalisierten) Dynamikstufen bewirkt (sodass leise Dynamikstufen auch einen geringeren Pegel aufweisen). Dabei wurde sich an der natürlichen Dynamik der Crescendosamples (ca. 20dB) orientiert. Der anfangs exponentielle Verlauf ist für die Nachbildung des Tonansatzes (siehe unten) von hoher Bedeutung: *Abbildung 30* zeigt diesen Verlauf und den Verlauf des Modulators der Luftgeräusch-Gruppe ("stochastischer Anteil")

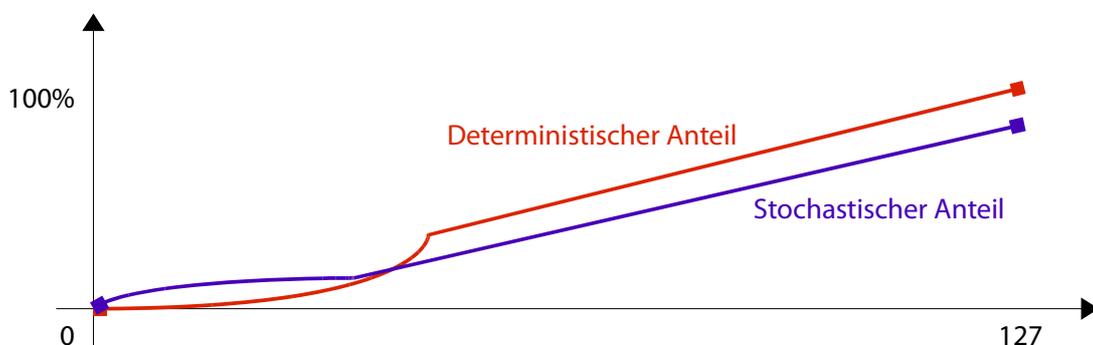


Abbildung 30: Vergleich zwischen Breathcontroller-Verläufen

Im ersten Viertel des Breathcontrollerwertebereichs ist das Luftgeräusch lauter als der

⁸ die Modulationen der programmierten Skripte nicht mit eingerechnet

deterministische Anteil. Dadurch wird die Trägheit der Luftröhre im Klarinettenkorpus nachgebildet, da der Blasdruck zur Ausbildung der stehenden Welle (und damit des Tones) einen gewisser Schwellwert überschreiten muss. Bei langsamen Tonansätzen ist dies deutlich wahrnehmbar, sodass bei der virtuellen Reproduktion der Ton ebenfalls erst nach Überschreitung einer gewissen Schwelle erklingt. In diesem Beispiel ist der Ton ca. ab dem Breathcontrollerwert 20 hörbar. Da die Implementierung dieses Phänomens den für die Dynamikkontrolle an sich schon zu geringen verwendbaren Wertebereich nochmals verkleinert (bei einem Schwellwert von 20 bleiben noch 108 Stufen), muss bei der Wahl der Kurvenverläufe ein Kompromiss gebildet werden.

Zusätzlich zum normalen Luftgeräusch werden bei hohen Breathcontrollerwerten die Nebenluftgeräusche abgespielt, die zu einer zusätzlichen Klangänderung bei lauten Tönen führt.

B 2.4.2 – Tonansatz

Durch die Verwendung des Breath Controllers entstehen zwei Tonansatztypen:

Möglichkeit 1: Es wird bereits vor dem Tastendruck in den Breathcontroller geblasen. In der weiteren Arbeit wird diese Art des Tonansatzes "Tastaturansatz" genannt, da der Ton erst bei einem Tastendruck entsteht.

Möglichkeit 2: Zuerst wird die gewünschte Tonhöhe auf der Tastatur gedrückt und dann in den Breath-Controller geblasen (fortan "Breathcontrolleransatz" genannt). Der Ton soll dabei erst erklingen, wenn der Wert des Breath-Controllers einen von Null verschiedenen Wert annimmt.

Der Tastaturansatz ermöglicht eine schnellere Tonwiederholung, da die Breathcontrollerdaten geglättet werden müssen, um Klangartefakte⁹ zu vermeiden. Der Breathcontrolleransatz dagegen eignet sich für langsame Tonansätze in Legatophrasen.

Die in dieser Arbeit entwickelten Art der Wavetablesynthese bietet von sich aus keine Berücksichtigung des speziellen Klangverhaltens des Instruments in seiner Einschwingphase. Gleichzeitig ist genau dieser kurze Zeitabschnitt (die ersten 100 – 200ms des Tons) aufgrund seiner starken Charakteristik für die Erkennung des jeweiligen Instrumententyps von essentieller Bedeutung. Die korrekte Reproduktion des Klangverhaltens am Tonansatz ist daher die kritischste Aufgabe der Modulatorprogrammierung.

⁹ der sogenannte "Reißverschluss effekt", der bei Wertesprüngen aufgrund der zu geringen Abtast-rate und -auflösung der MIDI-Daten entsteht.

Sollte eine einigermaßen realitätsgetreue Nachbildung der Einschwingphase gelingen, stößt man auf das nächste Problem, das auch von herkömmlichen Sampleinstrumenten gut bekannt ist: bei (schnellen) Tonwiederholungen macht sich der statische Charakter der virtuellen Tonansätze bemerkbar¹⁰. Um dies zu beheben, ist man auf die Anwendung weiterer Modulatoren / Verfahren angewiesen, die die Variation im Tonansatz erhöhen. Zunächst soll sich jedoch auf die Reproduktion eines Tonansatzes konzentriert werden.

Modellierung des "Tastaturansatzes" mithilfe Hüllkurven

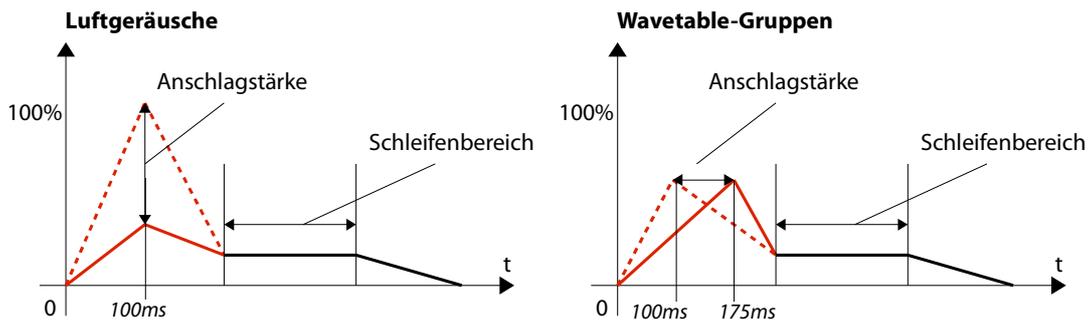


Abbildung 31: Unterschiedliche Hüllkurven der Klanganteile

Abbildung 31 zeigt die Hüllkurven, die getrennt auf die jeweiligen Klanganteile angewendet werden. Um der Komplexität des Einschwingvorgangs Rechnung zu tragen, wird der "Flexible Envelope" verwendet, dessen Parameter von der Anschlagstärke moduliert werden: Bei den Luftgeräuschen (Abb. 31 links) bestimmt die Anschlagstärke die Intensität des ersten Maximums, dessen Zeitpunkt nach subjektivem Einschätzen auf ca. 100ms festgelegt wurde. Je stärker eine Taste gedrückt wird, desto lauter ist demnach der erste Luftimpuls. Bei den identischen Hüllkurven der sechs Dynamikgruppen (Abb. 31 rechts) wird dagegen die Zeit bis zum ersten Maximum von der Anschlagstärke moduliert: je höher der Velocity-Wert, desto kürzer ist die Zeitspanne zum ersten Maximum (maximal 175ms). Der Minimalwert der Zeit entspricht dagegen exakt der Einschwing-Zeitspanne der Luftgeräuschhüllkurve (100ms)

Die Hüllkurven bewirken eine Veränderung des Tonansatzes: Von weichen, langen Tonansätzen zu kurzen, impulsiven Einschwingphasen für Staccato-Passagen kann mithilfe der Anschlagstärke der Charakter des Tonansatzes geformt werden. Die unterschiedlichen Modulationsziele (Intensität bei Luftgeräuschen und Einschwingzeit der Wavetables) erhöhen die Komplexität und damit die Natürlichkeit des Klangs.

Zusätzlich zur Lautstärke-Hüllkurve wird bei den Wavetablegruppen ein Tiefpassfilter mit Cutoff-Hüllkurve eingesetzt, dessen Attack Time von der Anschlagstärke gesteuert wird (in ähnlichen zeitlichen Dimensionen wie die Lautstärkehüllkurve). Dies soll eine grobe

¹⁰ siehe "Machine-Gun"-Effekt, Kapitel 1.2 - Sampling

Nachbildung der spektralen Dynamik in der Einschwingphase bewirken, da leise Töne weniger Obertöne besitzen, und bei einem Tonansatz der Ton aus dem Ruhezustand des physikalischen Systems entsteht. Die ideale Lösung wäre ein Durchlaufen der unterschiedlichen Wavetable-Stufen, allerdings ist dessen Implementierung innerhalb der KONTAKT-Architektur recht aufwändig und würde zu einer Konfliktbildung mit der "normalen" Dynamiksteuerung (Wertesprünge etc.) führen. Aufgrund der kurzen Zeitspanne, in der der Filtereffekt auftritt (namentlich die Attack-Zeit von ca. 100ms), stellt diese Lösung aber keine große Beeinträchtigung der Klangqualität dar.

Verarbeitung der Luftimpuls-Samples

Die aufgezeichneten Luftimpuls-Samples (siehe B 2.3) dienen der Unterstützung der Hüllkurven-Modellierung der Luftgeräusche. Die Anschlagstärke steuert hier die Attack-Zeit eines AHDSR-Envelope sowie die Grundlautstärke, sodass die Wahrnehmbarkeit mit der Anschlagstärke deutlich zunimmt. Gerade bei schnellen Tonwiederholungen ist dieses Geräusch sehr charakteristisch.

Um die Wiedergabe der Luftgeräusch auch beim "Breathcontroller"-Ansatz zu ermöglichen, muss auf die Skriptprogrammierung zurückgegriffen werden. Das folgende Skript hat nur Auswirkung auf die Gruppen der Luftimpulssamples.

```
1  on init
2    declare $voice_count := 0
3    declare $note_on_switch := 1
4    declare $last_note := 50
5    declare const $ATTACK_VELOCITY := 50
6  end on
7  on note
8    inc($voice_count)
9    if ($voice_count > 1 or $note_on_switch = 1)
10     ignore_event($EVENT_ID)
11   end if
12   $last_note := $EVENT_NOTE
13 end on
14 on release
15   wait(50000)
16   dec($voice_count)
17 end on
18 on controller
19   if($CC_NUM = 2)
20     if(%CC[2] # 0)
21       if($note_on_switch = 1 and $voice_count = 1)
22         play_note($last_note, $ATTACK_VELOCITY + random(-5,5), 0, -1)
23         $note_on_switch := 0
24       end if
25       $note_on_switch := 0
26     else
27       $note_on_switch := 1
28     end if
29   end if
30 end on
```

Codebeispiel 2: Implementierung des Breathcontrolleransatzes

Das Skript erfüllt drei Funktionen:

1. Wenn der Breathcontrollerwert = 0 ist, wird bei einem Tastendruck die Note ignoriert (Zeile 10), nachdem die Tonhöhe gespeichert wurde (Zeile 12). Dies verhindert, dass das Luftgeräusch vor dem eigentlichen Tonansatz abgespielt wird.
2. Wenn der Breathcontroller nun einen von Null verschiedenen Wert annimmt (Zeile 20) wird eine künstliche Note mit der gespeicherten Tonhöhe und einer mit einer Zufallsvariablen versehenen konstanten Anschlagstärke $\$ATTACK_VELOCITY$ erzeugt (Zeile 22), sodass der Luftimpuls der jeweiligen Tonstufe abgespielt wird. Obwohl der Luftimpuls keine tonalen Anteile hat, ist er chromatisch auf der Tastatur angeordnet, um die Variation zu erhöhen.
3. Bei einem Intervallübergang sollen die Impulssamples nicht abgespielt werden, da es sich hierbei nicht um einen Tonansatz im klassischen Sinne handelt. Dazu wird die Anzahl der gedrückten Tasten $\$voice-count$ gespeichert (Zeile 8 + Zeile 16). Ist bei einer neu gedrückten Note die Anzahl der Tasten größer als 1 (Zeile 9), so wird die Note ebenfalls ignoriert.

