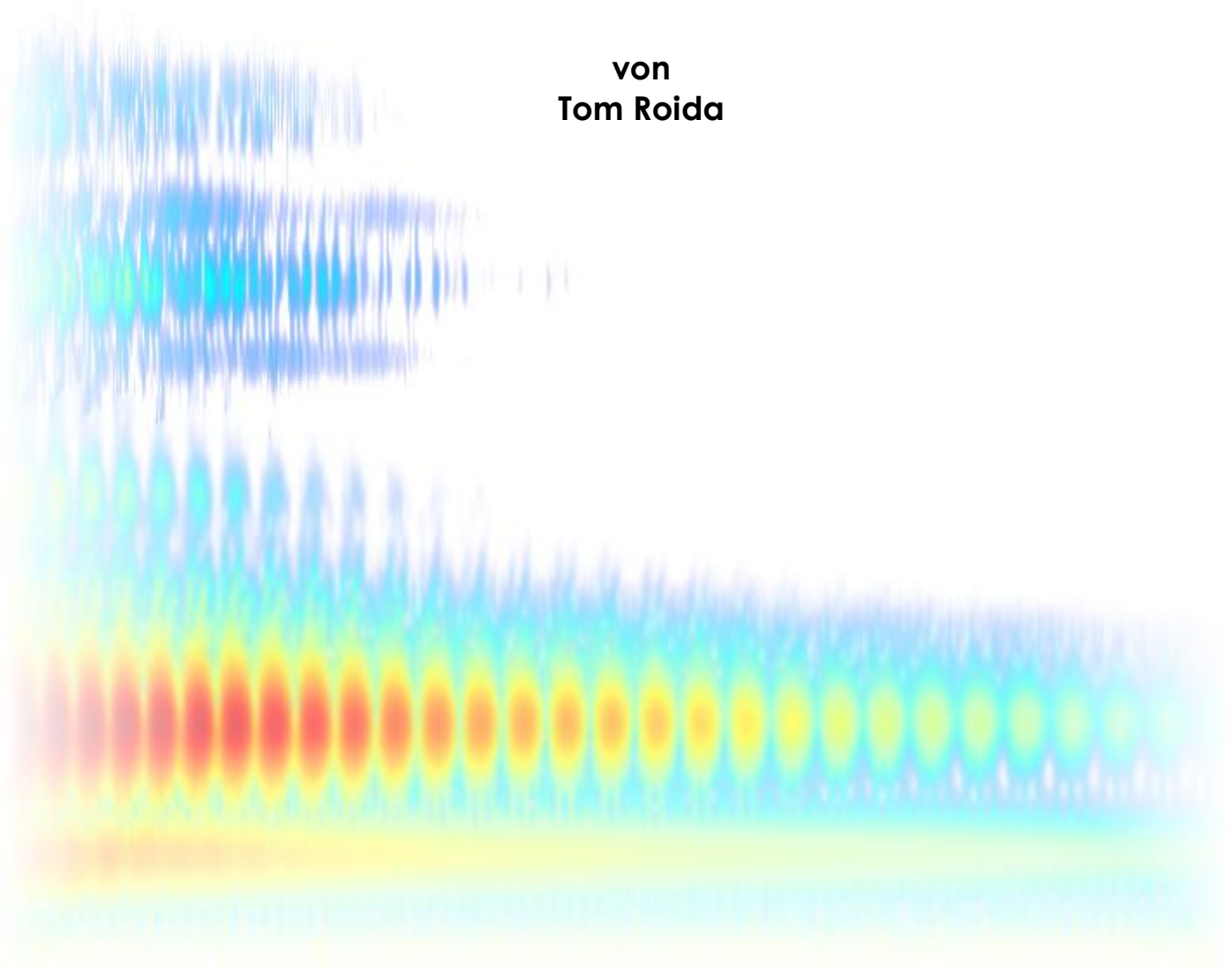


Gehörorientierte Schallanalyse mit VIPER

2. Studienarbeit

von
Tom Roida



Universität der Bundeswehr München
Fakultät für Luft- und Raumfahrttechnik
Institut für Mechanik
Prof. Dr.-Ing. H. Fleischer
85579 Neubiberg

6. Februar 2002
2. Studienarbeit Tom Roida

Zusammenfassung

Das Softwarepaket *VIPER* („Visual Perception“) mit dem Programm *VIPER Perceptual Explorer* der Firma *Neutrik Cortex Instruments*¹ soll eine visuelle Untersuchung von Schallsignalen und Geräuschen ermöglichen. Der Schall wird dabei grafisch so dargestellt, wie er vom menschlichen Gehör wahrgenommen wird. In der folgenden Arbeit wird die Analyse von diskreten Klängen, wie zum Beispiel Gongs und Gitarrensaiten, mit der Software beschrieben und exemplarisch durchgeführt. Geschildert werden die zur Verfügung stehenden Möglichkeiten und eine typische Arbeitsweise mit dem Programm.

Kritisch betrachtet wird die Wirkung der wählbaren Parameter der gehörgerechten Analyse. Außerdem wird in diesem Bezug auf die Qualität der Analyse mit *VIPER* eingegangen. Die Fähigkeit der Software, „das hörbare sichtbar zu machen“, wird überprüft und aus verschiedenen Perspektiven heraus beurteilt.

In Bezug auf die Untersuchung der Musikinstrumente stellt diese Arbeit eine Ergänzung zu der Arbeit von Frau M. Valenzuela dar, die bereits umfangreiche Untersuchungen dieser Art mit einer anderen Software durchgeführt hat.

Diese Arbeit kann auch als Anleitung genutzt werden. Anhand von Beispielen ist die Vorgehensweise mit der Software Schritt für Schritt beschrieben, wodurch ein praktischer Einstieg in die Bedienung ermöglicht werden soll.

Vertieft behandelt wird in der folgenden Arbeit nur ein Teil des Funktionsumfangs von *VIPER* – die FFT-Analyse, die gehörorientierte Darstellung von Frequenzkomponenten („Frequency Contours“) sowie die Bearbeitungs- und Resynthesefunktionen, mit denen die Schallsignale verändert und wieder in Zeitfunktionen zurückgewandelt werden können. Die weiteren Fähigkeiten von *VIPER* werden in einem Überblick vorgestellt.

¹ <http://www.cortex-instruments.de/>

Inhalt

ZUSAMMENFASSUNG	2
INHALT	3
1 GEHÖRORIENTIERTE SCHALLANALYSE MIT VIPER	5
1.1 PRAKTISCHE DURCHFÜHRUNG AN EINEM BEISPIEL	10
1.2 SCHALLPEGEL-KALIBRIERUNG DER DATEIEN	14
1.3 SYNTHESE EINER GROßE-TERZ-GLOCKE AUS EINER KLEINE-TERZ-GLOCKE	16
2 FFT-ANALYSE MIT VIPER	20
3 FFT VS. GEHÖRGERECHTE SCHALLANALYSE	24
3.1 KANN DAS GEHÖR ENG BEIEINANDERLIEGENDE TEILTÖNE UNTERSCHIEDEN?	24
3.2 WELCHE DER SCHWINGUNGSFORMEN EINER GLOCKE ERGEBEN HÖRBARE SCHALLKOMponentEN?	27
4 VERGLEICH DER VIPER-ERGEBNISSE MIT ANDEREN ANALYSEN (VALENZUELA)	29
4.1 GONG III	29
4.2 GONG I	32
4.3 GITARRENSAITE HAEFNER E2	34
5 ZUR BEDEUTUNG DER PARAMETER DER GEHÖRORIENTIERTEN SCHALLANALYSE	35
6 UNTERSUCHUNG VERSCHIEDENER GONGS MIT VIPER	39
7 RESÜMEE UND AUSBLICK	41
ANHANG	44
I DIE BARK FREQUENZSKALA	45

II	GONGS MIT VIPER ANALYSIERT	47
II.1	GONG I, KLEINER KOLBERG GONG	48
II.2	GONG II, KOLBERG-GONG 'EWA'	49
II.3	GONG III, MITTLERER KOLBERG-GONG	50
II.4	GONG IV, MITTLERER KOLBERG-GONG 'TOBI'	51
II.5	GONG V, GROßER KOLBERG-GONG 53 CM	52
II.6	GONG VI, GROßER KOLBERG-GONG GIS 56	53
II.7	GONG VII, GROßER KOLBERG-GONG 'PAPPELSTRASSE'	54
II.8	GONG VIII, JAVA-GONG 'SUWOKAN'	55
II.9	GONG IX, JAVA-GONG 'AGENG'	56
II.10	GONG X, FLACHGONG 'FENG-LUO', SEHR LAUT ANGESCHLAGEN	57
II.11	GONG XI, TAM-TAM 'FIRMA PAISTE'	58
III	LITERATURVERZEICHNIS	59
IV	CD-ROM	60

1 Gehörorientierte Schallanalyse mit VIPER

Das Softwarepaket VIPER bietet die Möglichkeit, Schall auf eine besondere, dem Gehör angepasste Art zu analysieren und die Ergebnisse grafisch darzustellen. Möchte man ein Schallsignal mit VIPER analysieren, so beginnt man normalerweise damit, das Signal in eine für den Computer lesbare Datei umzuwandeln. Beispielsweise kann man es mit einer Soundkarte aufnehmen oder digital überspielen. Damit liegt das Signal als Zeitfunktion vor (üblicherweise als Microsoft[®] Windows[™] *Wave*-Datei mit der Endung „.wav“) und kann in VIPER importiert werden. Im folgenden wird das in VIPER umgesetzte Prinzip der grafischen Repräsentation von Schall vorgestellt, und in Kapitel 1.1 wird konkret auf die Bedienung des Programms eingegangen.

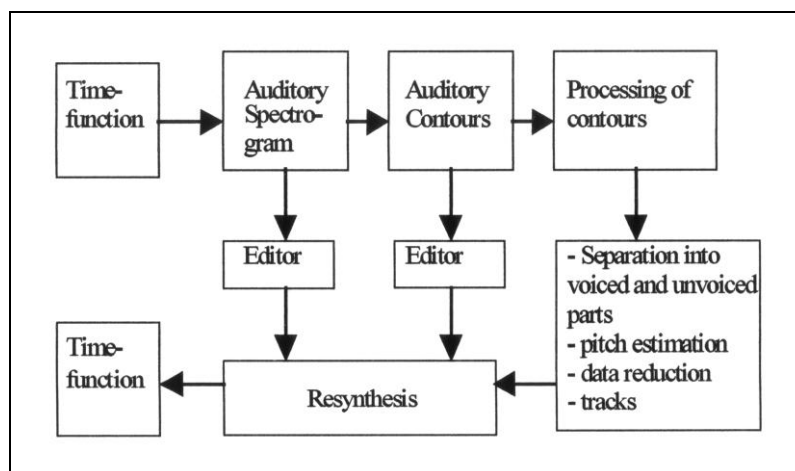


Abb. 1-1: Funktionsumfang und Arbeitsweise von VIPER²

Typischerweise erstellt man zunächst ein *Auditory Spectrogram* (ASP) des zu untersuchenden Schallsignals. Die Analyse basiert auf einer Short-Term Fourier-Transformation (STFT), der Ansatz ist im Handbuch beschrieben³.

Das ASP wird ähnlich wie eine FFT-Analyse dargestellt. Es sollten exakt alle hörbaren Informationen vorhanden sein.

Einzelne Signalkomponenten werden ihrer Stärke⁴ entsprechend farblich hervorgehoben. Man hat die Möglichkeit einige Parameter für die Darstellung zu ändern, jedoch werden vom Hersteller „default“-Werte vorgeschlagen.

² Quelle: VIPER PerceptualXplorer Manual V1.1, S.8

³ VIPER PerceptualXplorer Manual V1.1, S.22

⁴ Siehe auch Kapitel 1.2 *Kalibrierung*.

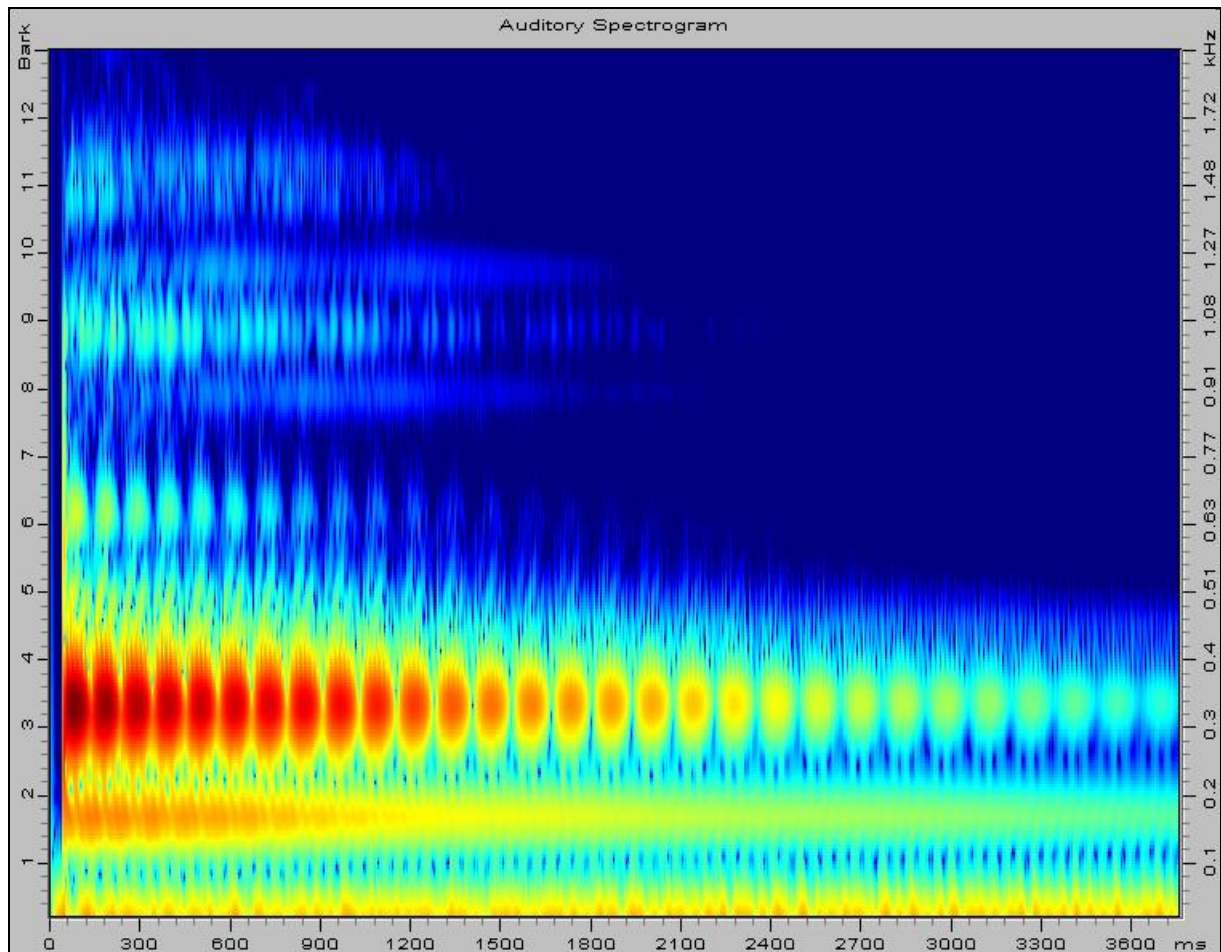


Abb. 1-2: Auditory Spectrogram

An der linken Skala des Diagramms ist die Tonheit in Bark aufgetragen (siehe Anhang), auf der rechten Skala die Frequenz in kHz⁵.