Verarbeitung der Multiphonics-Samples

Bei der Analyse realer Klarinettenaufnahmen stellt man fest, dass bei manchen Tonansätzen kurzzeitig unharmonische Schwingungen entstehen. Um dieses Verhalten zu reproduzieren, werden die aufgezeichneten Multiphonics-Samples verwendet. Da die Bedingungen für die Wiedergabe dieselben sind, wie die Bedingungen für die Wiedergabe der Luftimpuls-Samples, kann das oben angeführte Skript (Codebeispiel 2) ebenfalls auf diese Samplegruppe angewendet werden. Durch Modulation der Lautstärke mit Zufallsgeneratoren wird die Häufigkeit des "unsauberen" Tonansatzes verringert, da ansonsten ein künstlicher Effekt entsteht.

Erhöhung der Variation

Nun sollte der Grad der Variation erhöht werden, um bei schnellen Tonwiederholungen den natürlichen Klangeindruck zu erhalten. Dazu werden vor allem zwei Werkzeuge verwendet:

1. Das vom herkömmlichen Sampling bekannte Round-Robin Verfahren wird auf die Luftimpuls-Samples angewandt (in sechsfacher Ausführung, sodass bei Tonwiederholungen eine feste Reihenfolge von sechs verschiedenen Samples durchlaufen wird).

- Mithilfe der Zufallsmodulation bestimmter klanggebender Parameter (Intensität, Einschwingzeit der einzelnen Hüllkurven, Startpunkt der Luftgeräuschwiedergabe etc.) wird die stochastischen Abweichungen realer physikalischer Systeme vom idealisierten Modell nachgebildet. Die Intensität der Zufallsmodulation ist dabei rein willkürlich und hängt vom musikalischen Geschmack des Programmierers ab. Diese Entscheidung könnte jedoch auch an den Endnutzer übertragen werden, sofern mithilfe der Skriptprogrammierung eine geeignete Benutzeroberfläche erstellt werden würde. Im Rahmen dieser Arbeit wird allerdings nicht näher auf die Thematik der Individualisierung durch den Endnutzer eingegangen.

B 2.4.3 – Tonübergänge

Die Übergangsphase zwischen zwei (gebundenen) Tönen (das sogenannte "Legato-Intervall") ist ähnlich charakteristisch wie der eigentliche Tonansatz. Es existieren teilweise eigene musikalische Spielanweisungen ("Portamento", "Glissando"), die eine spezielle Technik beim Notenübergang fordern. Im Rahmen der Arbeit wird sich allerdings nur auf den "normalen" Tonübergang konzentriert, da dessen realistische Nachbildung kompliziert genug erscheint.

Beim herkömmlichen Samplingverfahren, bei dem nur aufgenommene Einzeltöne wiedergegeben werden, stösst man auf ein typisches Phänomen der Emergenz, da sich die Einzeltöne meist nicht zur einer musikalisch sinnvollen Phrase zusammenfügen lassen. Daher werden bei den aufwändigeren Samplelibraries die Intervallphasen separat aufgenommen, und an der entsprechenden Position abgespielt. Dies resultiert in einer enormen Anzahl an Samples: allein für die Nachbildung einer Dynamikstufe werden bei einem Tonumfang von n Halbtönen theoretisch $n \cdot (n-1)$ Intervallsamples benötigt. Beim Tonumfang der Klarinette (ca. 43 Halbtöne) käme man so auf ca. 1800 Samples, weshalb die meisten Hersteller die aufgezeichneten Intervalle auf ± 1 Oktave eingrenzen.

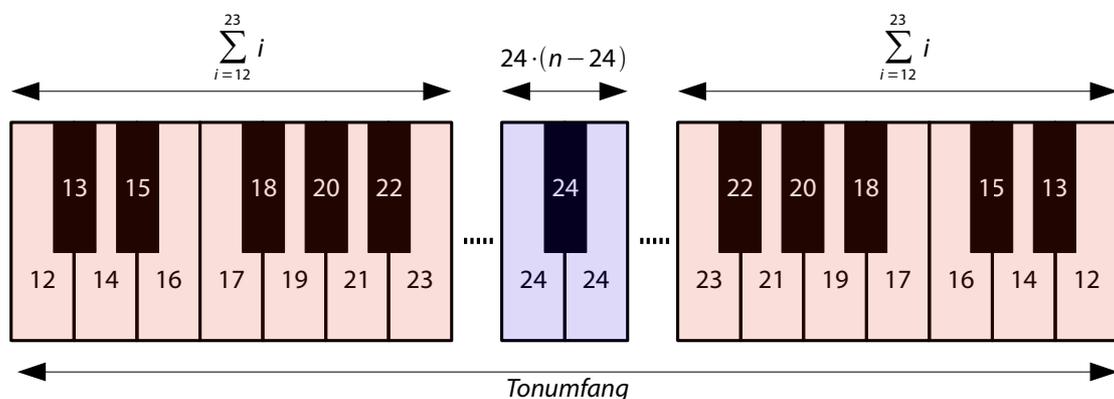


Abbildung 32: Anzahl der Intervallsamples pro Ton

Wie aus *Abbildung 32* hervorgeht, sinkt dadurch die Anzahl der benötigten Samples auf

$$(13) \quad O(n) = 2 \cdot \sum_{i=12}^{23} i + 24 \cdot (n - 24) = 24 \cdot n - 156, \quad n > 24$$

Bei der Klarinette müssten nach (13) noch ca. 900 Samples aufgezeichnet werden. Die Anzahl kann noch beliebig weiter gesenkt werden (z.B. durch das Weglassen einiger Halbtöne etc.), allerdings wird klar, was für ein Aufwand hinter einer Reproduktion der Tonübergänge per "Legato-Intervall"-Samples steckt.

Dieses Verfahren hat neben dem hohen Realisierungsaufwand noch einen weiteren Nachteil: der Übergang zwischen zwei Tönen klingt jedes Mal gleich, unabhängig von der Dynamikstufe, dem Tempo und dem Charakter der Melodie. Es ist offensichtlich, dass ein Tonübergang in einem schnellen Tonleiterlauf anders klingen muss, als ein Intervallsprung in einer langsamen Legatomelodie.

Das in dieser Arbeit vorgeschlagene Konzept dagegen basiert auf der Analyse der Klangbestandteile eines Tonübergangs und der anschließenden Nachbildung durch Kombination der einzelnen Elemente. Da diese Elemente einzeln vorliegen, kann durch Variation der Kombination unter Verwendung bestimmter MIDI-Daten und Zufallsmodulatoren ein natürliches und musikalisch kontrollierbares Klangverhalten im Tonübergang realisiert werden.

Die beiden wichtigsten Steuergrößen sollen dabei der Dynamikwert (d.h. der Breathcontrollerwert) und die Anschlagsdynamik sein. Der Dynamikwert steuert die Klangunterschiede zwischen leisen und lauten Tonübergängen, während die Anschlagsstärke die Geschwindigkeit und Intensität, mit der die Klappen der Klarinette geschlossen werden, nachbildet.

Zunächst soll jedoch die Abspielbedingungen der Intervallsamples bestimmt werden. Ohne spezielle Programmierung dieser Konditionen würde das Intervallsample bei jeder neuen Note (d.h. auch beim Tonansatz) abgespielt werden, und dessen Klangverhalten verfälschen. Das folgende Skript, das nur auf die Intervallsamplegruppen (siehe unten) angewendet wird, erledigt diese Aufgabe:

```

1  on init
2    declare $voice_count := 0
3    declare $last_note := 0
4  end on
5  on note
6    inc($voice_count)
7    ignore_event($EVENT_ID)
8    if($voice_count > 1)
9      play_note($last_note, $EVENT_VELOCITY, 0, -1)
10   end if
11   $last_note := $EVENT_NOTE
12 end on
13 on release
14   if($EVENT_NOTE = $last_note)
15     $voice_count := 0
16   end if
17 end on

```

Codebeispiel 3: Wiedergabebedingung der Intervallsamples

Die schon aus Codebeispiel 2 bekannte Variable `$voice_count` speichert wieder die Anzahl der gedrückten Tasten (Zeile 6 & 15). Die Variable `$last_note` enthält die Tonhöhe der letzten gedrückten Note (Zeile 11). Standardmäßig wird jede Note zuerst ignoriert (Zeile 7). Dies verhindert, dass die Intervallsamples bei jedem Tastendruck abgespielt werden. An einem gebundenen Übergang (d.h. die alte Taste ist noch gedrückt, während die neue Note gespielt wird) ist `$voice_count` grösser als 1. In diesem Fall (Zeile 8) wird eine künstliche Note mit der alten Tonhöhe `$last_note` erzeugt (Zeile 9), sodass die Intervallgeräusche erklingen.

Unter diesen Wiedergabekonditionen können nun die verschiedenen Intervallgeräusche abgespielt werden. Namentlich sind das Klappengeräusche (tonale und Klickgeräusche getrennt) und kurze Ausschnitte aus Multiphonics-Aufnahmen.

Klappengeräusche

Die nichttonalen Klappengeräusche werden chromatisch in dreifacher Round Robin Ausführung und mit 4 Lautstärkestufen auf die Tastatur gelegt. Dazu wären theoretisch 516 verschiedene Samples nötig. Durch eine Wiederholung der Samples ca. aller 7 Halbtöne und einer Verschiebung bei den Round-Robin Gruppen wurden effektiv nur 42 Samples eingesetzt. Die dadurch erzeugte Variation reicht schon aus, sodass weitere Samples nicht mehr sinnvoll waren.

Zusätzlich zur natürlichen Dynamik (da ja vier Lautstärkestufen verwendet werden) wurde die Lautstärke noch per Anschlagsdynamik moduliert, sodass bei leicht gedrückten Tasten die Klickgeräusche kaum hörbar werden. Ein Zufallsmodulator erhöht die Variation, da bei realen Instrumentenphrasen schätzungsweise nur 60% aller Übergänge einen Klick erzeugen.

Bei den tonalen Klappengeräuschen wurden nur zwei Aufnahmen angewendet: für das

Chalmeau-Register und für das Clarin- und Flageolett-Register zusammen jeweils ein Sample. Der Grund ist folgender: wegen der kurzen Dauer der Samples und der fast rein sinusoidalen Obertonstruktur aufgrund der starken Filterung kommt es nicht zu einer hörbaren Stauchung der Frequenzspektren. Zudem erzeugen die nichttonalen Klappengeräusche einen ausreichend hohen Grad der Variation, sodass hier auf zusätzliche Samples verzichtet werden kann.

Die Skriptfunktion, dass bei einem Übergang die letzte Tonhöhe (*\$last_note*) abgespielt wird, bewirkt bei den tonalen Klappengeräuschen, dass der alte Ton "nachklingt", sodass die Noten im Übergang nochmals verbunden werden.

Der wichtigste Modulator dieser Gruppe ist der Breathcontrollerwert, da sonst bei leisen Tönen die tonalen Klappengeräusche überproportional laut wiedergegeben werden. Zudem wird mithilfe einer Lookup-Tabelle die Lautstärke einzelner Tonhöhen unterschiedlich stark verändert.

Multiphonics

Die Aufnahmen der Multiphonics wurden schon bei der Reproduktion des unsauberen Tonansatzes verarbeitet. Bei den Intervallübergängen bietet sich eine erneute Verwendung der Samples an. Insgesamt werden sechs Multiphonicstöne eingesetzt.