Mit einer speziellen Cursorfunktion kann die Frequenz oder Zeit auf den Diagrammen genau abgelesen werden, außerdem können Verläufe einer Frequenzkomponente über die Zeit sowie das Frequenzspektrum zu einer bestimmten Zeit gesondert dargestellt werden

⁵ Siehe auch Anhang I: Die Bark-Skala

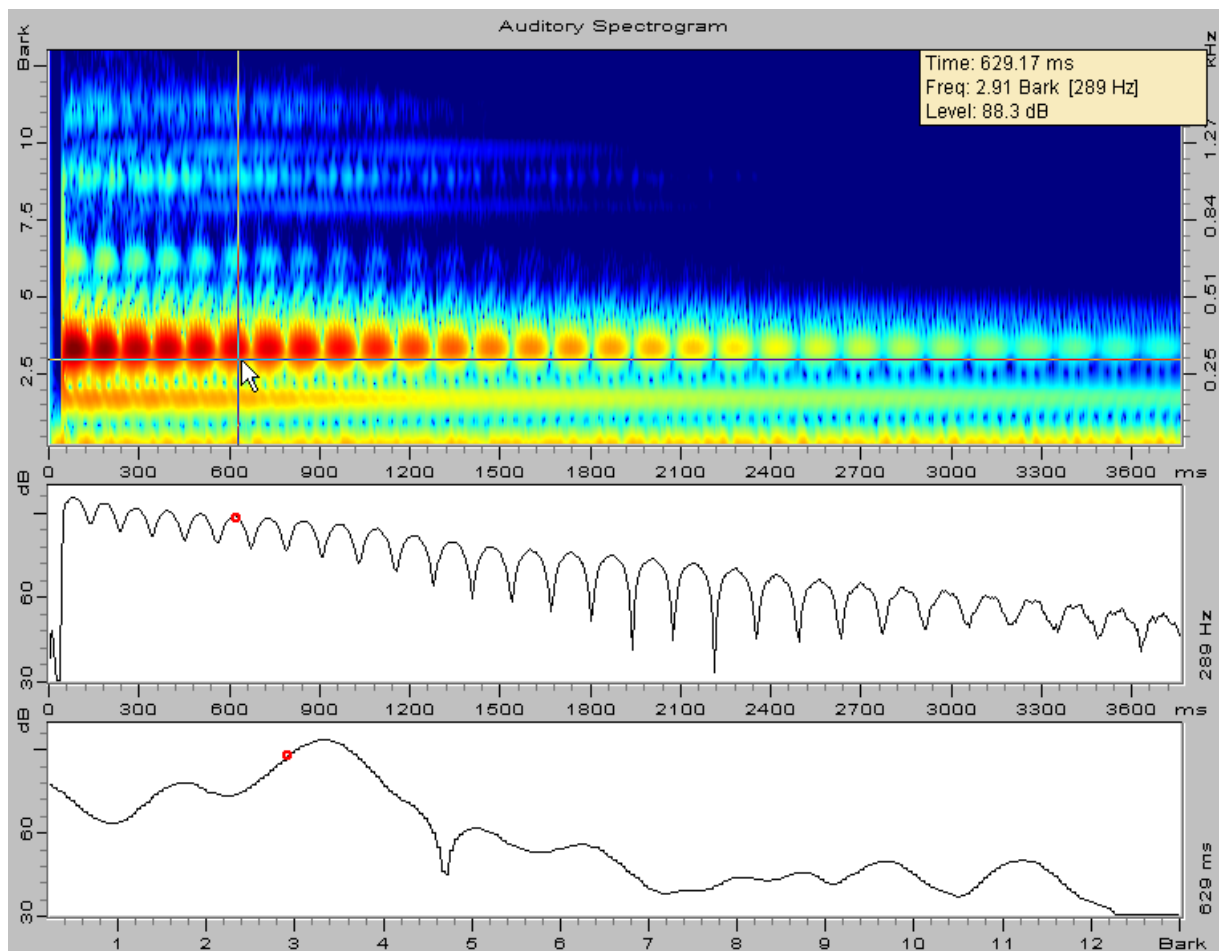


Abb. 1-3: Auditory Spectrogram

Der Benutzer der Software hat die Möglichkeit, mit dem Cursor genaue Informationen zu den Punkten des Diagramms zu erhalten (Frequenz, Zeit, Schallpegel). Ein Spezieller Cursor stellt gleichzeitig mehrere harmonische Vielfache der aktuellen Frequenz als Linien dar. Weiterhin besteht die Möglichkeit, aus dem *Auditory Spectrogram* zeitliche oder tonale Konturen herausrechnen zu lassen und gesondert darzustellen (*Time / Frequency Contours*).

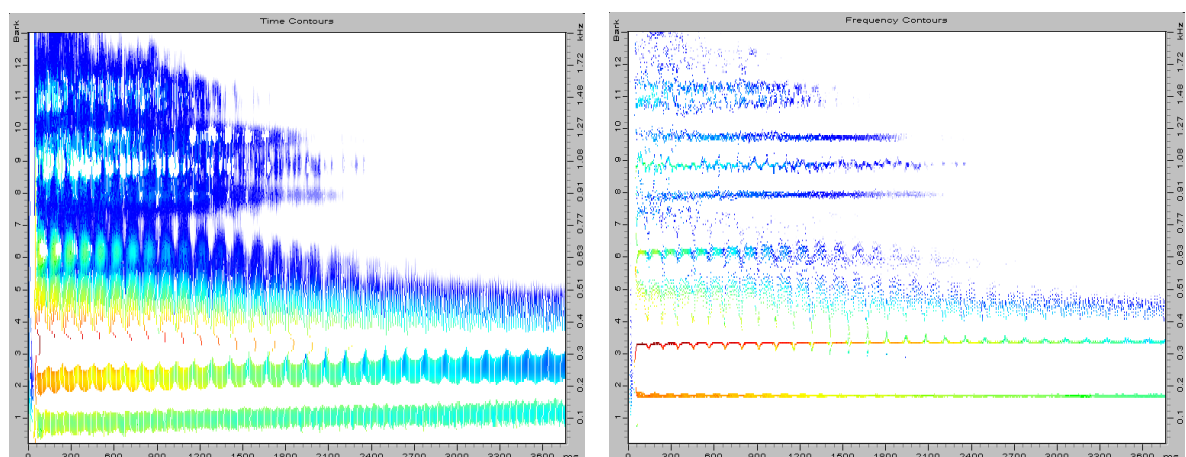


Abb. 1-4: Time Contours / Frequency Contours

Tonale Komponenten können als Frequency Tracks dargestellt werden, dabei werden nur die Teile der *Frequency Countours* übernommen, die zeitlich eine bestimmte Mindestlänge aufweisen. Diese Darstellung enthält nicht mehr alle hörbaren Anteile des ursprünglichen Schallsignals.

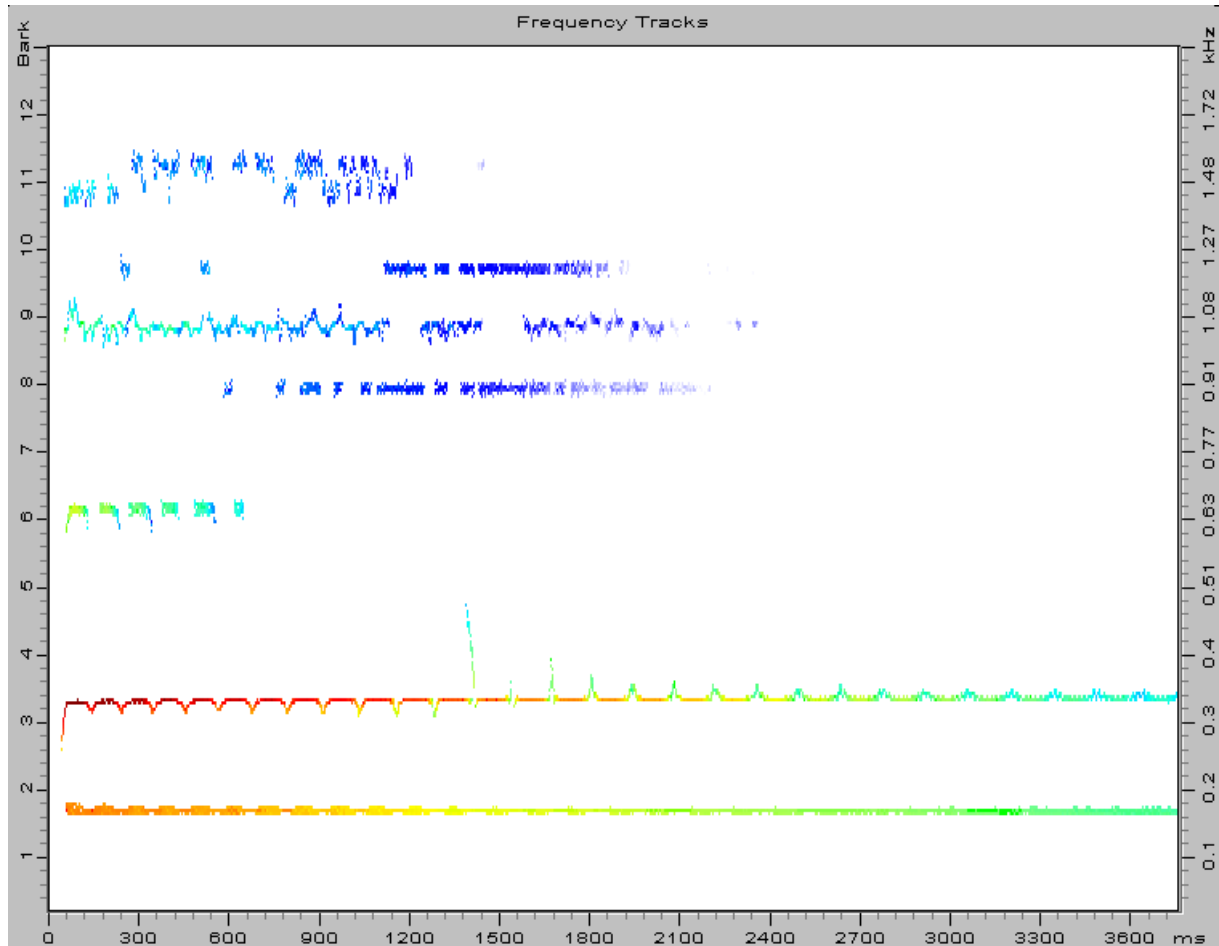


Abb. 1-5: Frequency Tracks

Nach dieser Analyse des Signals kann mit einem Editor eine Bearbeitung erfolgen. Man hat die Möglichkeit

- einen frei mit der Maus gezeichneten Bereich,
- einen rechteckigen Bereich oder
- eine Spur (*Track*)

zu markieren. Der markierte Teil kann nun

- im Pegel abgesenkt oder angehoben werden,
- im Frequenz- oder Zeitbereich verschoben werden oder
- an eine andere Stelle kopiert werden.

Danach kann das veränderte Schallsignal resynthetisiert werden (es entsteht wieder eine Zeitfunktion) und einem Hörvergleich unterzogen werden.

Neben den bis hierher angesprochenen Möglichkeiten bietet VIPER noch weitere Funktionen:

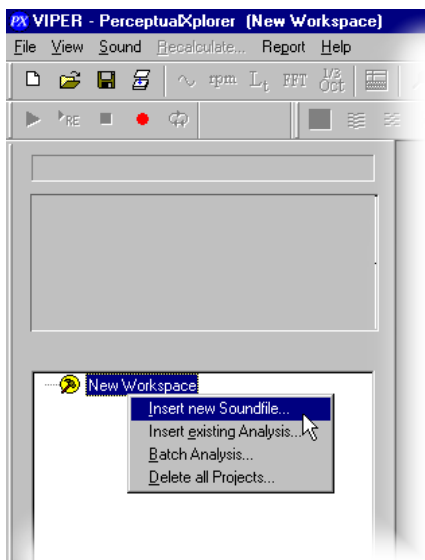
- Schallpegel des Gesamtsignals können errechnet und dargestellt werden (auch A, B oder C-Bewertet⁶).
- *Third-Octave* Analyse;
- Psychoakustische Analyse nach
 - spezifischer Lautheit (kann für Aufnahmebedingungen im freien oder diffusen Schallfeld berechnet werden);
 - Lautheit über die Frequenz in sone / Bark;
 - Lautheit in sone über die Zeit;
 - Schärfe über die Zeit in dA (dezi-Acum);
 - Rauigkeitsspektrum;
 - Rauigkeit über die Zeit.
- Eine Batch-Verarbeitungsfunktion erlaubt es, große Mengen an Schallsignalen automatisiert zu verarbeiten.
- VIPER-Daten können in MATLAB[®] und andere Software importiert werden und somit auf verschiedensten Wegen weiterverarbeitet werden.
- VIPER Speech Quality Analysis ist eine eigenständige Software, die umfangreiche Möglichkeiten zur Sprachanalyse bietet.

⁶ nach DIN EN 60651

1.1 Praktische Durchführung an einem Beispiel

Im folgenden Abschnitt soll exemplarisch ein Gong⁷ mit Viper analysiert werden.

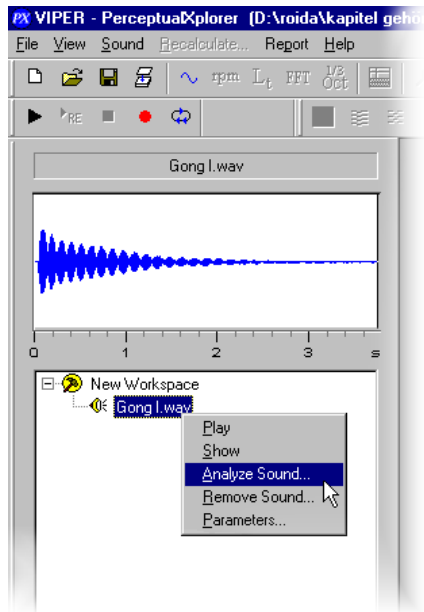
1. Der Gong wird mit einem Mikrofon, Verstärker und DAT Rekorder auf Band aufgenommen.
2. Über den Digitalausgang des Rekorders wird das Signal verlustfrei über den SPDIF-Eingang einer Soundkarte auf den PC-überspielt⁸. Dieser Vorgang hat zwei Vorteile:
 - im Vergleich zu einem analogen Überspielen wird die verlustbehaftete D/A A/D Wandlung vermieden;
 - im Vergleich zu einer direkten Aufnahme mit der Soundkarte ist die Aufnahme mit dem Mikrofonverstärker am DAT-Rekorder qualitativ besser.



3. Die Anwendung *VIPER PerceptuaXplorer* wird gestartet und im Menü *File* wird ein neuer „Workspace“ erstellt. Die Sounddatei wird durch einen Klick mit der rechten Maustaste auf „New Workspace“ importiert.

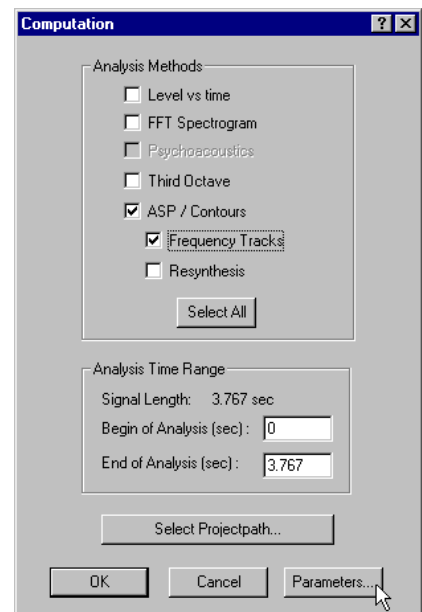
⁷ Siehe „\Beispiele\Kapitel 1_1 praktische durchf/“ auf der CD-Rom

⁸ Siehe Handbuch der Soundkarte.



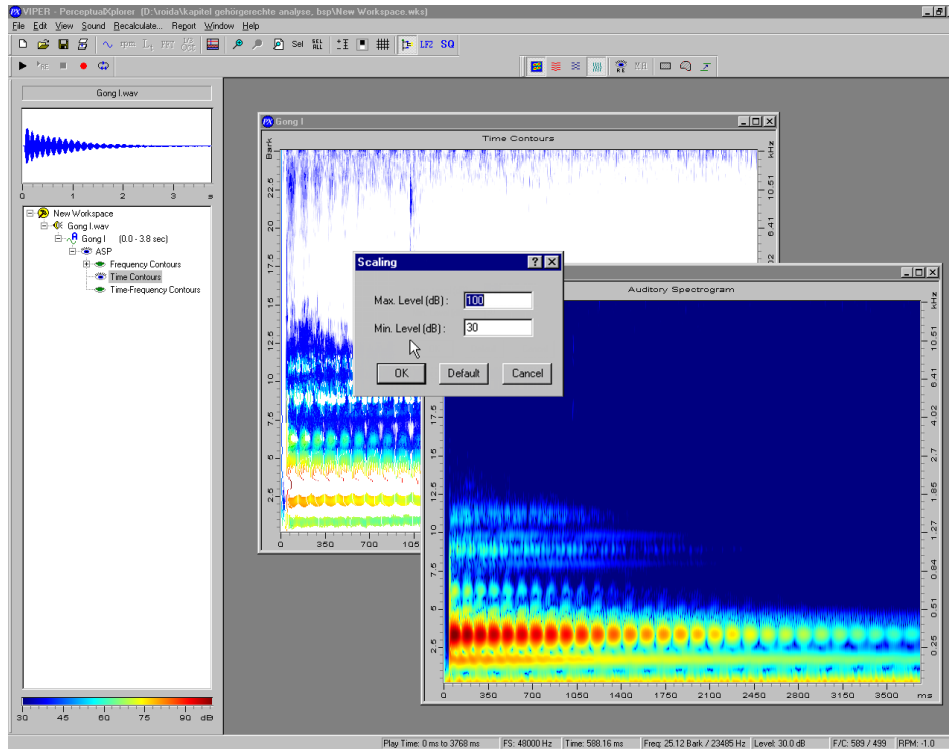
4. Mit einem rechten Mausklick auf die soeben eingefügte Datei und einem weiteren Mausklick auf den Menüpunkt „Analyze Sound...“ wird ein Dialogfeld mit den zu wählenden Analysemethoden eingeblendet.


5. Mit einem Klick auf das Schaltfeld „Parameters...“ werden die Parameter für die Berechnungen eingeblendet. In diesem Beispiel werden die Werte nicht verändert, da sie bereits vom Hersteller an das menschliche Gehör angepaßt sind⁹.
6. Ein Klick auf OK startet die Berechnung. Der Fortschritt wird im Hauptfenster links unten angezeigt.
 - Aktiviert man das Schaltfeld „Low Priority“, so kann man während der Berechnungen noch weiter an dem Computer arbeiten.

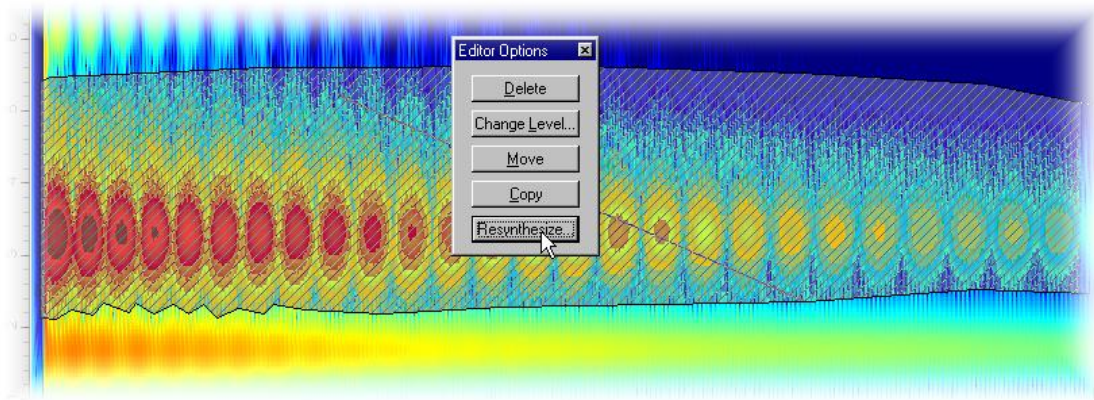


⁹ Siehe auch Kapitel 5 „Zur Bedeutung der Parameter...“

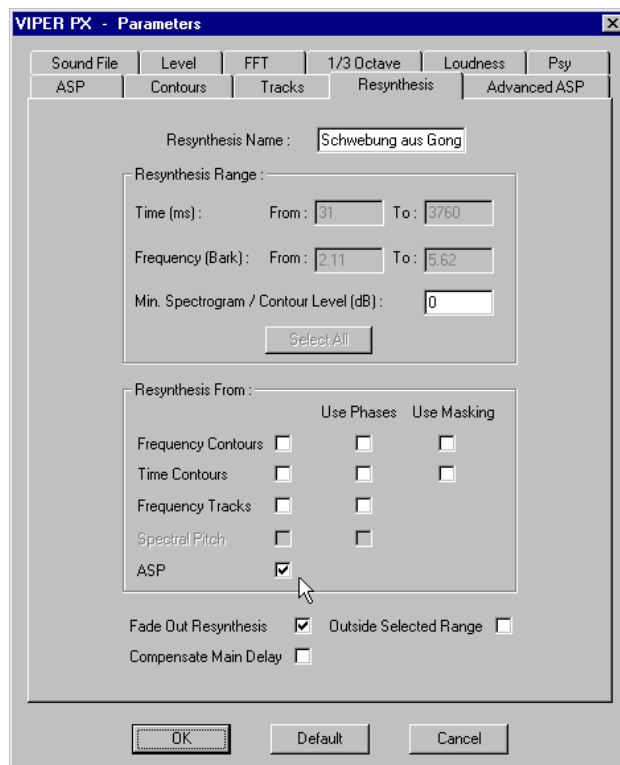
7. Nach Abschluß der Berechnungen werden die einzelnen Diagramme durch Doppelklick in der Projektübersicht (links im Hauptfenster) aufgerufen.





8. Durch einen Klick auf das Symbol „Scale Level Range“ sollte das Hintergrundrauschen ausgeblendet werden. Da der Minimalwert immer auf 0 dB voreingestellt ist, sollte er von Hand auf einen sinnvollen Wert angepasst werden (Siehe auch Kapitel 1.2 Kalibrierung).
9. Nach einem Klick auf die Symbole  wird die Bearbeitung der einzelnen Komponenten möglich. In diesem Beispiel wird der starke Teilton mit der Schwebung für sich alleine resynthetisiert:
 - a. In der Projektübersicht wird ASP ausgewählt.
 - b. Mit einem Klick auf das Symbol zur Polygonauswahl wird die Bearbeitung gestartet. Dann werden mit der linken Maustaste die Eckpunkte des Polygons gesetzt. Am Ende wird mit einem Rechtsklick das Polygon geschlossen und der Inhalt markiert.
 - c. Ein Linksklick auf den markierten Bereich öffnet das Bearbeitungsfenster. Hier wird „Resynthesize...“ ausgewählt.



- d. In dem nun erscheinenden Optionsfenster wird als „Quelle“ für die Berechnung nur ASP ausgewählt (denn nur in dieser Darstellung wurden Veränderungen vorgenommen), und der neuen wav-Datei kann ein Name gegeben werden. Ein Klick auf *OK* startet die Berechnung der Zeitfunktion.



Nach der Berechnung erscheint die neue Datei in der Projektübersicht mit einem roten Lautsprechersymbol. Mit einem Klick auf die Symbole   können die Originaldatei sowie der resynthetisierte Datei abgespielt werden.

1.2 Schallpegel-Kalibrierung der Dateien

Der Schallpegel bzw. Schalldruckpegel ist im allgemeinen definiert als:

$$L = 20 \cdot \text{Log}\left(\frac{\tilde{P}}{\tilde{P}_0}\right), \quad (1.2-1)$$

wobei \tilde{P} den effektiven Schalldruck (SI-Einheit $\frac{N}{m^2}$) darstellt. Der Bezugsdruck ist definiert als

$$\tilde{P}_0 := 20 \frac{\mu N}{m^2} \quad (1.2-2)$$

(entspricht in etwa der Hörschwelle zwischen 700 und 6000 Hz).

In Dateien von digitalisierten Schallsignalen sind allerdings normalerweise nicht der Schalldruck, sondern vielmehr Werte von elektrischen Spannungen gespeichert, die zum Zeitpunkt der Wandlung an einem A/D Konverter anlagen. Als Standard ist in VIPER definiert, dass die maximale Aussteuerung einer 16-Bit-Datei (der Wert $2^{15} = 32768$) genau 93dB entspricht. Dieser Wert kann angepasst werden und entspricht somit nicht dem wahren Dynamikumfang von 90 dB.

Mit folgenden Schritten kann ein Schallsignal kalibriert werden:

1. Ein Referenzsignal (zum Beispiel ein Kalibriersinuston mit definiertem Schallpegel) wird mit den gleichen Einstellungen auf den Rechner übertragen wie das zu untersuchende Schallsignal. Der physikalische Pegel spielt dabei keine Rolle (er sollte aus Qualitätsgründen jedoch nicht zu niedrig sein), beide Dateien müssen nur mit den gleichen Anhebungen bzw. Absenkungen übertragen werden.
2. Die Datei mit dem Referenzsignal wird in VIPER importiert. Im Menu *Sound* wird *Calibration* gewählt. Es öffnet sich ein Dialogfeld mit zwei Feldern für *Max. Peak Level (dB)* und *Sound Pressure Level (dB)*.

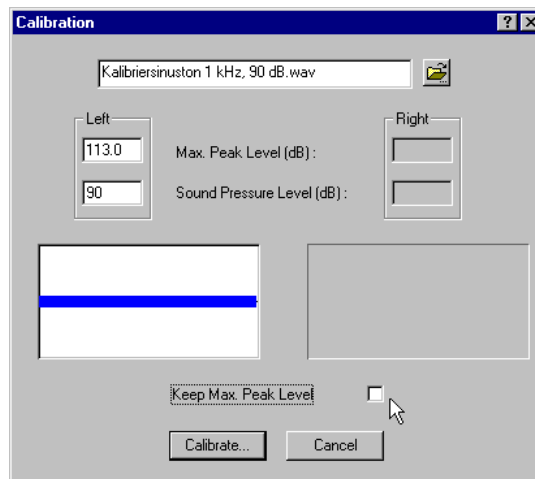



Abb. 1.2.1-1: Kalibrierungsdialog

3. Die Option „*Keep Max. Paek Level*“ wird deaktiviert. Dadurch kann im folgenden der Pegel neu definiert werden.
4. Im Feld „*Sound Pressure Level*“ wird der wahre Wert des Schalldrucks eingetragen. Dadurch verändert sich gleichzeitig der Wert für „*Max. Peak Level*“. **Dieser Wert dient der Skalierung des Schallpegels und muss für alle auf dieses Referenzsignal kalibrierte Dateien eingetragen werden!**
5. Nach einem Klick auf „*Calibrate...*“ erscheint ein Dialogfeld zum Abspeichern der Datei. Da die Informationen von VIPER direkt in die *Wave*-Datei geschrieben werden, muss diese neu gespeichert werden.
6. Die zu untersuchende Datei wird importiert und im Kalibrierungsdialog wird der gleiche Wert (in diesem Fall 113 dB) bei „*Max. Peak Level*“ eingetragen. Dieser Vorgang muss auch für alle weiteren Dateien manuell wiederholt werden.
7. Auch hier muss die Datei neu gespeichert werden. Wird sie auch unter einem neuen Namen gespeichert, so muss sie im Folgenden neu importiert werden!
8. Die Einstellungen können überprüft werden, indem das Signal mit der Methode *Level vs. Time* analysiert wird. Unter *Parameters...* muss die Option *Frequency Weighting* auf *none* eingestellt werden, um eine unbewertete Schallpegelanalyse zu erhalten. Der Schallpegel des Kalibriersinustones in der erscheinenden Grafik muss nun dem definierten Wert entsprechen.

Anmerkung: Um verschiedene Schallsignale mit VIPER zu vergleichen, ist es nicht nur sinnvoll, alle Dateien gleich zu kalibrieren, sondern außerdem sollte die Farbskala zur Darstellung bei allen Grafiken die gleichen Minimal- bzw. Maximalwerte haben. Dies muß manuell mit dem Symbol  durchgeführt werden. Im Gegensatz zu den Kalibrierdaten, die VIPER direkt in der *.wav*-Datei abspeichert und die man somit nur einmal für jedes Schallsignal einzugeben hat, muss man die Skala für jede Grafik einzeln einstellen.

1.3 *Synthese einer Große-Terz-Glocke aus einer Kleine-Terz-Glocke*

Am folgenden Beispiel sollen die Bearbeitungsmöglichkeiten der Schallsignale aufgezeigt werden. Dazu wird der Klang einer Perner Glocke (34 kg) zunächst analysiert (siehe dazu Kapitel 2.2) und ausgewertet. In der Abbildung 1.3-1 (*Frequency Contours*) sind die einzelnen tonalen Komponenten der Glocke zu sehen.

Die Teiltöne:

- Suboktave (0,51 kHz, Verhältnis 0,49),
- Prime (1,05 kHz, Verhältnis 1,00),
- **Moll- oder auch kleine Terze (1,26 kHz, Verhältnis 1,20),**
- Quinte (1,61 kHz, Verhältnis 1,52) und
- Oberoktave (2,12 kHz, Verhältnis 2,00).