Durch eine starke Tonhöhenvariation per Zufallsmodulator (± 6 Halbtöne) wurde eine ausreichende Variation des "Störgeräusches" erzeugt. Der Breathcontrollerwert sorgt wie bei den tonalen Klappengeräuschen für eine angemessene Einbindung in die Gesamtdynamik. Allerdings arbeitet der Anschlagstärkenmodulator der Lautstärke invers, sodass die Aufnahmen nur bei leicht gedrückte Tasten hörbar werden.

Die Kombination dieser drei Geräuschanteile bewirkt also folgendes:

- bei leisen Dynamikstufen sind die tonalen Klappengeräusche & Multiphonics leiser als bei lauten Tönen
- mit der Anschlagstärke verändert sich der Klang der Tonübergänge: von weichen Übergängen, die durch die Multiphonicsverarbeitung teilweise unsauber klingen, bis zu rein perkussiven Klickgeräuschen bei starken Tastenanschlägen können je nach musikalischer Situation unterschiedliche Intervallübergänge erzeugt werden.

Übergang der Wavetables

Bei Tonübergängen auf der Klarinette wird die Länge der Luftsäule verändert. Daher reicht es aus, die Wavetables verschiedener Tonhöhen ineinander überzublenden (mit einer

Überblendzeit von ca. 25ms). In Kombination mit den komplex programmierten Nebengeräuschen entsteht so ein ausreichend überzeugender Eindruck der Intervallübergänge.

B 2.4.4 – Intonation

Im Vergleich zu synthetisch erzeugten Klängen muss bei realen Blasinstrumenten stets auf die korrekte Intonation geachtet werden, da bei unsauberer Spieltechnik Frequenzabweichungen vom Idealtone entstehen, die einen Ton "schief" klingen lassen. Selbst bei professionellen Musikern treten minimale Abweichungen von der idealen Tonhöhe auf, die entweder charakteristisch für das jeweilige Instrument sind, oder sogar als musikalisches Stilmittel bewusst eingesetzt werden. In dieser Arbeit sollen drei verschiedene Intonationsphänomene nachgebildet werden. Dabei muss auf eine identische Modulation der sechs Dynamikstufen geachtet werden, da ansonsten die Phasenlage korrumpiert würde.

1. Intonationsschwankung unterschiedlicher Dynamikstufen

Eine Tonhöhenanalyse der Crescendoaufnahmen für die Wavetablesynthese ergab eine relativ große Differenz der Tonhöhen zwischen den Dynamikstufen: Während die fortissimo-Lautstärke nahezu perfekt intoniert war (± 4 Cent), wich die Tonhöhe bei der pianissimo-Lautstärke kurz nach dem Tonansatz um bis zu +35 Cent vom Idealtone ab.

Um dieses Verhalten nachzubilden, wurde die Tonhöhe der Wavetables vom Breathcontrollerwert in ähnlicher Dimension moduliert.

2. Manuelle Kontrolle der Intonation

Für die Vibratokontrolle (siehe unten) wurde das Pitchwheel mit einer maximalen Modulationsintensität von ± 20 Cent ausgestattet. Jedoch kann diese Steuermöglichkeit auch dazu eingesetzt werden, die Intonation manuell zu steuern.

3. Intonationsschwankungen bei schnellen Tonwiederholungen

Mit zunehmender Spielgeschwindigkeit wird bei Blasinstrumenten die korrekte Intonation immer schwieriger, da Intonation nach dem Prinzip eines rückgekoppelten Systems funktioniert: die Tonhöhe wird ständig "überwacht", und im Falle einer Abweichung korrigiert. Da die Korrekturgeschwindigkeit von der Reaktionszeit des Musikers abhängt, kommt es bei kleineren Notenwerten zu einer höheren, da unkorrigierten Abweichung.

Für die mathematische Modellierung dieses Phänomens wurde eine abschnittsweise gebrochen-rationale Funktion $\hat{p}(t)$ verwendet (Abbildung 33), deren Gleichung (14) die maximale Abweichung der Tonhöhe in Abhängigkeit des Zeitintervalls Δt zwischen zwei Noten angibt.

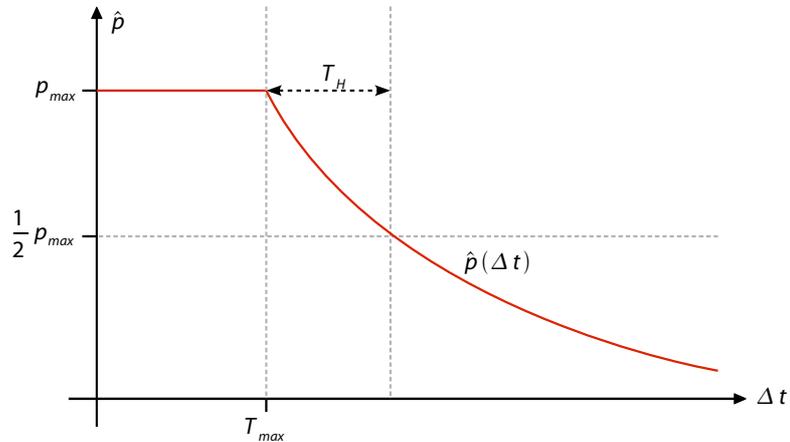


Abbildung 33: Funktionsgraph der Tonhöhenabweichungsfunktion

$$(14) \quad \hat{p}(\Delta t) = \begin{cases} p_{max} \cdot \frac{T_H}{t + (T_H - T_{max})} & , t > T_{max} \\ p_{max} & , t < T_{max} \end{cases}$$

$$(15) \quad p(\Delta t) = \text{rand}(\pm \hat{p}(\Delta t))$$

Der Effekt der Funktion wird mit drei Konstanten gesteuert: p_{max} ist die maximal mögliche Tonhöhenmodulation, die erreicht werden kann, wenn zwischen zwei Tönen das Zeitintervall T_{max} unterschritten wurde. T_H entspricht (ungefähr) der Halbwertszeit einer exponentiellen Funktion, und gibt an, mit welcher Intensität die Maximalabweichung bei größeren Zeitabständen als T_{max} abfällt.

Die tatsächliche Tonmodulation $p(\Delta t)$ soll dann ein Zufallsmodulator erzeugen (15), der als positiven und negativen Maximalwert den für jedes Zeitintervall Δt berechneten Wert $\pm \hat{p}(\Delta t)$ von (14) verwendet.

Die Funktion wird mithilfe des folgenden Skripts implementiert:

```

1  on init
2    declare const $T_max := 140
3    declare const $p_max := 15
4    declare const $T_H := 30
5    declare $last_note := 0
6    declare $speed := 0
7    declare $pitch_var := 0
8  end on
9  on note
10   if($speed # 0 and $last_note = $EVENT_NOTE)
11     $speed := $ENGINE_UPTIME - $speed
12     if($speed > $T_max)
13       $pitch_var := $p_max * $T_H / (($speed + $T_H) - $T_max)
14     else
15       $pitch_var := $p_max
16     end if
17     $pitch_var := random(-$pitch_var, $pitch_var)*1000
18     change_tune($EVENT_ID,$pitch_var,1)
19   end if
20   $speed := $ENGINE_UPTIME
21   $last_note := $EVENT_NOTE
22 end on

```

Codebeispiel 4: Tonhöhenabweichung bei schnellen Tonwiederholungen

In der Variable *\$speed* wird mithilfe der internen Variable *\$ENGINE_UPTIME* das Zeitintervall zwischen zwei Note-On Befehlen gespeichert (Zeile 11 & 20). Da der Effekt nur bei Tonwiederholungen auftreten soll, wird die letzte Tonhöhe unter *\$last_note* (Zeile 21) gespeichert. Bei zwei aufeinanderfolgenden Noten wird deren Zeitabstand berechnet, der als Basis für die Berechnung des Maximalwerts (Zeilen 12-16, vgl. Formel (14)) dient. Die Fallunterscheidung der abschnittsweise definierten Funktion wird mit einer If-Kontrollstruktur implementiert. Die tatsächliche Tonhöhenabweichung *\$pitch_var* (Zeile 17 vgl. Formel (15)), muss allerdings noch mit dem Faktor 1000 multipliziert werden, da der Befehl *change_tune* (Zeile 18) im Millicentbereich arbeitet.

Die Werte der drei Konstanten *\$T_max*, *\$T_H* und *\$p_max* werden willkürlich festgelegt, sodass der gewünschte Effekt hörbar, aber nicht störend wird. Die Werte von 140ms für *\$T_max*, 30ms für *\$T_H* und 15 Cent für *\$p_max* liefern das gewünschte Resultat.

Das vorliegende Modell bildet das Phänomen übrigens nicht vollständig nach: die mit (14) und (15) berechnete Tonhöhendifferenz wird statisch auf den kompletten Ton angewendet. Die Rückkopplungsfähigkeit (d.h. dass die falsche Intonation im Verlauf des Tons korrigiert wird) wird nicht implementiert. Die Umsetzung dieser Eigenschaft könnte erreicht werden, indem anstelle einer konstanten Tonhöhenmodulation die Intensität einer abfallenden Hüllkurve moduliert wird, sodass nach einer kurzen Zeit (der Reaktionszeit des Musikers) wieder die Ausgangstonhöhe erreicht würde. Da der Effekt jedoch hauptsächlich bei schnellen (und damit kurzen) Tönen auftritt, birgt die Vereinfachung des Modells keine hörbare Minderung des Klangeindrucks, sodass der

zusätzliche Aufwand nicht nötig erscheint.

B 2.4.5 – Spieltechniken

Auf der Klarinette lässt sich wie auch bei anderen Blasinstrumenten der Ton durch spezielle Spieltechniken verändern, was die musikalische Gestaltungsfreiheit erhöht. Die wichtigsten Spieltechniken der Klarinette sind Vibrato, Growl und Flatterzunge. Bei der Reproduktion der Klarinette beschränke ich mich auf die Nachbildung der Vibrato- und der Growlspieltechnik, da die Flatterzungentechnik fast ausschliesslich in der Neuen Musik Anwendung findet und das Verfahren der Reproduktion der Growlspieltechnik ähnelt.

Vibrato

Das Vibrato ist eine periodische Schwankung der Tonhöhe (und Lautstärke). Beim konventionellen Klarinettenspiel in der Klassikmusik wird auf Vibrato vollständig verzichtet, im Jazz- und Klezmerbereich ist es jedoch ein gebräuchliches Mittel zur Steigerung des musikalischen Ausdrucks.

Die Implementierung arbeitet auf zwei verschiedenen Ebenen: Mit dem Modulationsrad lässt sich ein LFO steuern, der die Tonhöhe (± 9 Cent), die Lautstärke (zu ca. 40%) und die Lautstärke der Luftgeräusche moduliert (ca. 50% um -90° phasenverschoben, um die Trägheit der Luftsäule zu simulieren). Die Frequenz des LFOs wurde nach Gehör eingestellt (ca. 5 Hertz), wobei der Breathcontrollerwert die Frequenz leicht erhöht, um eine Steigerung der Intensität bei hohen Dynamikstufen zu erzielen. Auf die Schwingung des LFOs wurde ein Fade-In gelegt (ca. 400ms) da das echte Vibrato ebenfalls erst nach ungefähr dieser Zeit einsetzt.

Diese recht simple Methode ist für eine grobe Nachbildung der Vibratospielechnik ausreichend. Dass es noch ungleich komplexer geht, belegen spezielle Patente¹¹, um Vibratosimulationen noch realistischer zu machen.

Mithilfe einer zusätzlichen Steuermöglichkeit kann der Charakter des Vibratos weiter an die musikalischen Vorgaben angepasst werden: wie schon kurz im Kapitel "Intonation" erwähnt, steuert das Pitchwheel eine Tonhöhenabweichung von ca. ± 20 Cent. Zusätzlich wurden sämtliche Modulatoren des Vibrato-LFOs (siehe oben) auch an das Pitchwheel weitergegeben. Durch schnelles Auf- und Abbewegen des Pitchwheels kann somit ein "manueller" LFO erzeugt werden, dessen maximale Frequenz aufgrund der menschlichen

11 "Determination of modal resonances and body impulse response of a musical instrument by analysis of sounds performed with pitch changes. Application to the synthesis of vibrato & portamento with samplers", eingereicht von Giorgio Tommasini und Stefan Lucato

motorischen Beschränkungen bei ca. 3 – 4 Hz liegt. Langsame Vibratomodulationen können mit dieser Methode besser nachgebildet werden, da die Intensität und die Geschwindigkeit größeren Schwankungen ausgeliefert sind (und damit ein natürlicheres Klangbild erzeugen), als beim mathematisch exakten LFO des Modulationsrades. Dafür setzt die Methode ein größeres musikalisches Verständnis sowie etwas motorisches Geschick voraus.