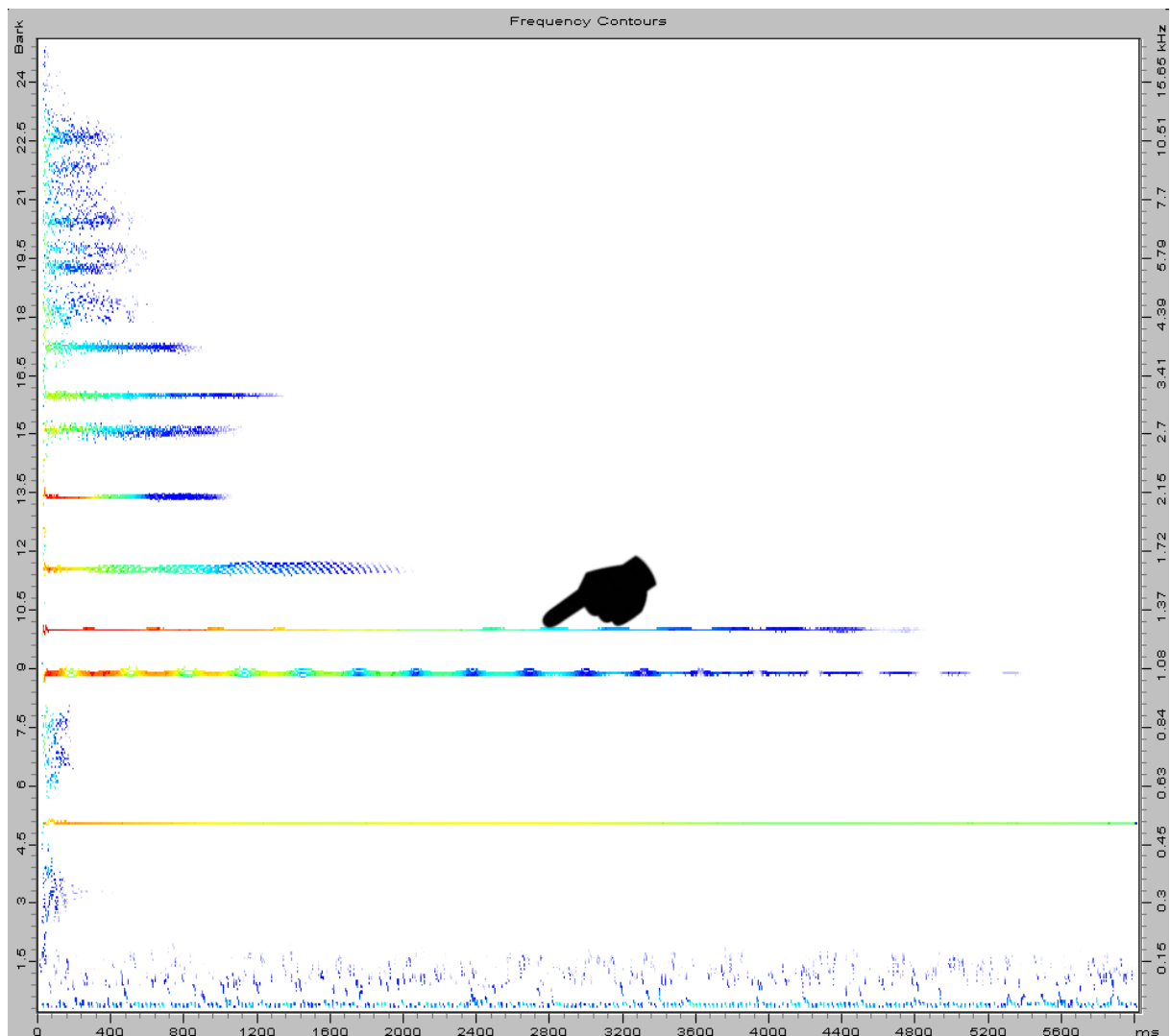



Abb. 1.3-1: Perner Glocke mit markierter kleiner Terze

Um nun einen Klang zu erzeugen, der dem einer Großen-Terz-Glocke ähnelt, muss die Terze ein Frequenzverhältnis von 1,25 aufweisen, das heißt die Spur muss auf 1,31 kHz verschoben werden:

1. VIPER wird gestartet, ein neuer *Workspace* erstellt und die zu bearbeitende Datei importiert.
2. *Frequency Contours* werden mit den gewünschten Parametern berechnet.
3. Nun wird als Referenz das unbearbeitete Schallsignal anhand genau dieser Darstellung resynthetisiert. Dazu wird mit der rechten Maustaste auf das Diagramm *Frequency Tracks* geklickt und *Resynthesis...* ausgewählt. Es erscheint das bekannte Fenster mit den Optionen und als Quelle ist nur *Frequency Contours* ausgewählt. Als Dateiname wird ein aussagekräftiger Name eingetragen, zum Beispiel „*Glocke unbearbeitet mit kleiner Terz.wav*“.
4. Mit *OK* wird die Berechnung gestartet, und es erscheint ein rotes Lautsprechersymbol mit dem Dateinamen in der Projektansicht. Klickt man darauf mit der rechten Maustaste und dann auf *Play*, kann das Ergebnis angehört werden.
5. Jetzt wird wieder das Diagramm *Frequency Contours* ausgewählt und mit der Lupe wird in den Frequenzbereich um 1,3 kHz gezoomt.
6. Im Menü *Ansicht* wird *Quick Info* eingeschaltet – dadurch wird in einem kleinen Feld die exakte Frequenz an der Cursorposition eingeblendet. Dies wird das Positionieren der Spur erleichtern.
7. Mit den Symbolen  wird in den Bearbeitungsmodus gewechselt.
8. Die ganze Terz-Spur wird ausgewählt (zum Beispiel durch Umrandung) und im daraufhin erscheinenden Menü (linke Maustaste auf den markierten Bereich) wird *Move* ausgewählt.

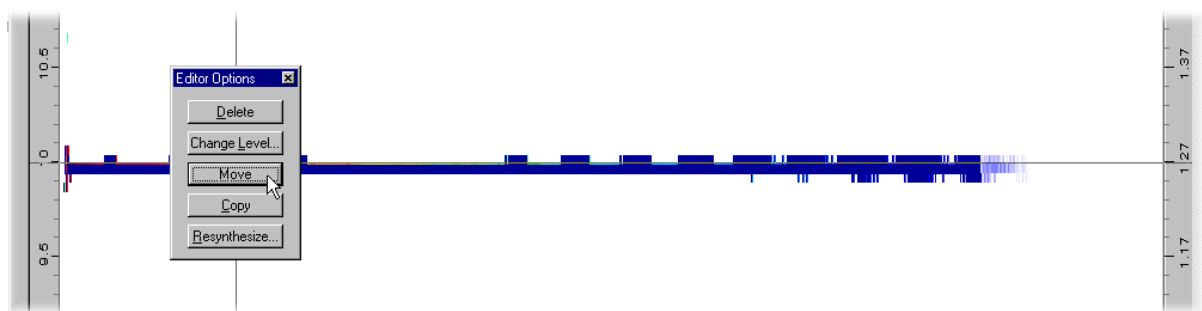


Abb. 1.2-2: Die tonale Komponente bei 1,27 kHz wird bearbeitet

9. Die Spur wird auf ca. 1,31 kHz verschoben (linker Mausklick zum „loslassen“).

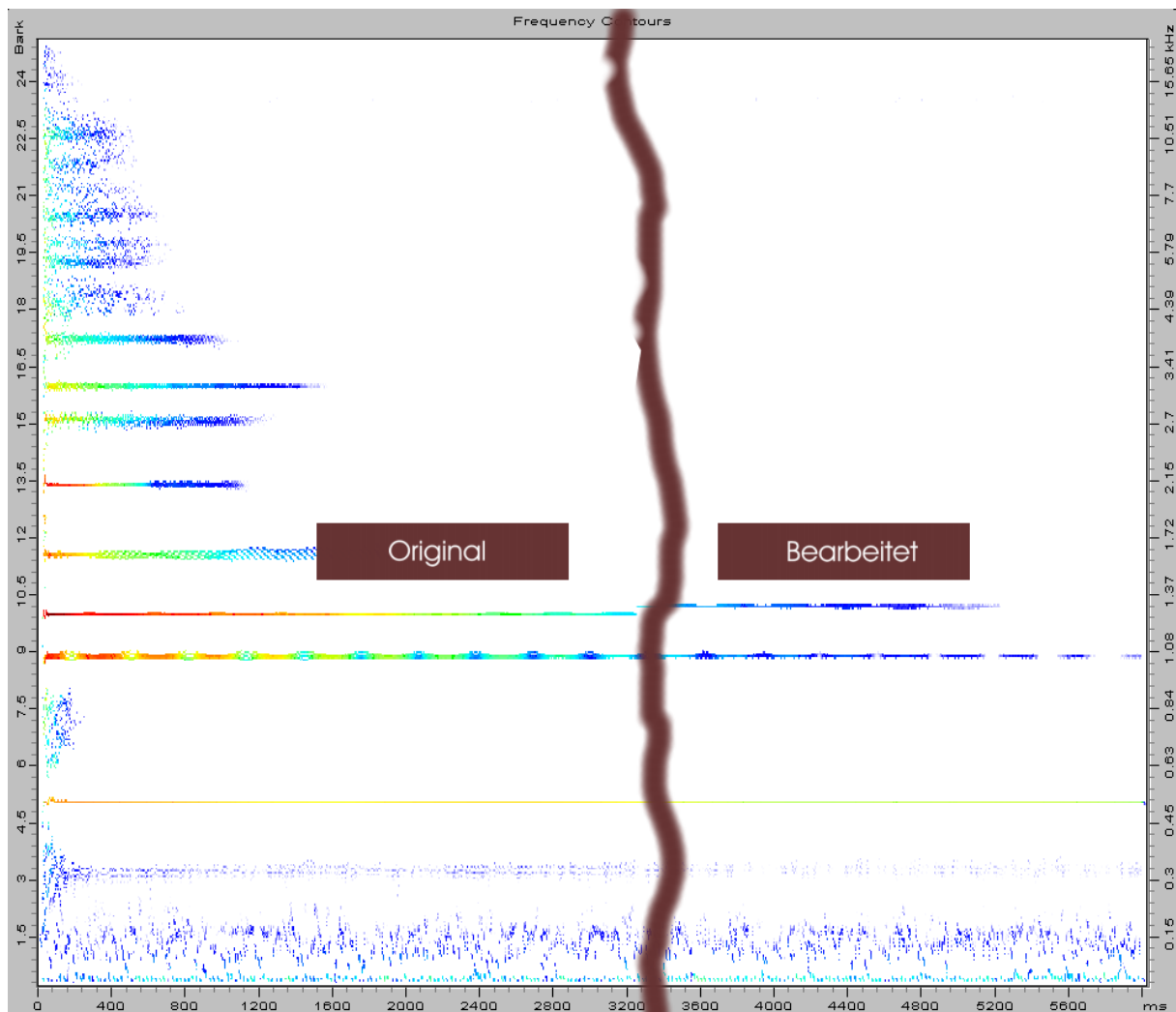


Abb. 1.2-3: Die tonale Komponente bei 1,31 kHz wurde verschoben

10. Wieder wird nach einem Klick mit der rechten Maustaste auf *Frequency Tracks* in der Projektansicht *Resynthesis...* ausgewählt.
11. Im Optionsfenster wird die Datei benannt (zum Beispiel „*Glocke editiert mit grosser Terz.wav*“) und die Resynthese gestartet.
12. Es erscheint ein zweites rotes Lautsprechersymbol mit der neuen Datei¹⁰. Auch sie kann mit der rechten Maustaste abgespielt werden.

Anmerkung:

Der charakteristische Klang des Anschlags besteht aus vielen Frequenzkomponenten, die sich mit der Zeit stark ändern. Mit der hier durchgeführten Bearbeitung ändert er sich zwar auch, aber sicher wird er nicht so wie der Anschlag einer „echten“ Große-Terz-Glocke klingen.

¹⁰ Die *.wav*-Dateien zu diesem Projekt finden sich unter „\Beispiele\Kapitel 1_3 Terze\“.

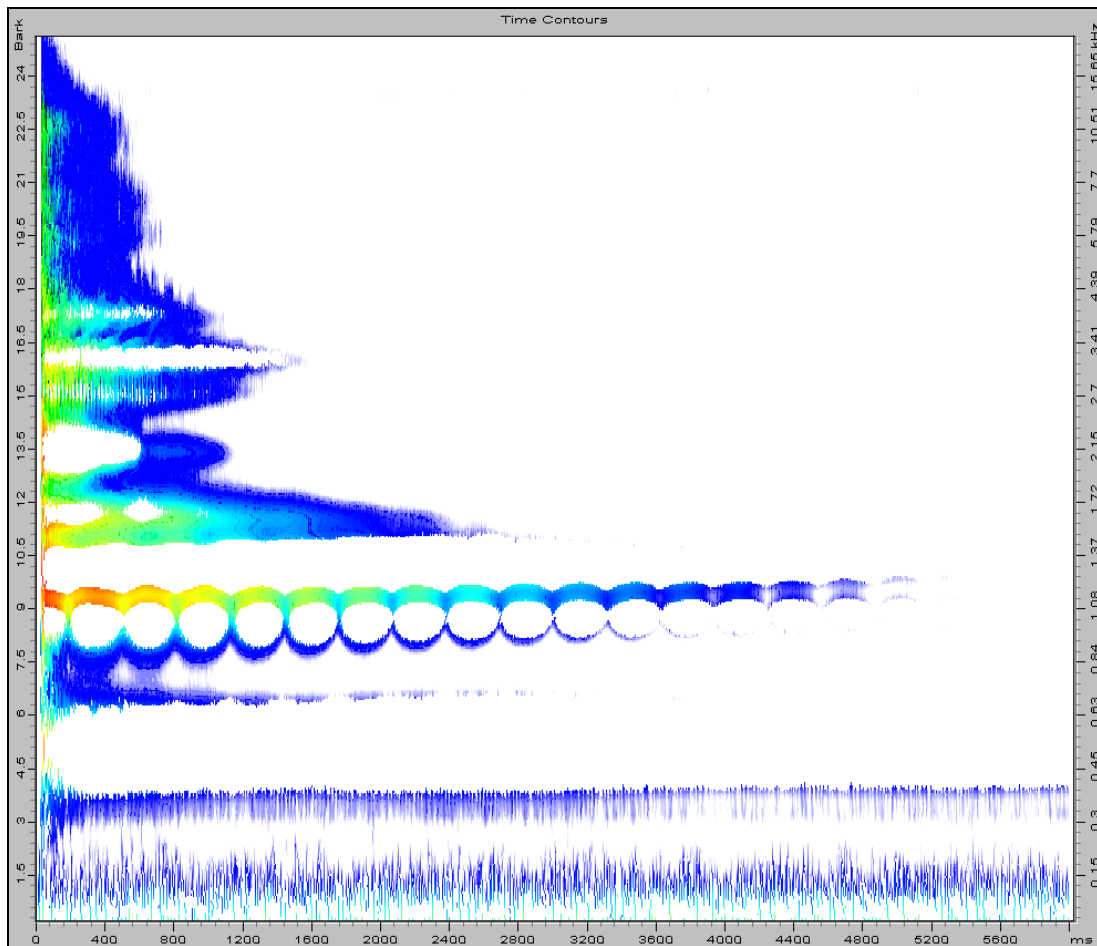


Abb. 1.2-4: Time Contours

Gezielte Änderungen in dieser Hinsicht könnten zum Beispiel in der Darstellung *Time Contours* vorgenommen werden und dann in die Resynthese mit eingebunden werden.

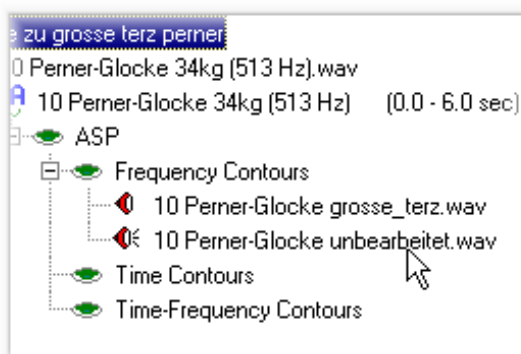



Abb. 1.2-5

Anmerkung:

Sind mehrere Sounddateien in einem Projekt vorhanden, so werden bei einem Klick auf die Symbole  genau die abgespielt, die ein aktives Lautsprechersymbol vor ihrem Dateinamen haben (siehe Abbildung 1.2-5).

2 FFT-Analyse mit VIPER

Neben der gehörgerechten Analyse bietet VIPER auch die Möglichkeit zur „klassischen“ FFT-Analyse. Die Parameter sind hierbei auf den Frequenz- und Auflösungsbereich für akustische Anwendungen beschränkt.

Die Überprüfung der FFT-Analysefunktionen von VIPER wurde anhand des Schallsignals eines Burma-Gongs¹¹ vorgenommen. Als Referenz wurde eine Analyse mit einem Mehrzweck-FFT-Analysator¹² zugrunde gelegt. Abbildung 2-1a zeigt das Wasserfall-Diagramm mit sehr hoher Frequenzauflösung im Bereich zwischen 25 Hz bis 500 Hz. Das Zeitfenster hat eine Länge von 1600 ms.

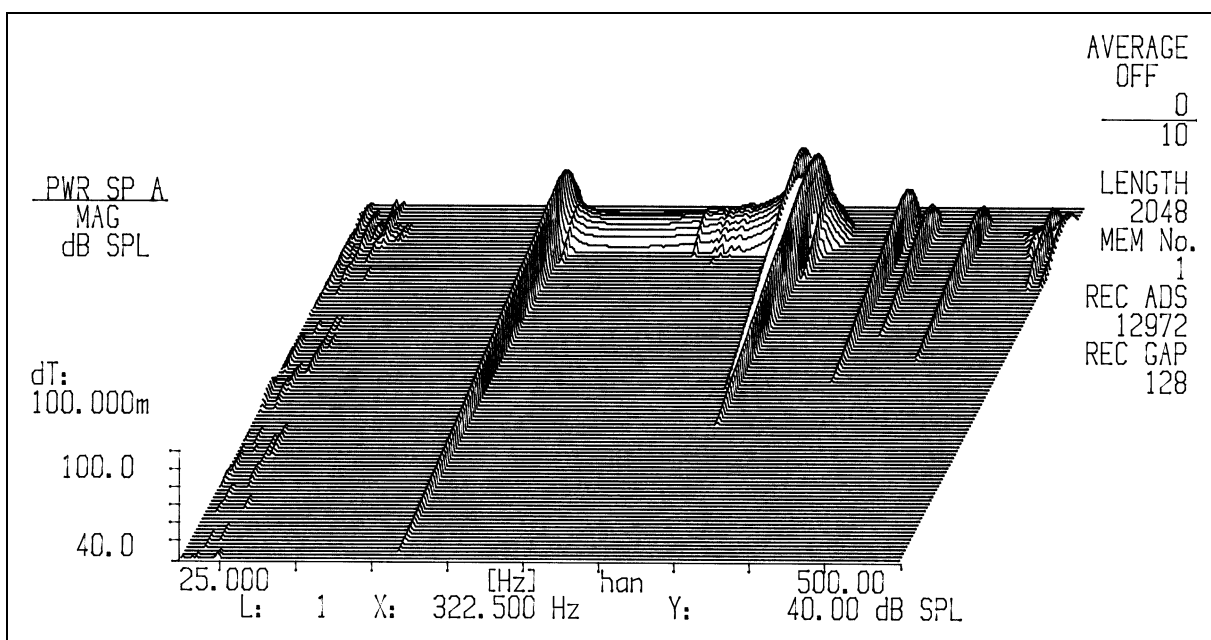


Abb. 2-1a: Burma-Gong FFT Wasserfall-Diagramm¹³

¹¹ Verwendet wurde das vom DAT-Tape digital überspielte Schallsignal „\Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62\61 Gong I, kleiner Kolberg-Gong, frontal“ bzw. „\Beispiele\Kapitel 2 Gong I\61 Gong I, kleiner Kolberg-Gong, frontal, 11khz + silence calibrated.wav“.

¹² ONO SOKKI CF-5210

¹³ Quelle: Prof. Fleischer, UniBw München

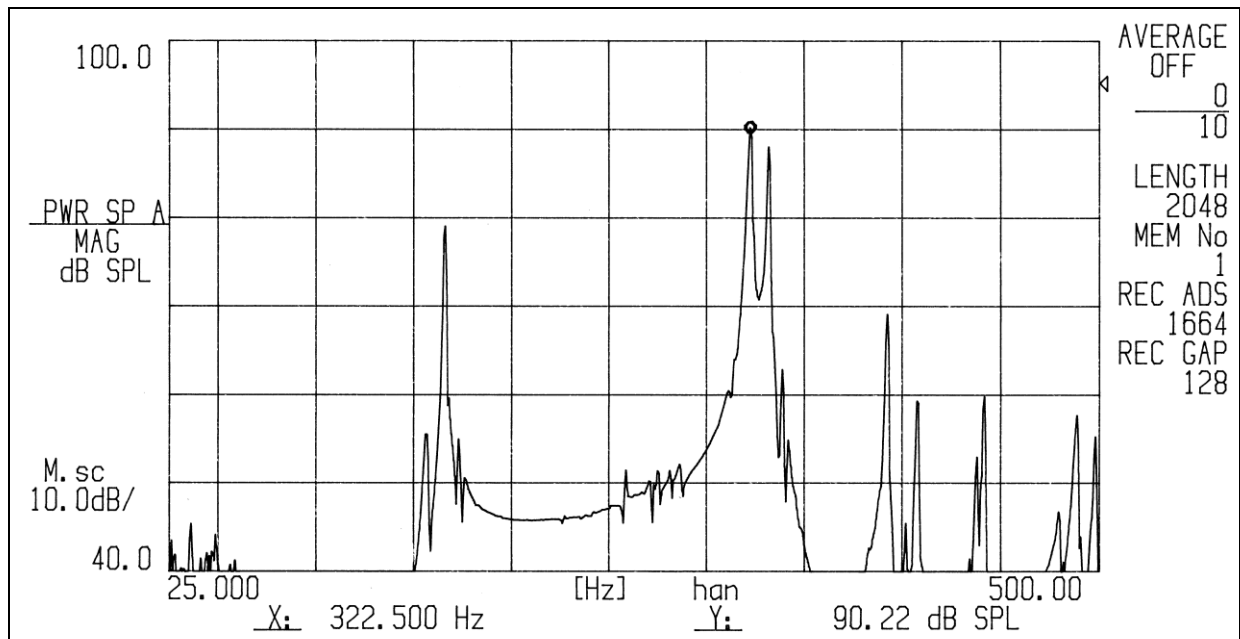


Abb. 2-1b: Burma-Gong FFT Diagramm¹⁴

Abbildung 2-1b zeigt einen Schnitt zur Zeit des maximalen Schallpegels. Abbildung 2-2a zeigt das *FFT-Spectrogramm* des gleichen Schallsignals¹⁵, untersucht mit VIPER, im entsprechenden Frequenzbereich und mit äquivalenten Parametern. Der obere Teil entspricht dabei dem Wasserfall Diagramm, allerdings als 2D-Ansicht mit farblicher Kodierung des Schallpegels. Der mittlere Teil zeigt den zeitlichen Verlauf der Schallkomponente bei 330 Hz, und der untere Teil entspricht der Darstellung aus obiger Abbildung 2-1b. Die Parameter wurden gemäß Abbildung 2-2b eingestellt.

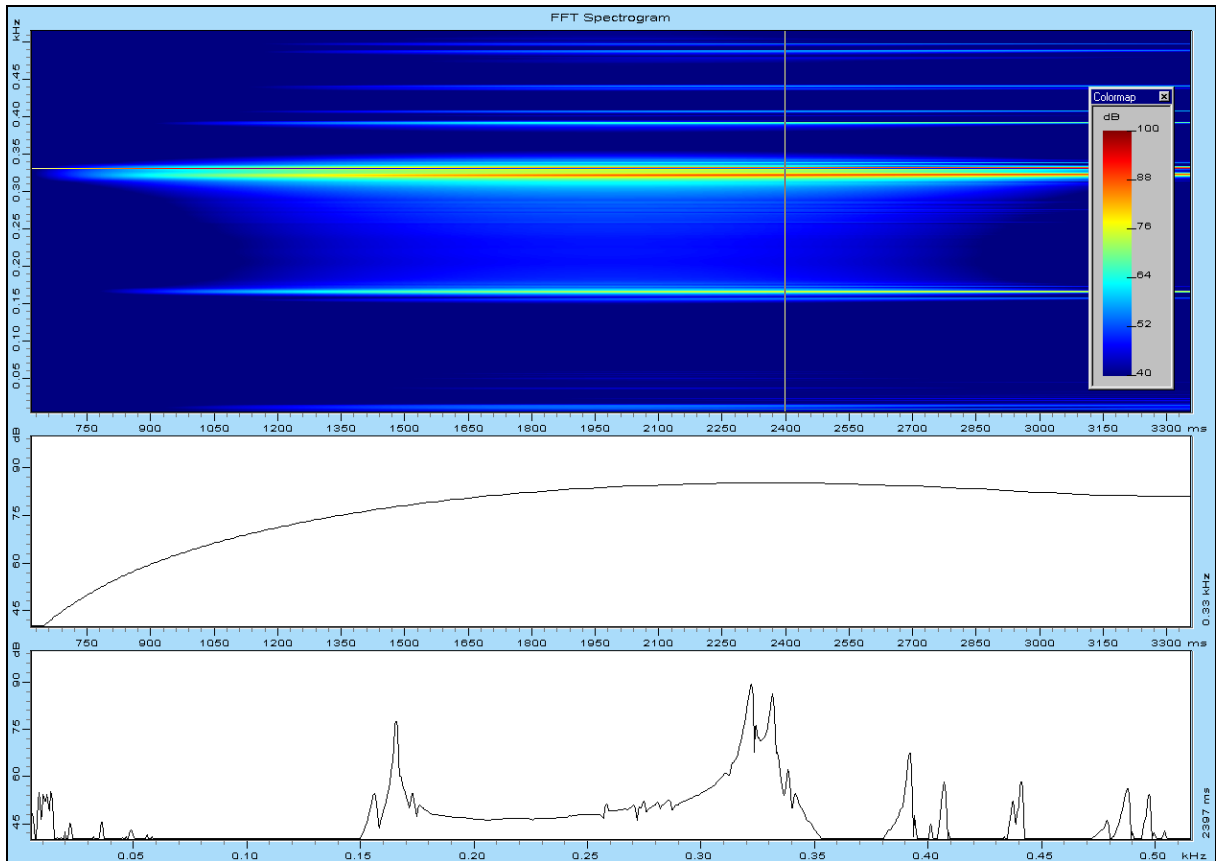
Gemeinsamkeiten und Unterschiede der beiden Analysen:

- Die Verläufe über den Frequenzbereich sind fast identisch. Unterschiede sind geringer als 2dB.
- Die Maxima über den Frequenzbereich (z.B. 165 Hz, 322 Hz, 332 Hz) finden sich an der gleichen Stelle (bei einer Ablesegenauigkeit von ca. 2 Hz).
- Der Verlauf über die Zeit ist ähnlich, soll hier aber wegen des langen Zeitfensters nicht weiter betrachtet werden.

¹⁴ Quelle: Prof. Fleischer, UniBw München

¹⁵ Um die Berechnung mit VIPER zu ermöglichen – siehe Kasten „Anmerkungen“ in diesem Kapitel – wurde die Samplingrate mit dem Programm „Cool Edit 2000“ auf 11 kHz reduziert.

2 FFT-Analyse mit VIPER



Oben: FFT-Spektrogramm

Mitte: Zeitlicher Verlauf der Frequenzkomponente bei ca. 330Hz

Unten: Frequenzspektrum bei ca. 2400 ms

Abb. 2-2a: FFT-Analyse mit VIPER (mit Schallpegel-Skala)

Hinweise zu den Optionen:

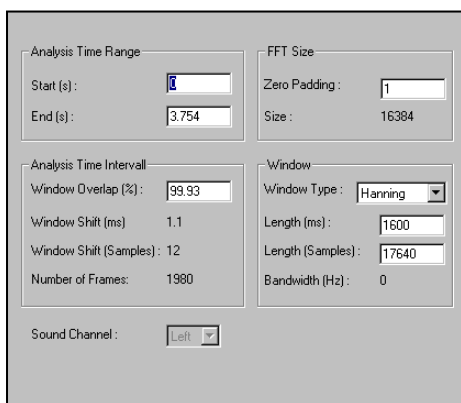


Abb. 1-2b: Optionen

- Nur die weiß hinterlegten Felder können manuell geändert werden, die anderen Werte errechnet das Programm automatisch.
- Bei einer Eingabe für „*Window Length (ms)*“ wird der Wert für „*Window Length (samples)*“ automatisch eingefügt und umgekehrt.
- Die Fensterbreite wird auf ganze Zahlen in Hertz gerundet.
- Die Werte sind größtenteils limitiert, siehe auch weiter unten in diesem Kapitel.

Anmerkungen zu der praktischen Benutzung:

- VIPER kann nur eine begrenzte Anzahl von Samples innerhalb eines FFT-Zeitfensters bearbeiten. Wünscht man eine Analyse mit hoher Frequenzauflösung, so muß man unter Umständen die 48-kHz-DAT-Soundfiles in eine niedrigere Samplingrate konvertieren (dazu wird ein externes Programm benötigt). Beispiel: für eine Frequenzauflösung von ca. 1 Hz wird ein Zeitfenster von über 1½ Sekunden benötigt; das sind beim DAT-Tape 72000 Samples, VIPER erlaubt aber maximal 32768 Samples¹⁶.
- Die Option „ZERO padding“ (*FFT Size, FFT Parameters*) in VIPER bewirkt im Fall ≥ 1 ein Vergrößern des FFT-Zeitfensters mit Nullen. Dadurch wird die Auflösung im Frequenzbereich scheinbar erhöht (feinere Übergänge zwischen den Frequenzlinien), der Informationsgehalt bleibt allerdings gleich.
- Die FFT-Size darf bei VIPER den Wert 32768 nicht übersteigen.
- Das erste Zeitfenster zur Analyse beginnt bei $t=0.00$ s und damit würde bei einer Dauer von ca. 1,6 s der größte Teil des Gongs bereits mitverarbeitet werden, wenn die .wav-Datei nicht eine „Pause“ von 1.6 s am Anfang enthalten würde (siehe Abbildung 2-3). Der FFT-Analysator ONO SOKKI beginnt die Analyse dagegen im „leeren“ Bereich vor dem eigentlichen Schallsignal.

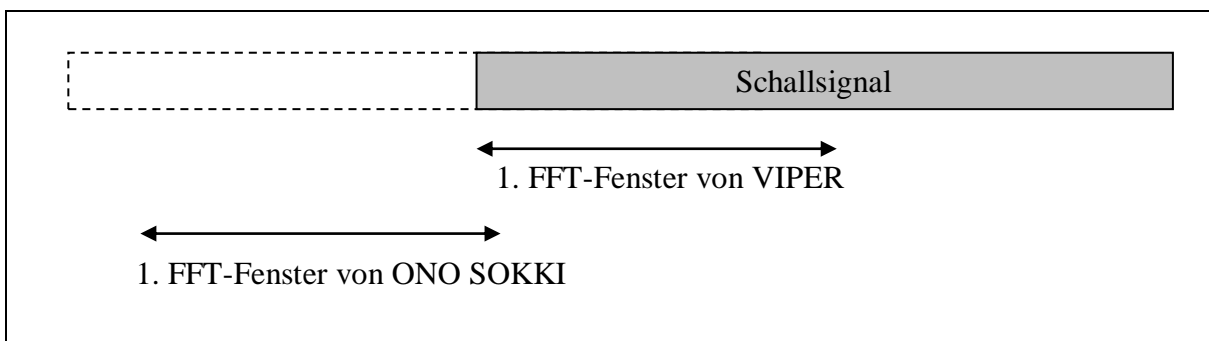


Abb. 2-3: Arbeitsweise der FFT-Analyse, 1. Zeitfenster

¹⁶ Ein von mir experimentell bestimmter Wert.

3 FFT vs. Gehörgerechte Schallanalyse

Anhand zweier typischer Fragestellungen werden im Folgenden die Vorteile der gehörgerechten Schallanalyse gegenüber der klassischen FFT-Analyse herausgearbeitet.

3.1 Kann das Gehör eng beieinanderliegende Teiltöne unterscheiden?

Ein Klangbild, beispielsweise das eines Gongs, besteht meist aus mehreren Frequenzkomponenten, die sich mit der Zeit ändern. Eine FFT-Analyse (**F**ast **F**ourier **T**ransformation) zerlegt das Signal immer in Teile gleicher Frequenzbreite und Länge. Der Benutzer muss sich für einen Fenstertyp sowie für eine bestimmte Auflösung entscheiden. Zeit- sowie Frequenzauflösung stehen dabei in einem festen Verhältnis. Die Abbildungen 3.1-1a und 3.1-1b zeigen je eine Analyse des gleichen Schallsignals¹⁷. Abbildung 3.1-1a zeigt eine hohe Zeit- und geringe Frequenzauflösung. Abbildung 3.1-1b umgekehrt.