Growl

Die Growlspieltechnik wird v.a. In der Jazz- und Klezmermusik eingesetzt. Durch Summlaute in die Klarinette erhält der Ton einen schreienden Charakter. Diese "schmutzigen" Töne sind ein populäres Stilmittel, um den musikalischen Phrasen mehr Expressivität zu verleihen.

Für die virtuelle Nachbildung wurde ein relativ schnell oszillierender LFO mit ca. 40 Hz erzeugt, der die Grenzfrequenz eines Tiefpassfilters steuert (im Vergleich zur Lautstärkemodulation kann so das Dynamikverhalten grob nachgebildet werden, siehe "Tonansatz"). Die Intensität des LFOs wird mithilfe des Aftertouch gesteuert. Gleichzeitig moduliert selbiger auch die Tonhöhe mit maximal -10 Cent, sodass eine deutliche Klangänderung bei verstärktem Tastendruck auftritt.

B 2.4.6 – Modellierung der akustischen Umgebung

Bis zu diesem Zeitpunkt wurde die Klarinette als abgeschlossenes Klangerzeugersystem betrachtet. Durch die Interaktion des Musikers mit dem Instrument und der raumabhängigen Ausbreitung des Schalls entstehen jedoch weitere Faktoren, deren Nachbildung für eine überzeugende Reproduktion von entscheidender Bedeutung sind. Im letzten Teil dieses Kapitels soll daher beschrieben werden, wie die Interaktion der virtuell modellierten Klarinette mit ihrer Umgebung simuliert wird.

Reproduktion der Umgebungsgeräusche

Neben dem eigentlichen Ton entstehen weitere Geräusche, die das Klangbild der Musikinstrumente mitprägen. Bei Blasinstrumenten sind das unter anderem Atemgeräusche, Bewegungsgeräusche (Kleiderrascheln etc.) und Mundgeräusche (Schmatzer). Beim herkömmlichen Samplingverfahren werden diese Geräusche vermieden, da sie sonst als Bestandteil des Instrumentenklanges bei jeder identischen Note auftreten würden (z.B. würde bei jedem mittellauten c" ein Kleidungsrascheln zu vernehmen sein.) In musikalischen Pausen, wenn also kein Sample abgespielt wird, ist absolute Stille ("Digital Null"), solange keine künstliche Hallfahne erzeugt wird. Gerade in

Solopassagen entsteht daher ein steriler und künstlicher Klang.

Bei dem hier vorgestellten Verfahren wurden die Geräusche separat aufgezeichnet, und können nun dem Instrumentenklang hinzugemischt werden. Die oben beschriebene Kopplung wird dadurch aufgehoben. Da diese Geräusche in der Realität unregelmäßig auftreten, muss bei der Programmierung ebenfalls auf eine stochastische Wiedergabe der Geräusche geachtet werden. Dies wird durch eine Zufallsmodulation der Lautstärke (bzw. der Abspielposition) erreicht.

Bewegungsgeräusche

Für die Reproduktion der Bewegungsgeräusche wurde eine ca. eine Minute lange "Stille" aufgezeichnet, in der sich der Musiker leicht bewegen sollte. Neben den so auftretenden Bewegungsgeräuschen wurden gleichzeitig noch andere Hintergrundgeräusche aufgezeichnet (Lüfterrauschen der Preamps und Mikrofonrauschen). Die so entstandene Klangkulisse dient als Ausgangsmaterial für die Gestaltung der Umgebungsgeräusche, dessen spezielle Wiedergabekonditionen das folgende Skript implementiert:

```

1  on init
2    declare $ambience_note := 0
3    declare $hold_time := 2000000
4    declare $last_ID
5    declare $start_ID
6    declare $is_On := 0
7  end on
8  on note
9    $last_ID := $EVENT_ID
10   if($is_On = 0)
11     $start_ID := play_note($ambience_note, 100,0, 0)
12     $is_On := 1
13   end if
14 end on
15 on release
16   wait($hold_time)
17   if($last_ID = $EVENT_ID)
18     note_off($start_ID)
19     $is_On := 0
20   end if
21 end on

```

Codebeispiel Nr. 5: Wiedergabekonditionen der Umgebungsgeräusche

Sobald eine Taste gedrückt wird, speichert die Variable *\$last_ID* die ID des Ereignisses. Wenn das Umgebungsgeräusch noch nicht abgespielt wird (Zeile 10), wird eine künstliche Note mit der Nummer *\$start_ID* erzeugt, die die Wiedergabe des Umgebungsgeräuschs startet. Wenn eine Taste losgelassen wird, wartet das Skript eine vorgegebene Zeit von 2 Sekunden (Zeile 3 & 16). Dies stellt sicher, dass zwischen kurzen Pausen das Geräusch aktiv bleibt. Wenn in dieser Zeit keine neue Taste gespielt wird, d.h. Die letzte Ereignisnummer immer noch die selbe Nummer ist (Zeile 17), wird ein Note-Off Befehl erzeugt (Zeile 18), der das Umgebungsgeräusch beendet.

Das Geräusch wird mit einer Hüllkurve mit langer Ausklingzeit (ca. 1,5 Sekunden) moduliert, sodass nach einem Note-Off Befehl die Lautstärke langsam ausgeblendet wird. Zusätzlich bestimmt ein Zufallsmodulator die Startposition innerhalb des Samples, um die Variation zu erhöhen.

Schmatzgeräusche

Die Schmatzgeräusche sollen am Ende einer gespielten Phrase abgespielt werden, d.h. wenn das Mundstück abgesetzt wird, um die Lippen zu benetzen. Durch die Release-Trigger Methode (d.h. Der Note-Off Befehl startet die Note) kann das Verhalten nachgebildet werden. Damit das Geräusch nicht während Intervallübergängen auftritt, wird mithilfe eines Skriptes die Lautstärke entsprechend angepasst.

```
1  on init
2    declare $voice_count := 0
3  end on
4  on note
5    inc($voice_count)
6  end on
7  on release
8    dec($voice_count)
9    if($voice_count > 0)
10     set_engine_par($ENGINE_PAR_VOLUME, 0, 0, -1, -1)
11  else
12     set_engine_par($ENGINE_PAR_VOLUME, 250000, 0, -1, -1)
13  end if
14 end on
```

Codebeispiel 6: Deaktivierung der Schmatzgeräusche bei Intervallübergängen

Mithilfe der bekannten Variable `$voice_count` wird die Anzahl der aktiven Noten gespeichert. Bei einem Note-Off Befehl wird geprüft, ob ein Intervallübergang vorliegt (Zeile 9), d.h. ob noch eine andere Note gedrückt ist. Wenn dies der Fall ist, wird die Lautstärke auf 0 gesetzt (Zeile 10), andernfalls wird das Sample mit der normalen Lautstärke abgespielt (Zeile 12)

Da die Schmatzgeräusche nur bei ca. 30% aller Phrasenenden auftreten, wird mithilfe eines Zufallsmodulators die Lautstärke der Schmatzgeräusche abgesenkt.

Atemgeräusche

Die Atemgeräusche wurden aus Aufnahmen realer Phrasen extrahiert (ca. 50 Stück) und unterhalb des Tonumfangs der Klarinette auf verschiedene Velocityebenen verteilt (von leicht nach stark). Diese Geräusche können nun im Nachhinein in den MIDI-Sequencer eingetragen werden. Eine Implementierung in das Echtzeitspiel (z.B. die Berechnung der Phrasenlänge mittels Differenz zwischen erster gedrückter und letzter losgelassener Note und dementsprechender Lautstärkemodulation für die Wiedergabe in musikalischen Pausen) mithilfe der Skriptsprache wäre durchaus möglich, soll aber im Rahmen der Arbeit

nicht weiter verfolgt werden, da eine manuelle Eintragung der Geräuschnoten ein besseres Aufwand / Ertrag-Verhältnis bietet.

Nachbildung der Raumakustik

Die Wechselwirkung eines Klangerzeugers mit seinem umgebendem Raum kann (grob) in zwei Komplexe eingeteilt werden: die zeitliche Veränderung durch diffuse Reflexionen (Nachhall) und die Entstehung von nichtlinearen Frequenzverläufen aufgrund der Ausbildung von Raummoden.

Da für die Wavetablesynthese nur kurze Ausschnitte der Aufnahmen verwendet werden, spielt der Nachhall nahezu keine Rolle mehr: Zwar wird die eigentliche Wellenform von nachhallenden Tönen überlagert, allerdings ist deren Lautstärke deutlich unterhalb der eigentlichen Wellenform. Theoretisch bleibt noch die raumabhängige Frequenzlinie durch Raummoden, jedoch werden durch Normalisierung der einzelnen Wellenformen unterschiedliche Pegel fast vollständig nivelliert. Die Geräusche sind aufgrund ihres atonalen Klangcharakters nicht anfällig für Raummoden. Durch Nahmikrofonierung wurde ausserdem der Raumanteil weitestgehend eingeschränkt.

Der vorliegende Klarinettenklang ist demnach fast vollständig von Rauminformationen "gesäubert"¹². Um eine natürliche Nachbildung des Klangverhaltens zu ermöglichen, wird mithilfe des eingebauten Faltungshalls und einer speziell erstellten Impulsantwort des Aufnahmeraums die Raumakustik simuliert.

Abbildung 34 zeigt die Signalkette innerhalb des KONTAKT-Instruments.

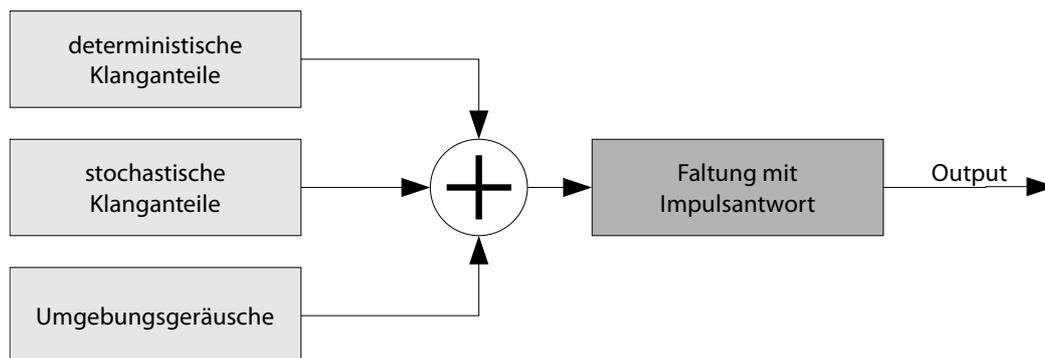


Abbildung 34: Signalkette der virtuellen Klarinette

Die verschiedenen Klanganteile (Wavetablesynthese, Nebengeräusche und Umgebungsgeräusche) werden zusammengemischt und mit der Impulsantwort des Aufnahmeraumes gefaltet. Die Differenz zwischen Dry- und Wet-Pegel beträgt dabei 6 dB, kann aber je nach gewünschtem Klang variiert werden.

¹² Ideal wäre allerdings schon die Aufnahme in einem schalltoten Raum, wie es z.B. bei den Samplemodeling Instrumenten geschieht.

TEIL C: Diskussion der Ergebnisse

C 1 – Analyse

Im letzten Kapitel der Arbeit soll die Klarinette analysiert und die Anwendbarkeit des Verfahrens auf weitere Instrumente untersucht werden. Die in den vorherigen Kapiteln verwendete Methode der FFT-Analyse ist dabei nur teilweise von Bedeutung, da die subjektiv empfundene Klangqualität / der Realismus nur schwer messbar sind. Daher soll neben einer Analyse der Frequenzspektren der Vergleich einiger nachgespielter Melodiephrasen vorgenommen werden.