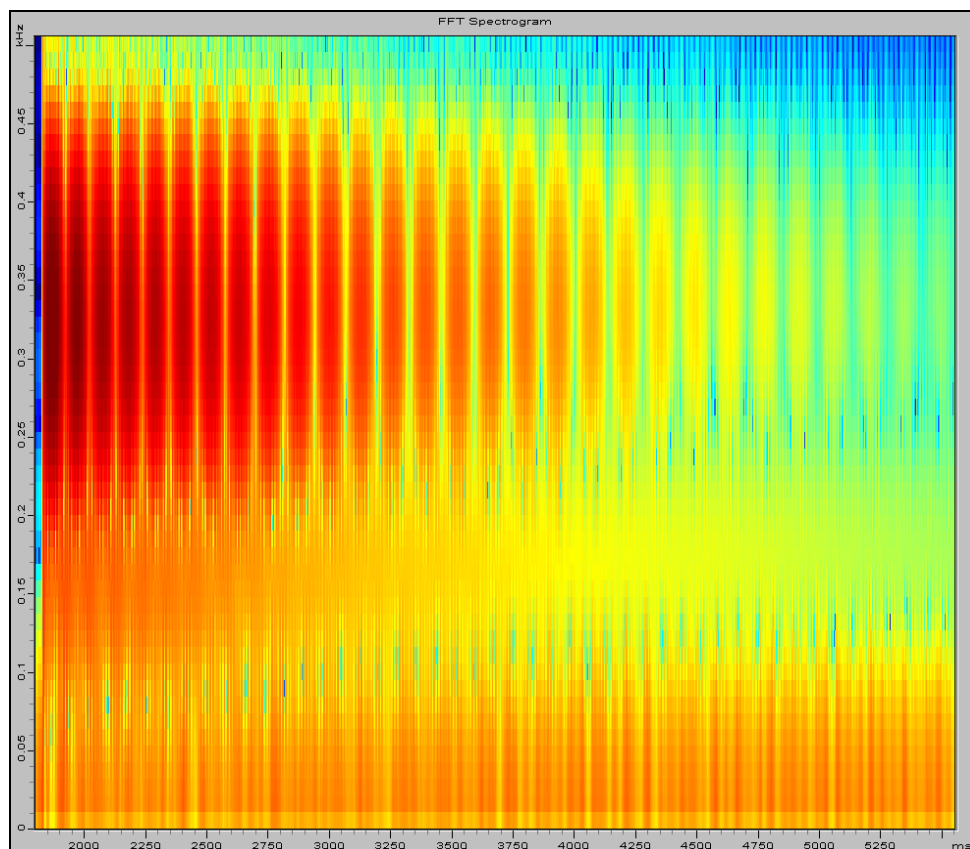


Abb. 3.1-1a: *FFT-Spectrogram* mit hoher Zeitauflösung

¹⁷ Gong I, siehe Archiv auf der CD-Rom

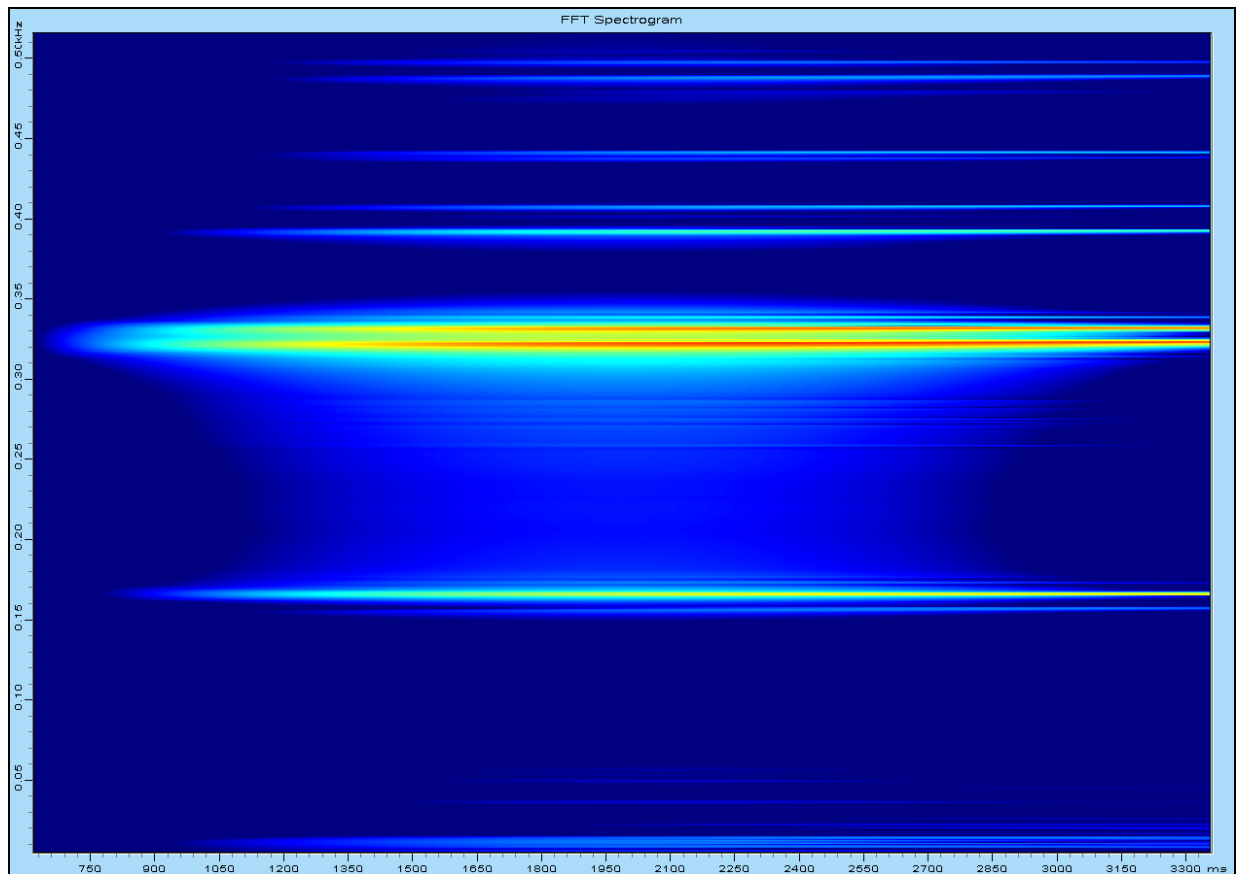


Abb. 3.1-1b: *FFT-Spectrogram* mit hoher Frequenzauflösung

Selbst wenn physikalisch eine hohe *Zeit* *und* eine hohe Frequenzauflösung möglich wäre, so entspräche das Bild sicherlich nicht dem, was das menschliche Gehör wahrnehmen kann. In jedem Fall bleiben dem Benutzer bei der FFT-Analyse willkürlich wählbare Parameter, die zu sehr unterschiedlichen Ergebnissen führen können.

Die gehörgerechte Schallanalyse bietet einen Ansatz, der zwar die Wahl von Parametern nicht ganz ausschließt¹⁸, aber in jedem Fall reduziert. Durch die besondere (nichtlineare) Skalierung im Frequenz- und Zeitbereich sowie eine exponentiell abklingende Fensterfunktion wird den physischen Gegebenheiten des menschlichen Ohres eher entsprochen.

¹⁸ siehe Kapitel 5 „Zur Bedeutung der Parameter der gehörorientierten Schallanalyse“

Abbildung 3.1-1c zeigt die mit VIPER erstellte gehörorientierte Analyse des gleichen Schallsignals. Folgende Informationen lassen sich hieraus gewinnen:

- Die beiden Schwingungen bei ca. 330 Hz werden nicht getrennt voneinander wahrgenommen. Statt dessen ergibt sich ein schwebender Teilton, das heißt ein Ton mit stark schwankender Amplitude (mit einer Frequenz von ca. 8 Hz).
- Nach dem Anschlag sind hauptsächlich zwei tonale Komponenten hörbar, eine bei ca. 330 Hz und eine bei ca. 170 Hz.
- Im tieffrequenten Bereich ist ein Rauschen hörbar.

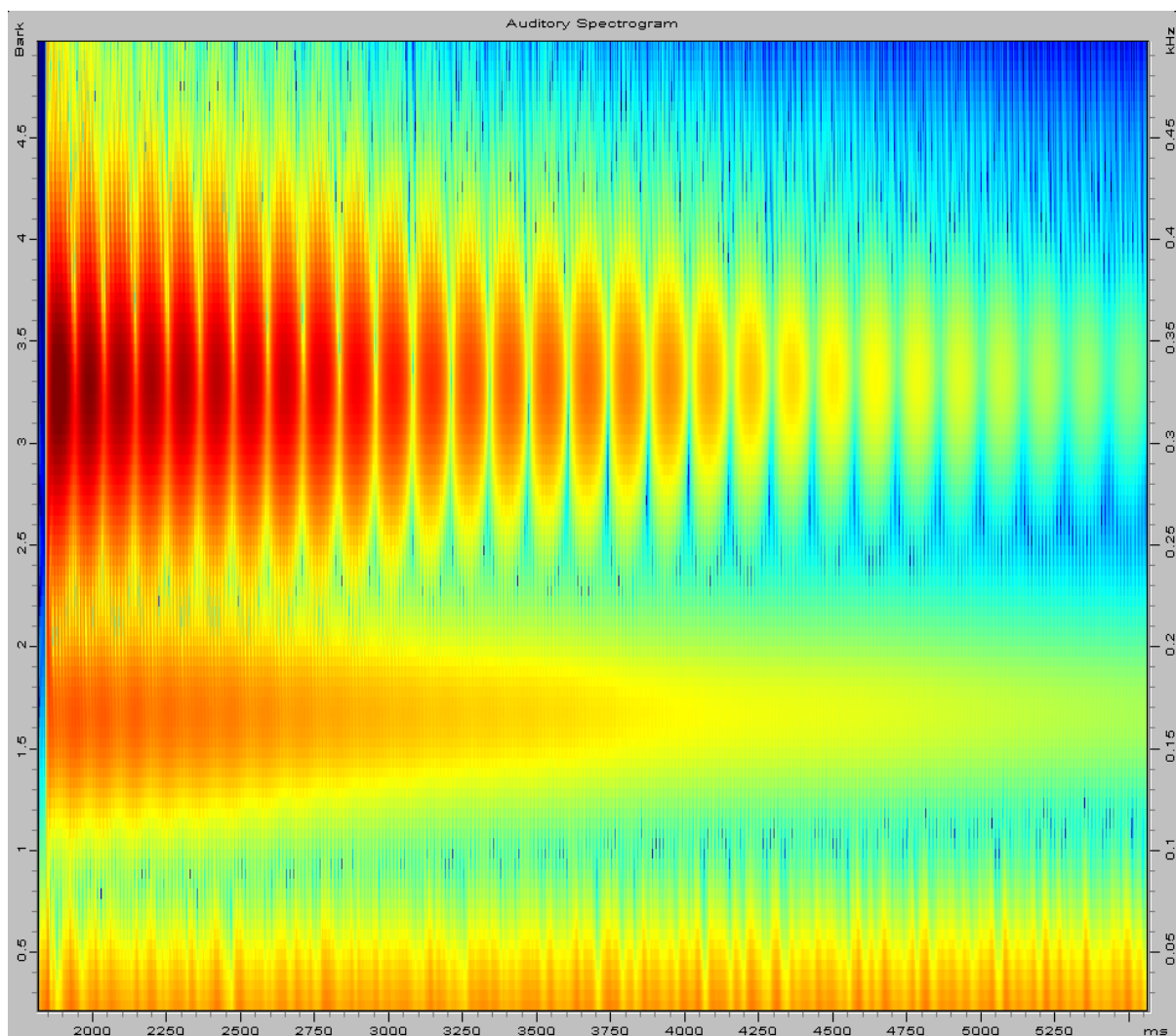


Abb. 3.1-1c: Gehörorientierte Analyse des gleichen Signals

3.2 Welche der Schwingungsformen einer Glocke ergeben hörbare Schallkomponenten?

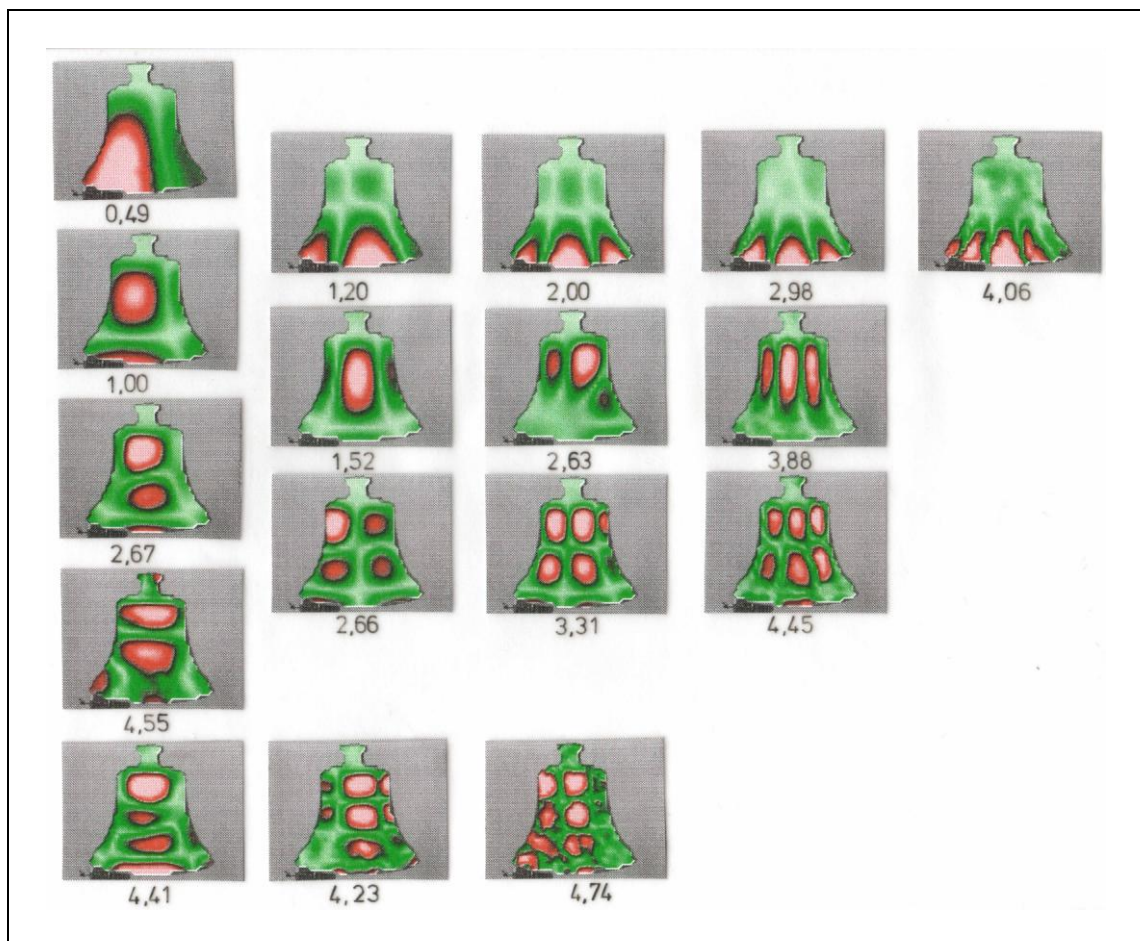


Abb. 3.2-1: Eigenschwingungsformen (Perner Glocke 34kg), die Zahlen unter den einzelnen Bildern geben das Frequenzverhältnis zur Prime an.

Eine Glocke schwingt nach dem Anschlag in verschiedenen Eigenschwingungsformen. Jede dieser Schwingungsformen tritt mit einer bestimmten Frequenz und Intensität auf. Durch Anregung der Luftmoleküle wird Schall erzeugt, der sich im Wesentlichen aus Frequenzkomponenten dieser Eigenschwingungen zusammensetzt.