C 1.1 – FFT-Analyse

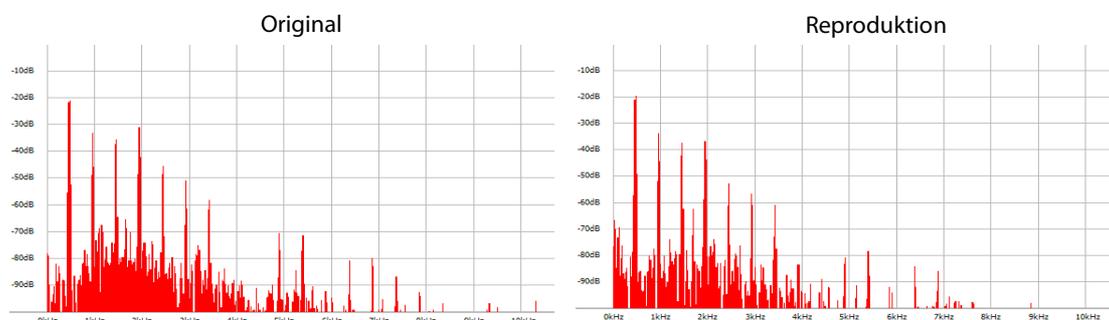


Abbildung 35: FFT-Vergleich zwischen Original und reproduziertem Klarinetten-ton

Abbildung 1 zeigt die Frequenzspektren eines real gespielten B3 in mittlerer Lautstärke (links) und der virtuellen Reproduktion (rechts). Auf den ersten Blick scheint der virtuelle Ton eine ausreichend genaue Übereinstimmung zu bieten. Die stochastischen Anteile scheinen etwas geringer auszufallen, was sich aber aufgrund der getrennten Erzeugung der Klanganteile nachträglich korrigieren liesse.

Für eine genauere Untersuchung wurden nun alle sechs Dynamikstufen analysiert und mit dem realen Vorbild verglichen. Dabei wurde der stochastische Anteil stummgeschaltet, um die Übersichtlichkeit zu erhöhen. Zusätzlich wird die Pegeldifferenz der einzelnen Teiltöne berechnet und in einem Diagramm dargestellt. Um die Aussagekraft dieses Schaubildes zu erhöhen wurde die Differenz der Grundtonpegel von den restlichen Obertonpegeln subtrahiert. Dadurch werden die Grundtonpegel nivelliert und fungieren als Eichgröße. Eine arithmetische Mittelwertbildung dieser (vorzeichenunbewerteten) Pegeldifferenzen liefert den durchschnittlichen Abweichungswert der Obertöne vom Originalspektrum, der als Richtgröße für die Genauigkeit der Reproduktion betrachtet werden kann. Um eine Vorstellung von den Größendimensionen dieser Richtgröße zu erhalten, wurden das pp-Sample und das ff-Sample nach den selben Vorgaben verglichen. Der Vergleich (und damit die ungefähre "Gesamtspektraldynamik") ergab eine Differenz von 29dB.

Die Spektraldiagramme der Dynamikstufen finden sich in Abb. 36 (blau= virtuell; rot= real).



Abbildung 36: FFT-Vergleich der einzelnen Dynamikstufen

Die orangene Kurve gibt die jeweils normalisierten und unbewerteten Pegeldifferenzen an. Die dB-Größe neben der Dynamikbezeichnung ist der wie beschrieben berechnete Abweichungswert. Die Werte zwischen 1,9dB und fast 6dB können als ausreichend genau interpretiert werden: die Auswertung der AET-Spektren ergab dagegen 9,83dB für das mf-Sample mit aufgeprägtem pp-Spektrum und 9,88dB für das gleiche Sample mit dem ff-Sample als Spektralvorlage.

C 1.2 – Subjektive Beurteilung des Klangs

Um den Klang zu beurteilen, wurden einige Phrasen, die bei den Klarinettenaufnahmen als Referenzaufnahmen erstellt wurden, mit der virtuellen Reproduktion und einem Breathcontroller nachgespielt [HB Nr. 10 - 12]. Dabei sollen nicht nur die Unterschiede in der Obertonstruktur, sondern auch die unterschiedlichen Klangphasen beim Tonansatz und zwischen Intervallen untersucht werden.

Durch die Nebengeräusche erhält der Klang der virtuellen Klarinettenphrasen einen lebendigen Charakter, den man bei üblichen virtuellen Instrumenten oftmals vermisst. Die Unterschiede zwischen verschiedenen Dynamikstufen sind deutlich wahrnehmbar, sodass die Reproduktion der spektralen Dynamik als einigermaßen gelungen betrachtet werden kann.

Beim Vergleich der Aufnahmen kann man jedoch feststellen, dass der Gesamtklang der virtuellen Klarinette etwas dumpfer ist (vor allem bei lauten Passagen, vgl. die Ergebnisse aus C.1.1). Das hängt teilweise mit der Impulsantwort des Raumes zusammen, die hohe Frequenzen etwas abdämpft, kann aber auch daran liegen, dass bei der Superposition der Perioden für die Wavetableerstellung Phasenverzerrungen auftreten, die die hohen Obertöne dämpfen. Da die Klarinette ein verhältnismäßig obertonarmes Spektrum aufweist (im Vergleich etwa zu einer Trompete), ist dieser Effekt nicht allzu störend.

Ein weiterer Nachteil der virtuellen Klarinette ist der statische Klangcharakter: während der Ton bei den Realaufnahmen auf derselben Dynamikstufe zwischen verschiedenen Klangfarben wechseln kann, ist man bei der Klarinette auf den Klang der aufgenommenen Wavetables beschränkt. Für die musikalischen Gestaltungsmöglichkeiten stellt dies einen bedeutenden Nachteil dar.

Der Tonansatz und die Intervalle sind einigermaßen überzeugend, im Vergleich zu den echten Klarinettenphrasen kann man jedoch ohne weiteres erkennen, dass es sich um künstlich modellierte Klänge handelt. Mit etwas mehr Erfahrung (da die Klarinette das erste in dieser Weise hergestellte Instrument ist) lässt sich das Ergebnis für zukünftige Entwicklungen sicherlich noch steigern.

Der Raumanteil der virtuellen Klarinette ist im Vergleich zu den Aufnahmen deutlich höher, kann jedoch nach Belieben angepasst werden. Allerdings klingt die Klarinette bei geringerem Raumanteil etwas steril.

C 2 – Anwendbarkeit auf weitere Instrumententypen

Das in dieser Arbeit beschriebene Verfahren zur virtuellen Reproduktion realer Musikinstrumente kommt eher für Instrumente in Frage, bei denen die Dynamik und die Tongestaltung auch in der quasistationären Phase kontrolliert werden kann. Bei Tasten- und Schlaginstrumenten, die nur zu Beginn des Tons angeregt werden, und nach dem Prinzip einer gedämpften Schwingung ausklingen, ist die herkömmliche Samplingmethode besser geeignet, da die unterschiedlichen Dynamikstufen nicht überblendet werden müssen. Blas- und Streichinstrumente jedoch profitieren von der neuen Möglichkeit, ohne Klangartefakte zwischen Dynamikstufen wechseln zu können.

Da das Verfahren auf der getrennten Reproduktion deterministischer und stochastischer Klanganteile beruht, müssen die (realen) Instrumente die stochastischen Anteile erzeugen können, ohne in Schwingung versetzt zu werden, damit diese Klanganteile separat aufgenommen werden können. Im Falle einer Geige beispielsweise könnte das Bogengeräusch durch eine Abdämpfung der Saiten isoliert werden. Bei der Reproduktion einer Blockflöte könnten allerdings Schwierigkeiten auftreten, da es nicht möglich ist, in die Flöte zu blasen, ohne einen Ton zu erzeugen.

Da der Tonansatz und die Intervallübergänge manuell nachgebildet werden, muss die Vorgehensweise für jedes Instrument individuell erarbeitet werden. Dies könnte bei Instrumenten mit extrem charakteristischen Einschwingphasen (beispielsweise Fagott) zu Problemen führen, die eine Weiterentwicklung des Verfahrens erfordern.

Der dumpfe Klangcharakter der Klarinette könnte sich bei obertonlastigeren Instrumenten als problematisch erweisen, falls der Grund für die Dämpfung in der Verfahrensweise selbst liegt.

C 3 – Fazit und Ausblick

Die in dieser Arbeit entwickelte Klarinette kombiniert eine für virtuelle Instrumente beachtlich hohe Spielbarkeit mit einer Klangqualität, die mit der im Moment vorherrschenden Samplingmethode durchaus mithalten kann. Da für die Reproduktion eines Instruments nur noch eine Crescendoaufnahme pro Halbton und einige Nebengeräusche benötigt werden, kann im Vergleich zu normalen Sampleaufnahmen, bei denen jeder Halbton in mehreren Artikulationen und Dynamikstufen einzeln aufgenommen werden muss, der Aufwand während der Instrumentalaufnahmen enorm gesenkt werden. Ein angenehmer Nebeneffekt dieses Verfahrens ist der geringe Ressourcenverbrauch: Die komplette Klarinette belegt 40MB des Speichers, und liegt damit weit unterhalb des Speicherverbrauchs eines herkömmlichen Instruments (teilweise mehrere hundert Megabyte).

Durch die getrennte Modellierung der unterschiedlichen Klanganteile kann die Klarinette nachträglich an die musikalische Situation angepasst werden: als Ensembleinstrument würde man die Nebengeräusche sicherlich absenken, um das Klangbild homogen zu halten, während bei Solopassagen diese Geräusche etwas prominenter eingesetzt werden können, um den Realismus zu erhöhen.

Das größte Problem des Verfahrens ist die Modellierung der einzelnen Klangabschnitte wie Tonansatz und Intervallübergänge, die hauptsächlich von der Erfahrung des Programmierers abhängt. Ausserdem ist das Verfahren überdurchschnittlich stark von der Verwendung geeigneter Eingabegeräte abhängig: Während die Steuerung mithilfe des Breathcontrollers eine intuitive Dynamikkontrolle zulässt, ist eine natürlich klingendes Ergebnis mit einem Modulationsrad nur nach langwieriger und umständlicher Nachbearbeitung der Controllerdaten möglich. Für die Reproduktion von Streichinstrumenten nach diesem Verfahren sollte daher nach einer Entwicklung eines entsprechenden Controllergerätes gesucht werden.

Ein weiterer Nachteil des Verfahrens (aber auch der herkömmlichen Samplingtechnologie) ist die mangelnde Variabilität des Klangcharakters. Dies resultiert in einem stets gleichen Klang des virtuellen Instruments. Eine Anpassung des Klangcharakters in den musikalischen Kontext ist somit nur bedingt möglich (durch die künstlich modellierten Spielweisen Vibrato und Growl)

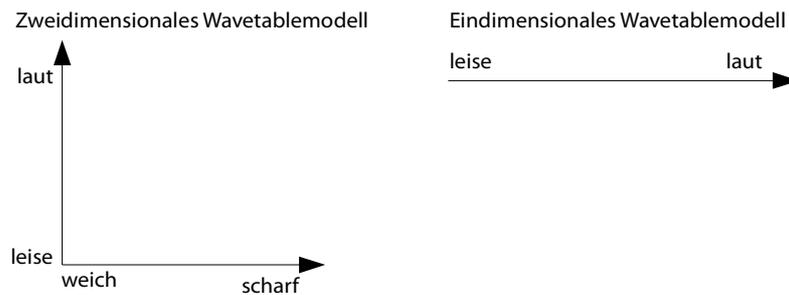


Abbildung 37: Entwurf eines mehrdimensionalen Wavetabekonzepts

Abbildung 37 skizziert eine Möglichkeit, die Komplexität der Reproduktion zu erhöhen. Das In dieser Arbeit entwickelte Modell entspricht einem eindimensionalen Raum, bei dem die Steuerung nur auf laut / leise beschränkt ist (wenn man die künstlichen Vibrato- und Growlspieltechniken ignoriert). Durch unterschiedliche Spielweisen entstehen jedoch unterschiedliche Klangcharakter, wodurch die realen Instrumentenaufnahmen an musikalischer Tiefe gewinnen. Streichinstrumente können beispielsweise durch Variation der Streichposition den Klang verändern: je näher die Seite am Steg gestrichen wird, desto "härter" wird der Ton. Man könnte dies durch eine "mehrdimensionale Wavetablestruktur" nachbilden, indem mehrere Crescendos des selben Tons (mit unterschiedlichen Streichpositionen) aufgenommen werden, und bei der Programmierung so verarbeitet werden, dass neben der Dynamiküberblendung eine weitere Steuergröße zwischen den verschiedenen Klangfarben überblenden kann. Der Herstellungsaufwand wächst natürlich mit der Anzahl an zusätzlichen "Klangdimensionen". Durch eine Automatisierung der Wavetableerstellung, die zur Zeit noch in Handarbeit geschieht, könnte dieser Aufwand jedoch gesenkt werden.