Die mit einem *Scanning Vibrometer* herausgearbeiteten Schwingungsformen einer Glocke sind in Abbildung 3.2-1 dargestellt. Hörbar sind jedoch nicht alle diese Schwingungen, es werden einige maskiert oder sind vom menschlichen Gehör nicht getrennt auflösbar. Um diese zu finden, nimmt man das Schallsignal auf und analysiert es gehörgerecht. Vor allem in der *Frequency Tracks*¹⁹ Analyse (entspricht einer modifizierten Teiltonzeitmuster-Analyse nach Terhardt) sind die wesentlichen tonalen Komponenten sofort ersichtlich.

¹⁹ Die *Frequency Tracks* sind ein Teilbild der *Frequency Contours*. Sie bestehen aus allen Konturen, die eine bestimmte zeitliche Mindestlänge aufweisen und entsprechen (mit gewissen Vereinfachungen) dem, was vom menschlichen Gehör als tonal (und nicht nur geräuschhaft) wahrgenommen wird.

Mit der Option „*Nonlinear Masking*“ lässt sich außerdem direkt überprüfen, ob es zu Überdeckungen kommt.

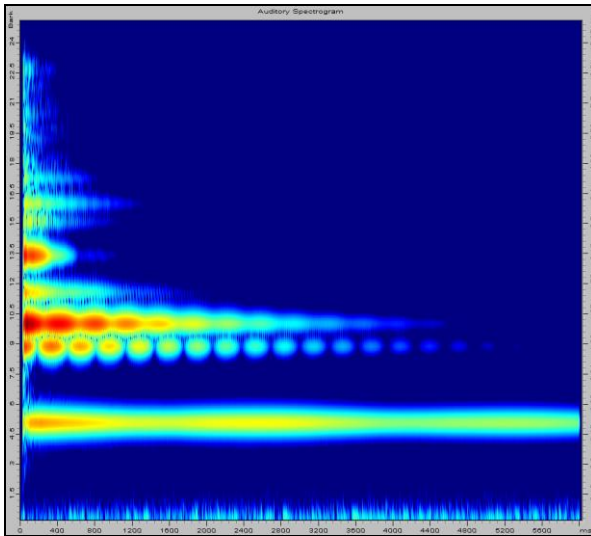


Abb. 3.2-2a: *Auditory Spectrogram*

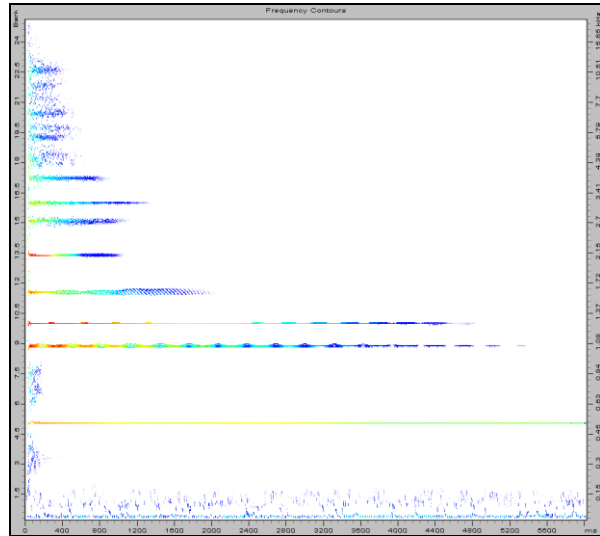


Abb. 3.2-2b: *Frequency Contours*

Deutlich sichtbar und somit auch hörbar sind

- Suboktave (0,51 kHz, Frequenzverhältnis 0,49 zur Prime),
- Prime (1,05 kHz, Verhältnis 1,00),
- Moll-Terze (1,26 kHz, Verhältnis 1,20),
- Quinte (1,61 kHz, Verhältnis 1,52) und
- Oberoktave (2,12 kHz, Verhältnis 2,00).

Weiterhin lässt sich in Abbildung 3.2-2b erkennen, dass für den charakteristischen Klang zum Zeitpunkt des Anschlags noch mehr als die 6 Grundtöne verantwortlich sind. Dies lässt sich mit der Resynthesefunktion leicht überprüfen – errechnet man eine Zeitfunktion nur aus diesen 6 Grundtönen, so ergibt sich beim Hörvergleich gerade beim Anschlag eine deutliche Abweichung vom Original.

Um die gleichen Informationen mit einem reinen FFT-Analysator zu gewinnen ist bei der Parameterwahl und der Analyse der Daten viel Erfahrung notwendig – VIPER erlaubt hingegen auch einem Laien einen solchen Einblick.

Gerade die Möglichkeit der Resynthese macht VIPER zu einem wertvollen Werkzeug, da die gewonnen Erkenntnisse schnell im Hörvergleich überprüfbar sind. Zwar bietet auch die inverse Fourier-Transformation allein bereits Möglichkeiten für Hörvergleiche, jedoch fehlen meist die komfortablen und schnellen Bearbeitungsmöglichkeiten.

4 Vergleich der VIPER-Ergebnisse mit anderen Analysen (Valenzuela)

An der Universität der Bundeswehr in München²⁰ sind bereits 1998/99 von M. Valenzuela Schallsignale von Musikinstrumenten „gehörorientiert“ untersucht worden²¹. Einige der Ergebnisse sollen mit Viper nachvollzogen bzw. verglichen werden.

Folgende Schallsignale wurden untersucht:

1. Gong III (bei Valenzuela²¹: „Gong 3“), Archivdatei: „*Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46\05 Kolberg-Gong Nr 3 (154 Hz).wav*“
2. Gong I (bei Valenzuela²¹: „Gong 1“), Archivdatei: „*Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46\05 Kolberg-Gong Nr 1 (164 Hz).wav*“
3. Neue Gitarrensaite E2, Gitarre Hoefner Nr. 1, Archivdatei „*Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46\28 Hoefner Nr 1, neue Saiten, E.wav*“

4.1 Gong III

Wie in Abbildung 4.1-1 zu erkennen ist, entspricht die VIPER-Darstellung „*Frequency Contours*“ dem von Valenzuela berechneten „Teiltonzeitmusters“ (TTZM) nach Terhardt²². Ein signifikanter Unterschied ist jedoch bei der Frequenz von ca. 325 Hz zu sehen. In der Darstellung von Valenzuela sind zwei eng beieinanderliegende Frequenzkomponenten sichtbar (ähnlich zu einer FFT-Analyse mit sehr langem Zeitfenster), bei der VIPER-Darstellung nur eine oszillierende Komponente, die außerdem in der Frequenz mit der Zeit leicht abnimmt. Dieser Teilton wurde extrahiert und isoliert von den anderen Schallkomponenten wieder resynthetisiert und daraufhin von mir selbst einem Hörtest unterzogen²³. Ich komme zu dem Entschluss, dass das Gehörte eher dem bei der VIPER-Darstellung sichtbaren Klangverlauf entspricht, aber zwei andere Testpersonen konnten sich meiner Meinung nicht anschließen. Eine endgültige Entscheidung kann an dieser Stelle nicht getroffen werden.

²⁰ Institut für Mechanik, Fakultät für Luft- und Raumfahrttechnik, Universität der Bundeswehr München, 85577 Neubiberg

²¹ Valenzuela, M.: Zur Rolle des Gehörs bei akustischen Untersuchungen an Musikinstrumenten

²² Valenzuela, M.: Zur Rolle des Gehörs..., S. 5

²³ Datei „\Beispiele\Kapitel 4 Vergleich\06 Kolberg-Gong Nr 3 (154 Hz)_resynt_01.wav“

4 Vergleich der VIPER-Ergebnisse mit anderen Analysen (Valenzuela)

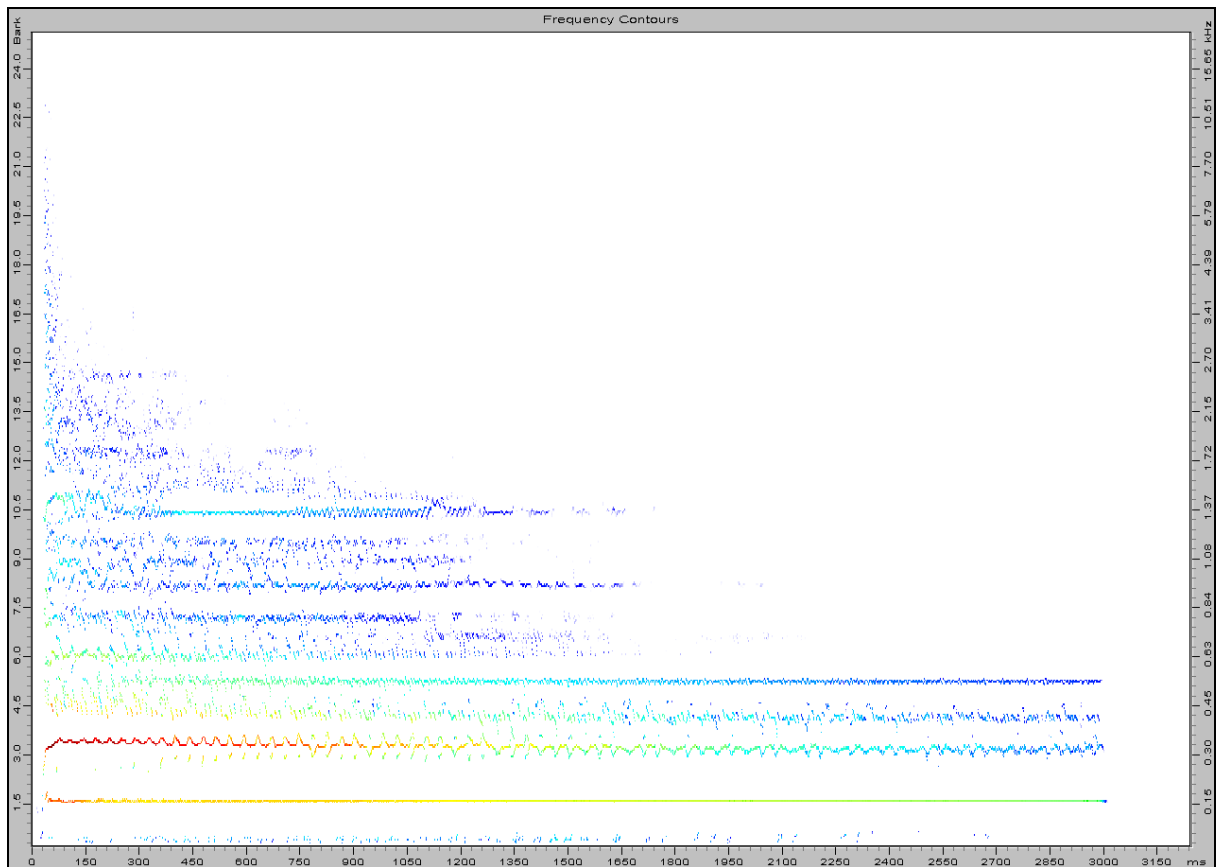


Abb. 4.1-1: Gong III VIPER *Frequency Contours*

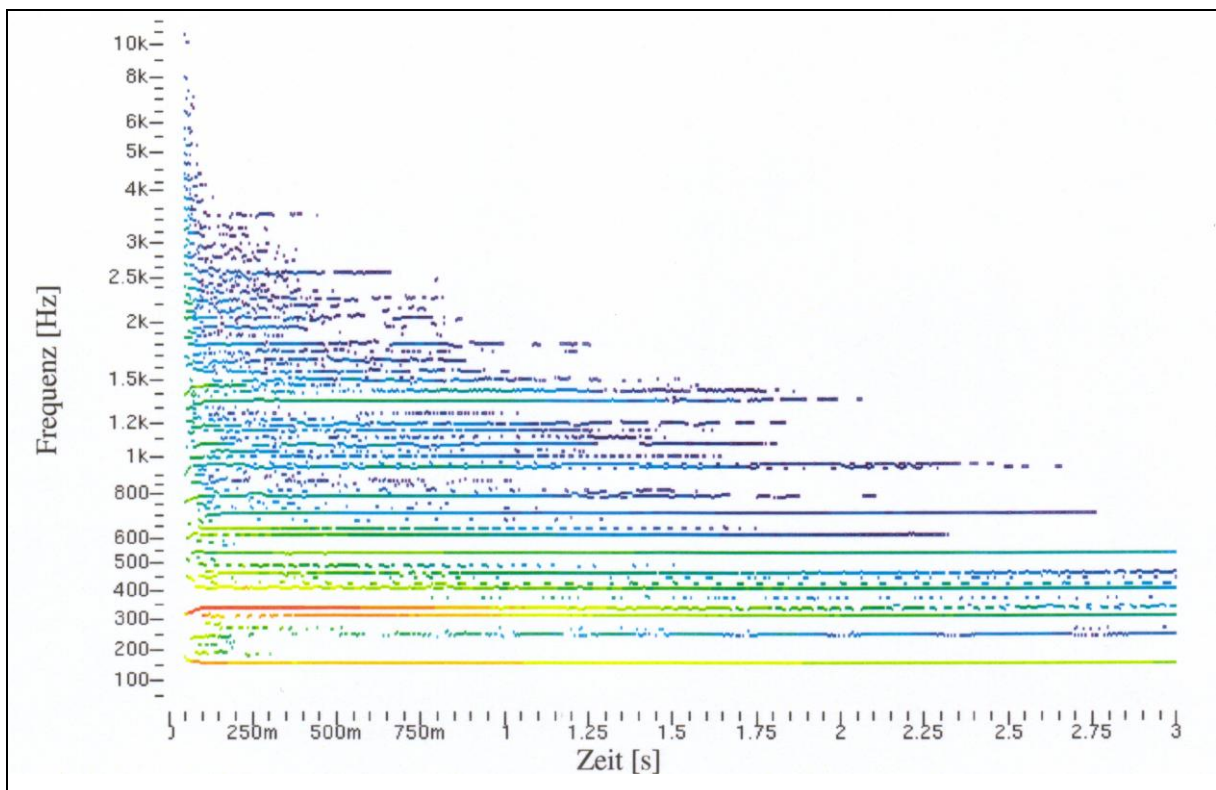


Abb. 4.1-2: TTZM von Valenzuela

	VIPER, Abbildung 4.1-1	TTZM von Valenzuela, Abbildung 4.1-2
Tonheitsabstand der Analysefrequenzen <i>Frequency interval</i>	0,05 Bark	0,03 Bark
Analysebandbreite <i>Analysis bandwidth</i>	0,5 Bark	0,1 Bark
Analysefrequenzbereich <i>Min. / max. frequency</i>	0 Hz bis 24 kHz	20 Hz bis 32 kHz
Auswertintervall <i>Time interval</i>	1ms	2,5ms
Filtertyp	4. Ordnung	1. oder 2. Ordnung
Abtastintervall <i>Sampling rate</i>	1/48000 Hz (DAT Standard)	1/32000 Hz

Tabelle 4-1: Verwendete Parameter

Die Unterschiede in den Parametern „Tonheitsabstand der Analysefrequenzen“ und „Analysebandbreite“ sind für den oben beschriebenen Effekt verantwortlich. In Kapitel 5 wird nochmals darauf eingegangen, welche Bedeutung die Parameter haben und welche dem menschlichen Gehör am ehesten entsprechen.

Verändert man die Parameter in VIPER gemäß denen von Valenzuela, so ergibt sich eine sehr ähnliche Darstellung (Abbildung 4.1-3).



Abb. 4.1-3: VIPER *Frequency Contours*, angepasste Parameter

4.2 *Gong I*

Die von Valenzuela errechneten Maskierungen²⁴ können mit Viper nicht in dem gleichen Maße nachvollzogen werden. Die VIPER-Option „*nonlinear masking*“ ist praktisch nicht dokumentiert und bietet keine Einstellmöglichkeiten²⁵.

²⁴ Valenzuela, M.: Zur Rolle des Gehörs..., S. 35

²⁵ Laut Aussage eines Mitarbeiters der Firma Cortex kommt bei der Berechnung nicht das von Valenzuela verwendete Verfahren der spektralen Maskierung zur Anwendung.