C 4 – Ethische Bewertung der technologischen Entwicklung

Die Arbeit soll mit einer ethischen Betrachtung des Forschungsgebiets abschliessen. Der Fortschritt auf dem Bereich der virtuellen Instrumente stellt auf den ersten Blick eine zunehmende Bedrohung der Lebensgrundlage vieler Instrumentalmusiker dar. Im Jahre 2003 erschienen Meldungen, dass die Broadwaymusiker "vom Computer ersetzt" werden sollen, was zu schweren Protesten der US-Musikergewerkschaft geführt hatte [Watkin, 2004]. Diese Befürchtungen gehen jedoch zurück bis in die Anfänge der Samplingtechnologie. Das Mellotron wurde von der englischen Musikergewerkschaft heftigst bekämpft: Die Musikgruppe "Tangerine Dream" musste sogar 2000 Pfund Schadensersatz an drei Kammerorchester zahlen, da der Einsatz des Instruments diese arbeitslos gemacht habe [Schelle, 2000].

Es gibt jedoch auch andere Standpunkte. Der Komponist und Dirigent Paul Henry Smith, der mit den Möglichkeiten der virtuellen Orchestern für die konzertante Aufführungspraxis experimentiert, spricht von einer technischen Revolution vergleichbar mit der Entwicklung des Buchdrucks von Johannes Gutenberg, der die Renaissance einläutete und damit einen der wichtigsten kulturellen Fortschritte der Menschheit ermöglichte [vgl. Lischke, 2007]. Sicherlich ist dieser Vergleich etwas überspitzt, dennoch sind einige Parallelen erkennbar: Während durch die Entwicklung des Buchdrucks die Produktion demokratisiert wurde, da das Anrecht auf die Verbreitung von Schriften nicht mehr auf die kirchliche Institution beschränkt war, bieten virtuelle Instrumente erstmals die Möglichkeit, ohne langjähriges Studium und Kontakte zu Sinfonieorchestern diesen Klangkörper für eigene Musikkompositionen zu verwenden. Eine ähnliche Entwicklung hat sich in den letzten Jahren im Bereich der Tontechnik abgespielt: durch die Preisentwicklung im Audioequipmentbereich ist ein eigener Markt für Homerecordingequipment entstanden, sodass nach einer Investition von wenigen hundert Euro eine Klangqualität erreichbar ist, die vor zwanzig Jahren nur großen Tonstudios vorbehalten waren. Das fast schon legendäre "Tonstudioterben" der letzten Jahre ist eine Seite der Entwicklung, andererseits entwickelte sich dadurch die Musikproduktion fast schon zum Volkssport, was man unabhängig von der Qualität der Erzeugnisse als durchaus zu bejahenden Schritt des musikinteressierten Menschen vom passiven Konsumenten zum Produzenten eigener Musik betrachten kann. Es gibt unzählige weitere Beispiele¹³, die von einer Auflösung hierarchischer Strukturen durch die Demokratisierung der Produktionsmittel zeugen.

Ein Nachteil dieser Entwicklung ist jedoch eine Nivellierung der Qualität. Aldous Huxley beschreibt schon 1933 anhand der Entwicklung im Buchmarkt, dass unter der Prämisse, dass der Anteil an talentierten Autoren in der Bevölkerung konstant bleibt, eine Erhöhung der Produktion eine Verminderung der literarischen Gesamtqualität zur Folge hat [zitiert nach Benjamin, 1937]. Diese elitäre Sichtweise ist sicherlich etwas eindimensional (und wird auch von Walter Benjamin als "wenig fortschrittlich" empfunden), allerdings kann man diesen Effekt in sämtlichen aufgeführten Beispielen erkennen: die Qualität der in den einschlägigen Internetforen hochgeladenen Homerecordingproduktionen lässt meist sehr zu wünschen übrig. Der recht überschaubaren Anzahl ambitionierter Bands auf Myspace steht eine fast unzählbare Masse von mittelmäßigen MP3-Uploads gegenüber. Im Bereich der virtuellen Instrumente manifestiert sich der Effekt in dem typischen Samplelibrary-

13 z.B. die Möglichkeiten der Selbstvermarktung in Internetportalen wie Myspace.com, die für die Krise der Major-Plattenfirmen teilweise verantwortlich sind

Sound vieler aktueller Filmmusikproduktionen (v.a. bei Fernsehproduktionen), die der Musik ihre klangliche Tiefe nimmt und auf billige Effekthascherei beschränkt.

Es lässt sich also sagen, dass die Demokratisierung der Produktionsmittel nicht unbedingt eine Erhöhung der künstlerischen Qualität der gesamten kulturellen Erzeugnisse bewirkt, dafür jedoch dem Einzelnen neue Möglichkeiten bietet, und ihn von der Konsumentenrolle in eine kreative Produzierendenrolle versetzt. Möglicherweise könnte diese neue Mündigkeit langfristig für einen Anstieg der Qualität sorgen, da der Durchschnittskonsument durch seine eigenen kreativen Erfahrungen die Leistung anderer Kulturschöpfenden besser beurteilen kann.

Dass die anfangs beschriebene Sorge der Musikergewerkschaft, vom Computer vollständig ersetzt zu werden¹⁴ unbegründet ist, zeigen folgende Überlegungen: die enorme Verbesserung der Authentizität der virtuellen Instrumente seit der Erfindung des Mellotrons findet nur auf gradueller Ebene statt. Der fundamentale Unterschied zwischen einem realen Musikinstrument und der digitalen Reproduktion kann als das beschrieben werden, was Vilém Flusser ein "modernes Paradox" nennt, das durch die Lücken zwischen den Zahlen der "denkenden Sache (res cogitans)" entsteht, die die "ausgedehnte Sache (res extensa)" abbilden möchte [Flusser 1991]. Diese "Lücken" in der virtuellen Realität, die durch die Rasterung der Steuerparameter entstehen, sind der Grund, weshalb die Reproduktion eines Musikinstruments nie an die Qualität seines Vorbildes herankommen kann.

Betrachtet man ausserdem die Wirkung virtueller Instrumente auf den Durchschnittskonsumenten, stellt man fest, dass die Hörgewohnheiten und Ansprüche an den Realismus der virtuellen Kopien parallel zu der technologischen Entwicklung steigen¹⁵. Während die Klänge des Mellotron heutzutage von praktisch jedem Zuhörer als künstlich erkannt werden und bestenfalls nostalgischen Charakter besitzen, zeigt die verängstigte Reaktion der damaligen Musikergewerkschaft, dass das Instrument als ernsthafte Konkurrenz betrachtet wurde. Die ersten elektronischen Samplingeräte aus den 1980er Jahren haben für die aktuelle Musikproduktion praktisch keine Relevanz mehr. Man kann

14 wie es beispielsweise in der Datenverarbeitung geschehen ist: Da die Organisation und Verwaltung von Daten mit elektronischen Mitteln kein fundamentalen Unterschied zur alten Datenverarbeitung mit Aktenschranken darstellt, kann der gleiche Arbeitsaufwand (Sortieren und Pflegen des Datenbestandes) von weniger Personal durchgeführt werden.

15 Ein weiteres Beispiel für die Gewöhnung an neue technologische Möglichkeiten lässt sich in den Anfängen der Cinematographie finden. Die ersten bewegten Bilder, die eine auf die Kamera zufahrende Lokomotive zeigt, sorgten damals für panische Reaktionen der Kinobesucher, die mit der ungewohnten Virtualität der visuellen Eindrücke nicht umgehen konnten. [Kroll 2008, S. 269].

davon ausgehen, dass aufgrund der aktuellen Entwicklungen den derzeit verwendeten Samplelibraries in wenigen Jahren ein ähnliches Schicksal widerfahren wird. Die objektiv lineare Entwicklung der klanglichen Authentizität führt daher zu einer zyklischen Entwicklung der subjektiven Empfindung: nachdem eine neue Technologie entwickelt wurde, die die alten Verfahren als überholt erscheinen lässt, steigt die Empfindlichkeit des Konsumenten für die Künstlichkeit der Klänge, sodass die Nachfrage nach einer weiteren technologischen Innovation entsteht. Hätte man als Hersteller virtueller Instrumente den Anspruch, reale Instrumente ein für allemal zu ersetzen, so wäre man einer Aufgabe konfrontiert, welche an die aussichtslose Tätigkeit von Sisyphos aus der griechischen Mythologie erinnert.

Man sollte daher virtuelle Instrumente eher als Kompositionswerkzeug denn als Ersatz für Instrumentalmusiker auffassen. Speziell mit dem in dieser Arbeit entwickelten Ansatz mit der Fokussierung auf die Echtzeitspielbarkeit soll dem Komponisten die Möglichkeit in die Hand gegeben werden, ohne jahrelange Übung Instrumente in seinen Kompositionen einsetzen zu können, ohne auf externe Musiker angewiesen zu sein, oder durch eine umständliche Nachbearbeitung und Programmierung der verschiedenen Sampleartikulationen im Kompositionsprozess gestört zu werden. Möglicherweise entsteht durch die Vereinfachung der Kompositionsarbeitweise eine Vergrößerung der Gesamtproduktion und damit auch ein neuer Bedarf an realen Musikaufnahmen, wodurch die Instrumentalisten letztendlich profitieren könnten. Denn eines bleibt klar: aus Sicht des Komponisten kann es nur von Vorteil sein, wenn seine Komposition von anderen Musikern eingespielt werden, da jede Interpretation unterschiedliche Aspekte der Komposition betont, und somit eine kritische Auseinandersetzung mit dem eigenen Werk ermöglicht. Sofern die Voraussetzungen vorhanden sind (d.h. Budget, Zeit und Kontakte), würde sich daher jeder Komponist für den realen Musiker entscheiden, unabhängig von der (immer weiter steigenden) Authentizität der virtuellen Reproduktion.

Quellenangaben

Literaturverzeichnis

- Benjamin, Walter*: Das Kunstwerk im Zeitalter seiner Reproduzierbarkeit, Frankfurt 1936.
- Bristow-Johnson, Robert*: Wavetable Synthesis 101, A Fundamental Perspective, Burlington 1996.
- Dickreiter, Michael*: Handbuch der Tonstudioteknik Band 1, München 1987.
- Flusser, Vilém*: Digitaler Schein, in: Rötzer, Florian (Hg.): Digitaler Schein. Ästhetik der elektronischen Medien. Frankfurt a.M. 1991.
- Hagenow, Henry*: Digitale Synthese komplexer Wellenformen zur Simulation akustischer, elektrischer und optischer Eigenzustände mehrdimensionaler Systeme, Berlin 2001.
- Kroll, Thomas*: Der Himmel über Berlin. Säkulare Mystagogie?, Berlin 2008.
- Marinic, Nicki*: Native Instruments KSP Reference Manual, Berlin 2010.
- Morgenstern, Jan*: Native Instruments KONTAKT 4 Referenzhandbuch, Berlin 2009.
- Runow, Bernfried*: Automatischer Stereo-Mixdown von 5.1-Mehrkanalproduktionen, Stuttgart 2008.
- Serra, Xavier*: Musical Sound Modeling with Sinusoids plus Noise, Barcelona 1997.
- Stephan Schelle*: Die Tangerine Dream Story, aus German Rock e.V. Heft 12/2000, Hasbergen 2000.
- Tolonen, Tero, Välimäki, Vesa, Karjalainen, Matti*: Evaluation of Modern Sound Synthesis Methods, Helisinki 1998.

Internetquellen

- Liberg, Nils*: Nils Liberg's Scripting Tools - <http://nilsliberg.se/ksp/>
- Lischke, Konrad*: Digitales Orchester – der Computer spielt die erste Geige, aus Spiegel Online – <http://www.spiegel.de/netzwelt/web/0,1518,482082,00.html> [2007]
- NI Kontakt 4 Produktseite*: <http://www.native-instruments.com/#/en/products/producer/kontakt-4/?page=969>
- Sample Modeling* - <http://www.samplemodeling.com/en/technology.php>
- Sengpiel, E.*: Berechnen der Summen-Pegel von zwei kohärenten Quellen - <http://www.sengpielaudio.com/Rechner-kohquellen.htm> [2010]
- Smith, J.O.*: Digital Audio Resampling Home Page - <http://www-ccrma.stanford.edu/~jos/resample/>
- Vienna Symphonic Library* - <http://vsl.co.at/en/211/442/478/1701/1702/1305.htm>
- Vintage Synth Explorer*: AKAI S1000 - <http://www.vintagesynth.com/akai/s1000.php>
- Wallander Instruments* - <http://wallanderinstruments.com/>
- Wakin, Daniel J.*: Concert Canceled Over Use of Electronic Musicmaker, aus New York Times – http://www.nytimes.com/2004/12/15/nyregion/15opera.html?_r=2&oref=login [2004]
- Wikipedia*, Artikel "Klarinette" – aufgerufen am 05.07.2010
<http://de.wikipedia.org/w/index.php?title=Klarinette&oldid=76240512>
- Wikipedia*, Artikel "MIDI" – aufgerufen am 01.07.2010
<http://de.wikipedia.org/w/index.php?title=Musical Instrument Digital Interface&oldid=76395276>

Anhang A1 – Hörbeispiele

Um die Ergebnisse nachvollziehbar zu machen, und einen Eindruck vom Klangverhalten der virtuellen Klarinette zu erhalten, wurden einige Hörbeispiele erstellt, die entweder auf der mitgelieferten CD gefunden werden können, oder unter

<http://christoph-hart.com/bachelor/audiodemos.zip>

heruntergeladen werden können. Die einzelnen Hörbeispiele sollen hier erläutert werden.

AET - Kapitel

01 – Original_Sinus – AET_Sägezahn_Sinus.wav

Man hört einen kurzen Ausschnitt einer Sinusschwingung, und das gefilterte Sägezahnspektrum.

02 – Original_Dreieck – AET_Sägezahn_Dreieck.wav

Hier hört man die Dreiecksschwingung und danach den damit gefilterten Sägezahnton.

03 – Original_Sägezahn – AET_Dreieck_Sägezahn.wav

In diesem Hörbeispiel kann man die Originalsägezahnschwingung und die Nachbildung aus der Dreiecksschwingung hören.

04 – Original_Klarinette_pp – AET_Klarinette_pp.wav

Das Hörbeispiel enthält eine Aufnahme eines pianissimo-Tons und eine fortissimo-Note, die mit dem pianissimo Sample gefiltert wurde

05 – Original_Klarinette_ff – AET_Klarinette_ff.wav

Hier hört man ein realen fortissimo Ton und ein mezzoforte-Sample mit dem aufgeprägten Spektralverlauf des fortissimo-Tons.

Wavetable - Kapitel

06 – Wavetable – Original Crescendo.wav

Ein real gespieltes Crescendo. Diese Audiodatei dient als Referenz für die kommenden Hörbeispiele

07 – Wavetable – ohne Dynamikcrossfade.wav

Ein virtuell erzeugtes Crescendo, allerdings ohne Überblendung zwischen verschiedenen Wavetables (sondern lediglich eine Pegelmodulation der lautesten Dynamikstufe).

08 – Wavetable – ohne Phaselocking.wav

Bei diesem Hörbeispiel wurden die Überblendungen aktiviert, allerdings sind die verschiedenen Wavetables nicht gleich lang (sondern manuell um einige Cent verschoben). Dadurch entstehen die Kammfiltereffekte, die bei normaler Überblendung zwischen zwei Samples auch auftreten würden.

09 – Wavetable – mit Dynamikcrossfade und Phaselocking.wav

Hier hört man das fertige Crescendo mit phasenstabiler Wavetableüberblendung. Die Artefakte aus Hörbeispiel 08 sind vollständig verschwunden, und die spektralen Veränderungen bei zunehmender Lautstärke ähneln durchaus Hörbeispiel 06.

Analyse - Kapitel

10 – Analyse – Jazz 1.wav, 11 – Analyse – Klassik 1.wav, 12 – Analyse – Klassik 2.wav

Die kurzen musikalischen Phrasen, die zur Analyse der virtuellen Reproduktion verwendet wurden. Man hört jeweils erst die real eingespielte und danach die virtuelle Phrase.

Bonus Tracks

Diese kurzen Musikstücke sollen die musikalische Verwendbarkeit der virtuellen Klarinette demonstrieren. Dabei wurden mehrere Musikstile eingearbeitet, um der hohen musikalischen Vielfalt des Instruments Rechnung zu tragen. Die Klarinettenspuren wurden mit dem Breathcontroller eingespielt und nur minimal nachbearbeitet.

13 – Pink Panther Theme.wav

Das weltbekannte Thema von Henry Mancini von einem virtuellen Jazz-Quartett interpretiert (Drums: Addictive Drums Jazz Kit, Piano: Galaxy Steinway, Kontrabass: AcousticSamples KontrBass)

14 – Alfred's Lied.wav

Der Titelsong des Films "Seniorenland" (Regie: Friederike Fiek). Ursprünglich ist das Stück als Duett für Cello und Klavier konzipiert, allerdings erlaubt der Charakter der Melodie auch eine Interpretation mit der virtuellen Klarinette.

15 – Fantasie für Klarinette und Klavier.wav

Ein Klarinettenstück mit Klavierbegleitung im spätromantischen Stil.

16 – Klezmer Pattern.wav

Ein kurzes Musikstück im Klezmer-Stil. Die akustische Gitarre und die Percussions wurden live eingespielt.

17 – Balkan Beatz.wav

Die Klarinette in einem unkonventionellen Umfeld in einer Mischung aus osteuropäischer Polka und Hip Hop. Das Akkordeon, die Gitarre und die Percussions wurden live eingespielt, das Schlagzeug wurde mit Addictive Drums erzeugt.

Anhang A2 – Modulatorübersicht

Gruppen:

- Wavetables: pp – ff
- Intervall Clicks Round Robin 1 – 3
- tonale Intervall Clicks
- Intervall Multiphonics
- Luftgeräusche
- Nebenluft
- Luftimpulse Round Robin 1 – 6
- Multiphonic Attacks
- Schmatzgeräusche
- Bewegungsgeräusche
- Atemgeräusche

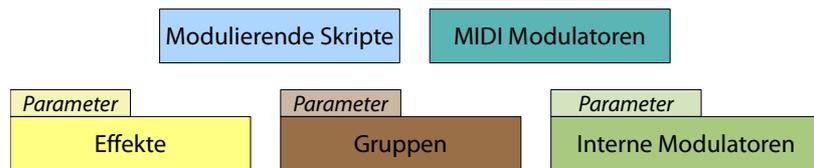


Abbildung 38: Legende

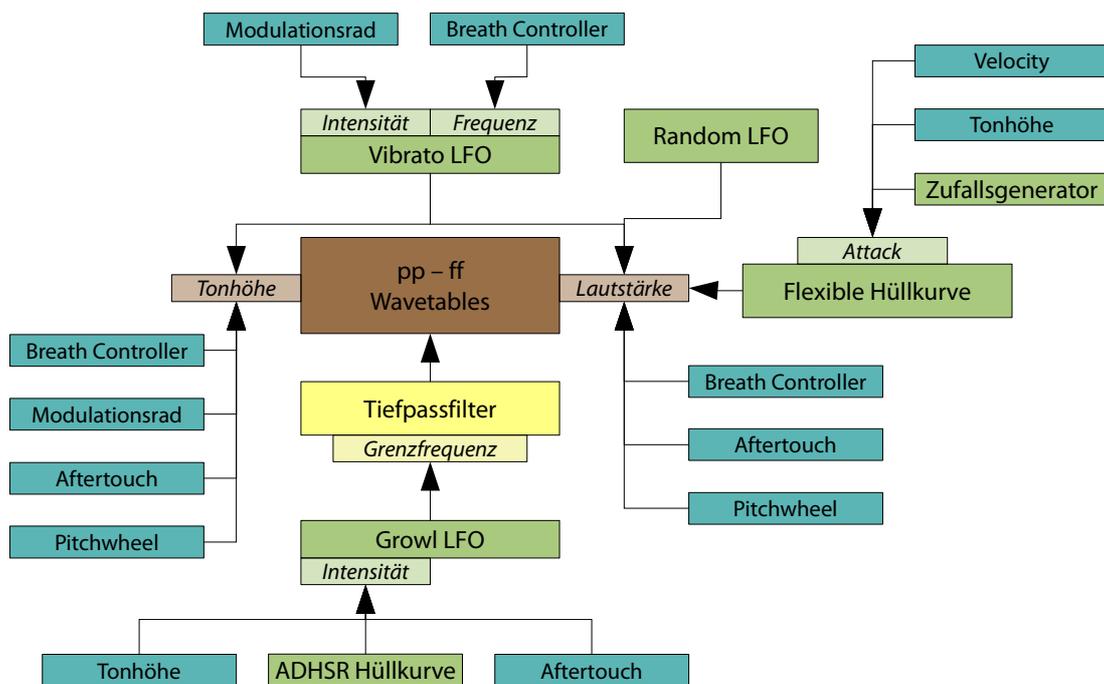


Abbildung 39: Wavetable-Modulatoren

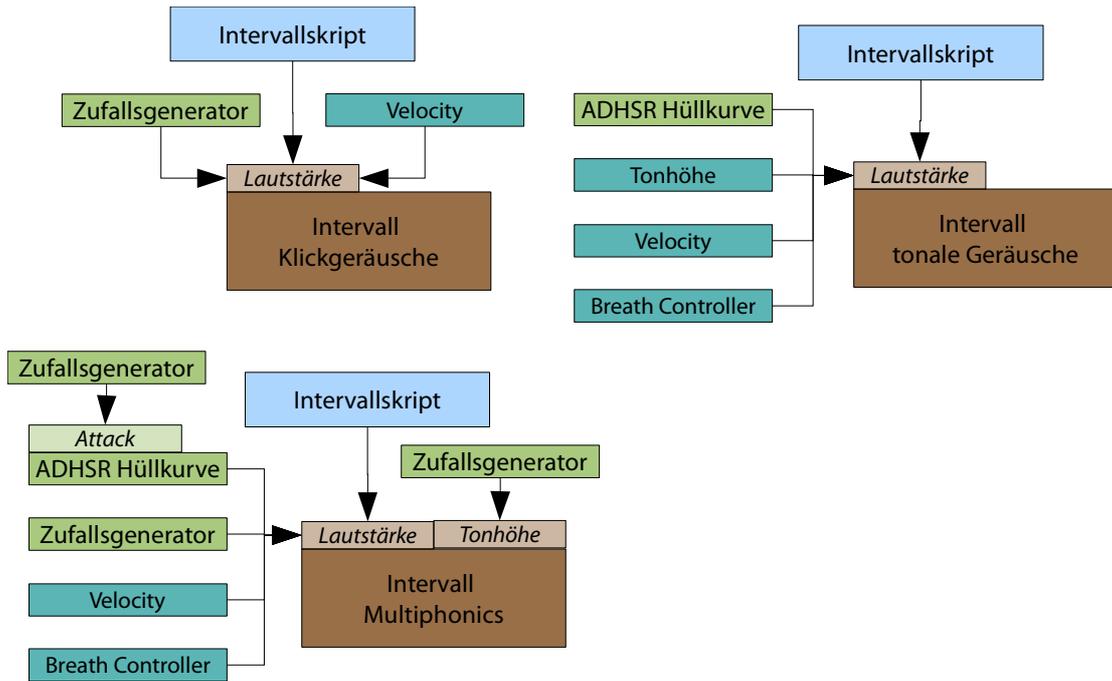


Abbildung 40: Intervall-Modulatoren

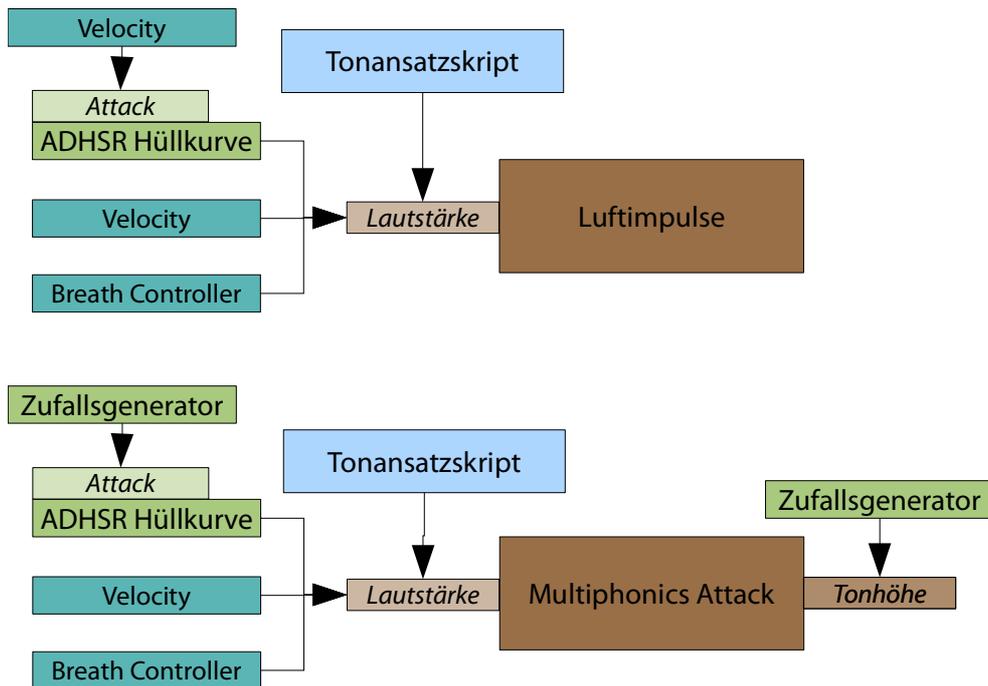


Abbildung 41: Tonansatzmodulatoren

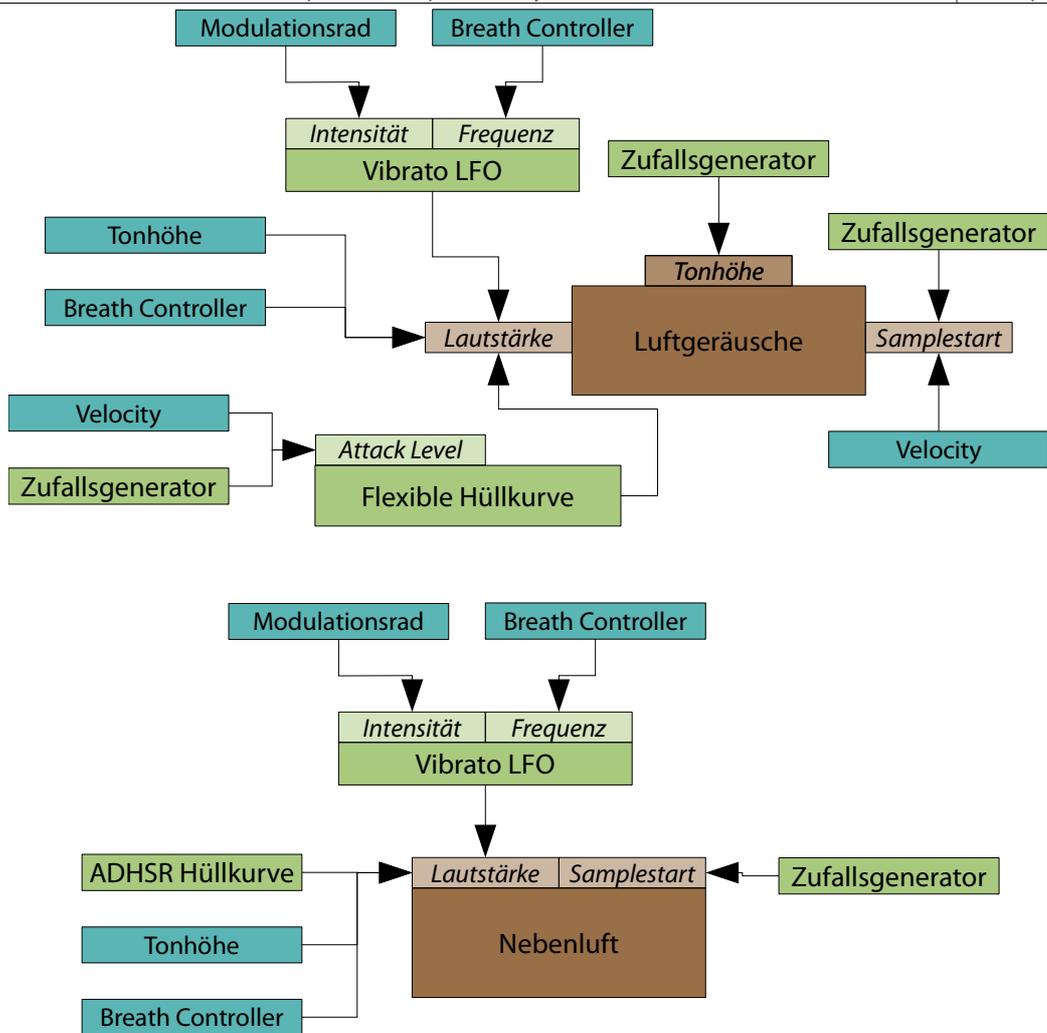


Abbildung 42: Luftgeräuschmodulatoren

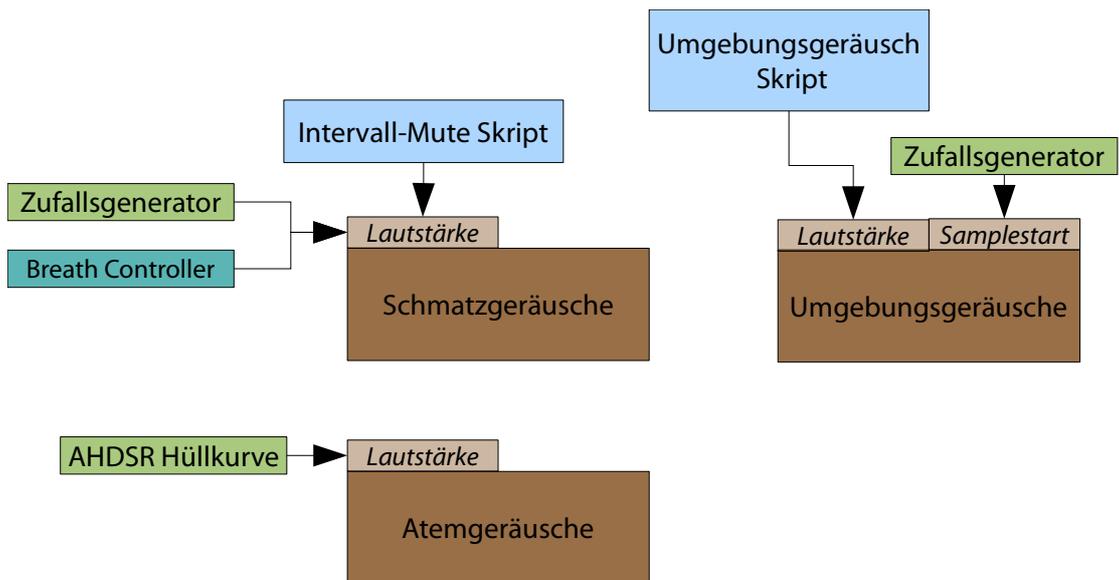


Abbildung 43: Nebengeräusch-Modulatoren

Abbildungsverzeichnis

Abbildung 1: (vereinfachter) Crossfade zwischen Dynamikstufen.....	10
Abbildung 2: Phasenauslöschung bei Überlagerung ähnlicher Frequenzen.....	11
Abbildung 3: Yamaha BC3.....	14
Abbildung 4: Organisation der Ebenen.....	15
Abbildung 5: Beispiel eines Sample-Mapping.....	15
Abbildung 6: FFT-Analyse verschiedener Klarinetttöne.....	18
Abbildung 7: exemplarische Anwendung des AET-Filters.....	19
Abbildung 8: Sägezahn / Sinus.....	20
Abbildung 9: Sägezahn / Dreieck.....	20
Abbildung 10: additive Anwendung Dreieck -> Sägezahn.....	21
Abbildung 11: Pegeldifferenzen der einzelnen Teiltöne.....	21
Abbildung 12: Auswirkungen unterschiedlicher Glättungsparameter.....	21
Abbildung 13: AET-Nachbildung des pianissimo-Tons.....	22
Abbildung 14: AET-Nachbildung des fortissimo-Tons.....	23
Abbildung 15: Pegeldifferenzen bei additiver Anwendung.....	23
Abbildung 16: statistische Verteilung des Tonhöhenfehlers.....	27
Abbildung 17: schematische Darstellung des zweiten Lösungsmodells.....	28
Abbildung 18: visuelle Repräsentation der unterschiedlichen Wellenformen.....	29
Abbildung 19: Vergleich zwischen Original Spektrum und Wavetable-Spektrum.....	29
Abbildung 20: Unterschiedliche Fensterfunktionen.....	31
Abbildung 21: Implementierung der Fensterfunktion innerhalb KONTAKT.....	31
Abbildung 22: Spektralvergleich mit / ohne Crossfade.....	32
Abbildung 23: Auswirkungen der Mehrfachabtastung auf Wellenformbild.....	34
Abbildung 24: Vergleich der Spektralverläufe von Original- und überlagerter Sinusschwingung.....	34
Abbildung 25: Pegeldifferenz zwischen einfacher und achtfach überlagerter Periode.....	34
Abbildung 26: Vergleich zwischen Original- und interpoliertem Spektrum.....	35
Abbildung 27: Verarbeitung der MIDI-Daten.....	40
Abbildung 28: Hüllkurventypen in KONTAKT.....	41
Abbildung 29: Überblendung der mit Breath-Controllerdaten.....	42
Abbildung 30: Vergleich zwischen Breathcontroller-Verläufen.....	43
Abbildung 31: Unterschiedliche Hüllkurven der Klanganteile.....	44
Abbildung 32: Anzahl der Intervallsamples pro Ton.....	48
Abbildung 33: Funktionsgraph der Tonhöhenabweichungsfunktion.....	52
Abbildung 34: Signalkette der virtuellen Klarinette.....	58
Abbildung 35: FFT-Vergleich zwischen Original und reproduziertem Klarinetton.....	59
Abbildung 36: FFT-Vergleich der einzelnen Dynamikstufen.....	60
Abbildung 37: Entwurf eines mehrdimensionalen Wavetabekonzepts.....	64
Abbildung 38: Legende.....	72
Abbildung 39: Wavetable-Modulatoren.....	72
Abbildung 40: Intervall-Modulatoren.....	73
Abbildung 41: Tonansatzmodulatoren.....	73
Abbildung 42: Luftgeräuschmodulatoren.....	74
Abbildung 43: Nebengeräusch-Modulatoren.....	74