4 Vergleich der VIPER-Ergebnisse mit anderen Analysen (Valenzuela)

Abbildung 4.2-1 bietet einen direkten Vergleich der Maskierung.

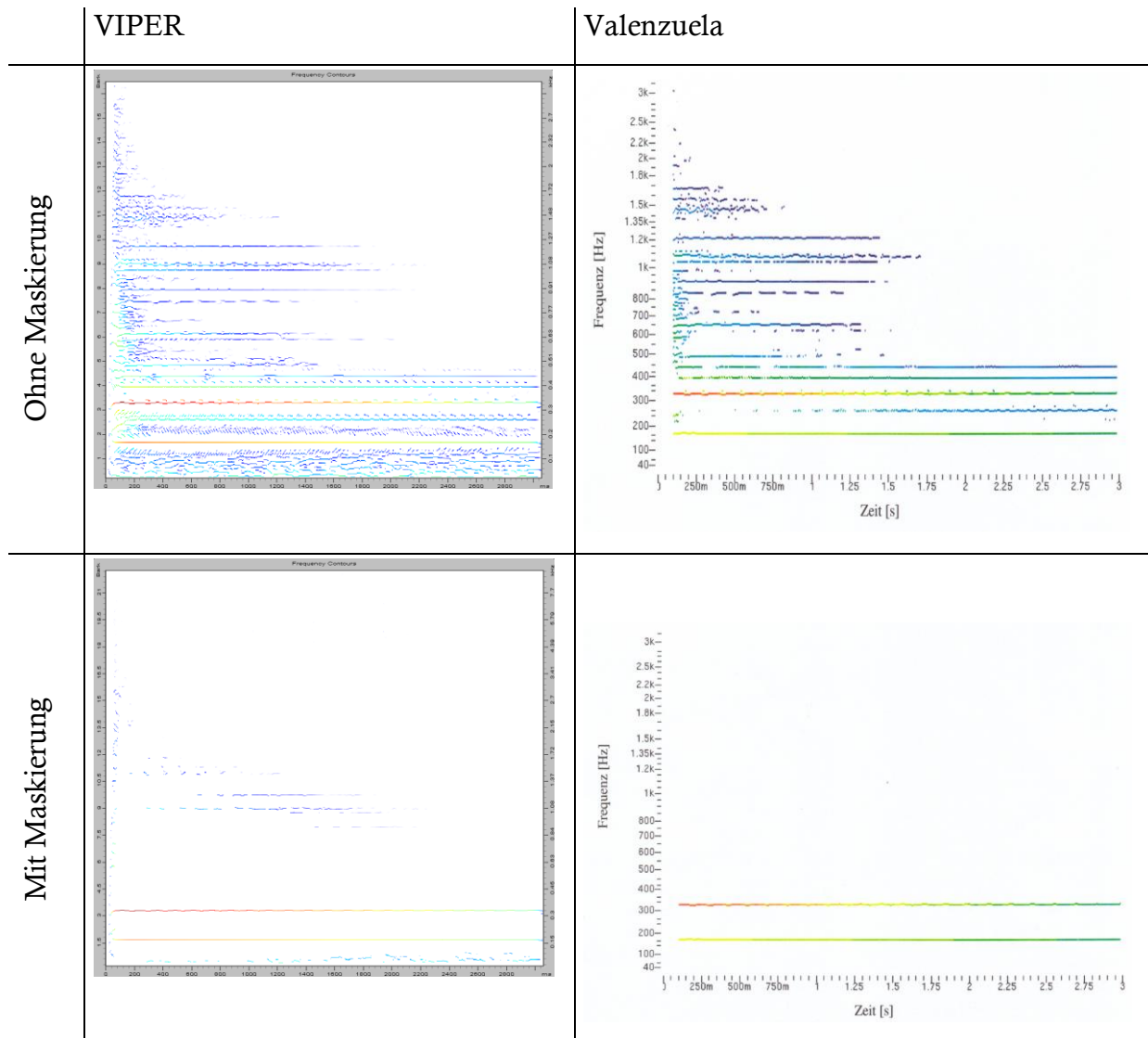


Abb. 4.2-1: Gong I

4.3 Gitarrensaite Haefner E2

Analog zu den beiden Gongs wurde der Klang einer Gitarrensaite mit und ohne Maskierung untersucht und mit denen von Valenzuela verglichen (Abbildung 4.3-1). Bis auf die Maskierung kann das Ergebnis von Valenzuela bei angepasster Parameterwahl reproduziert werden.

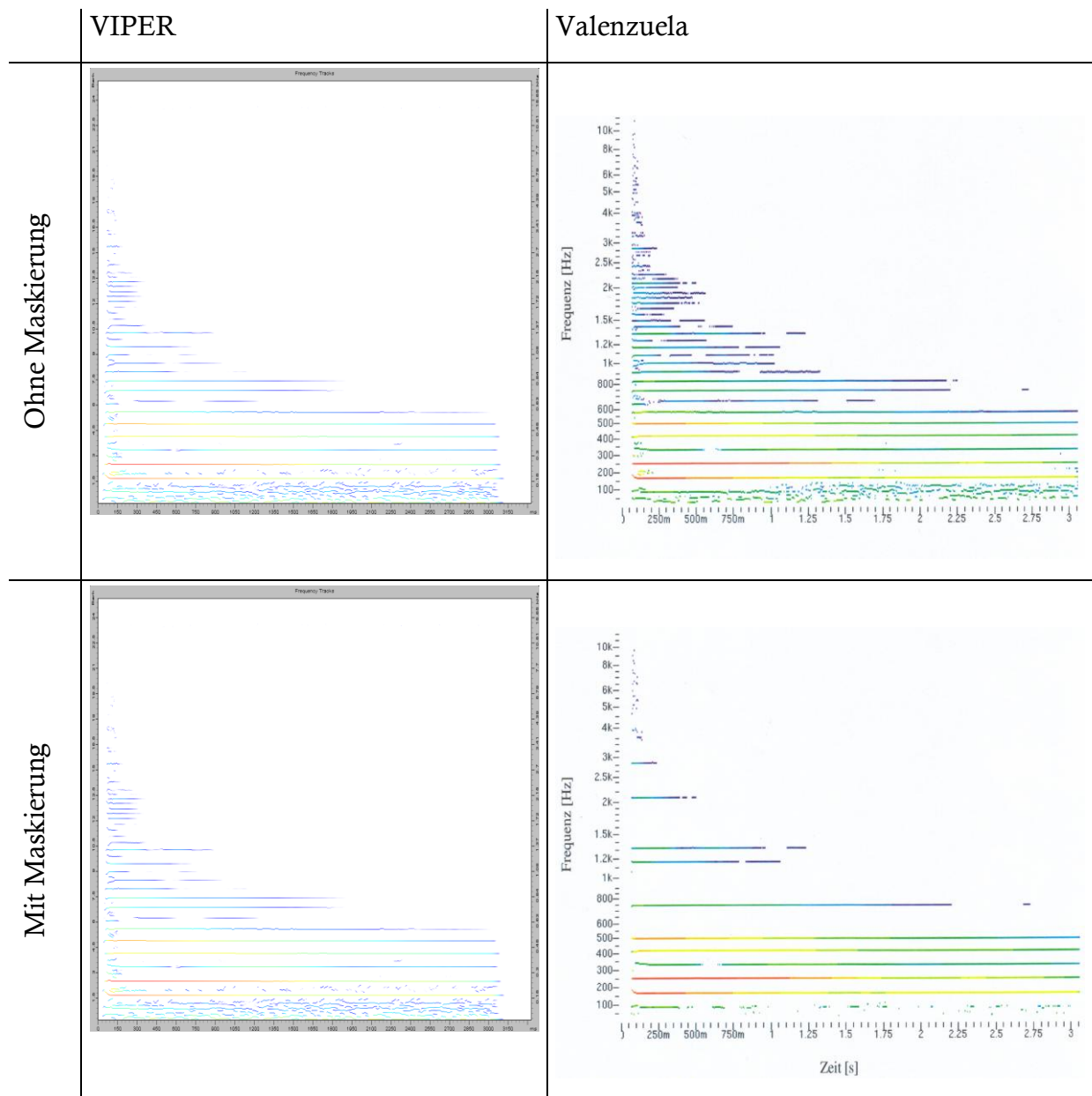


Abb. 4.2-1: Gitarrensaite Haefner E2

5 Zur Bedeutung der Parameter der gehörorientierten Schallanalyse

Bereits in Kapitel 4 hat sich gezeigt, dass das Problem der willkürlichen Wahl der Parameter für die gehörgerechte Schallanalyse nicht vollständig gelöst ist. Die Analyse von „Gong III“ mit unterschiedlichen Parametern führte zu signifikanten Unterschieden in der Darstellung. In der Fachliteratur werden verschiedene Ansätze für die gehörorientierte Repräsentation diskutiert, aber es gibt keine einheitliche Meinung über ein „einzig richtiges“ Verfahren.

Im Folgenden sollen die wichtigsten Parameter für das VIPER „Auditory Spectrogram“ und die „Frequency Contours“ einzeln erläutert werden (siehe Abbildung 5-1 und Tabelle 5-1).

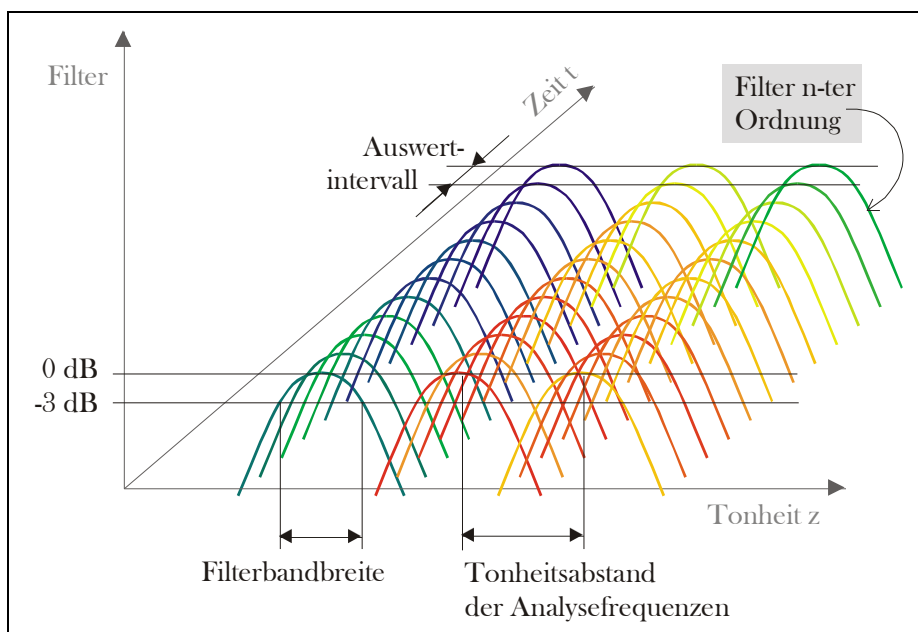


Abb. 5-1: Parameter bei der Analyse

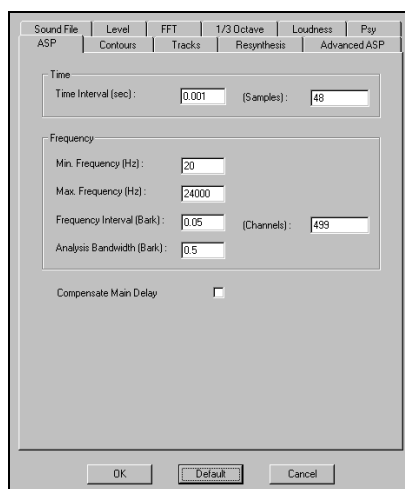


Abb. 5-2a

VIPER Optionen

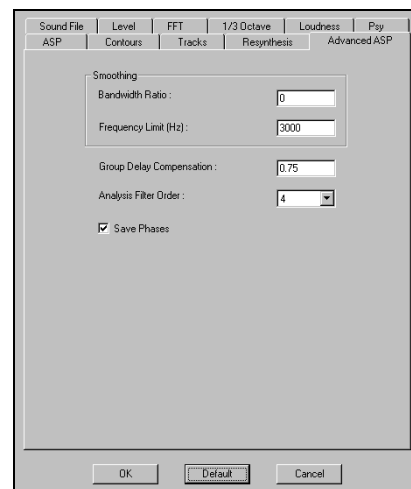


Abb. 5-2b

Parameter	Bedeutung	VIPER default
<i>Begin / end of analysis</i>	Hiermit kann ein Teil der Datei für die Analyse "ausgeschnitten" werden.	komplette Datei
<i>Time interval (sec) / samples</i>	Äquidistante Auswerteintervalle, betrifft also die Skalierung im Zeitbereich. Jedes Intervall entspricht einer vertikalen Linie der VIPER-Grafiken.	0,0125 Sekunden
<i>Min / max frequency</i>	Hiermit kann der relevante Frequenzbereich für die Analyse "ausgeschnitten" werden.	20 Hz bis 0,5 x Samplingfrequenz
<i>Frequency interval (Bark)</i>	Tonheitsabstand der Analysefrequenzen. Gibt den Frequenzabstand der einzelnen Filter an. Aus diesem Wert (zusammen mit <i>min / max frequency</i>) errechnen sich die Anzahl der „Channels“, also die Anzahl der horizontalen Linien der VIPER-Grafiken. Der Wert sollte auf den verwendeten Filtertyp sowie die verwendete Analysebandbreite abgestimmt sein.	0,05 Bark
Analysis bandwidth (Bark)	Die Analysebandbreite (3 dB-Bandbreite) des Filters.	0,5 Bark
Analysis filter type	Der für die Analyse verwendete Filtertyp.	4 x 1. Ordnung

Tabelle 5-1

Die Abbildung 5-3 zeigt zwei Analysen eines Schallsignals, das additiv aus zwei Sinussignalen zusammensetzt ist:

- Ein Sinussignal mit 300 Hz, 80 dB;
- ein zwei Sekunden später einsetzendes Sinussignal mit anfangs 390 Hz, dessen Frequenz sich auf 300 Hz verringert, auch bei konstanten 80 dB Schallpegel.

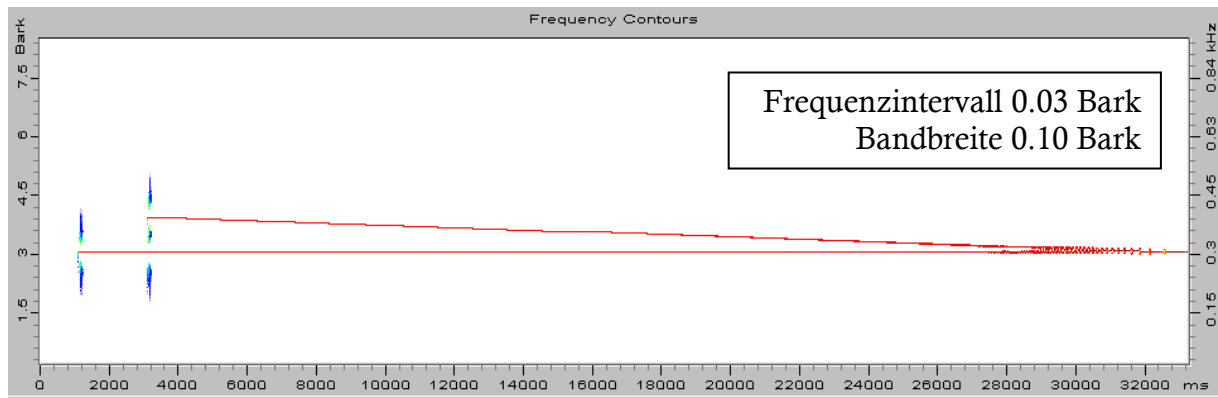


Abb. 5-3a: Analyse mit hoher Frequenzauflösung

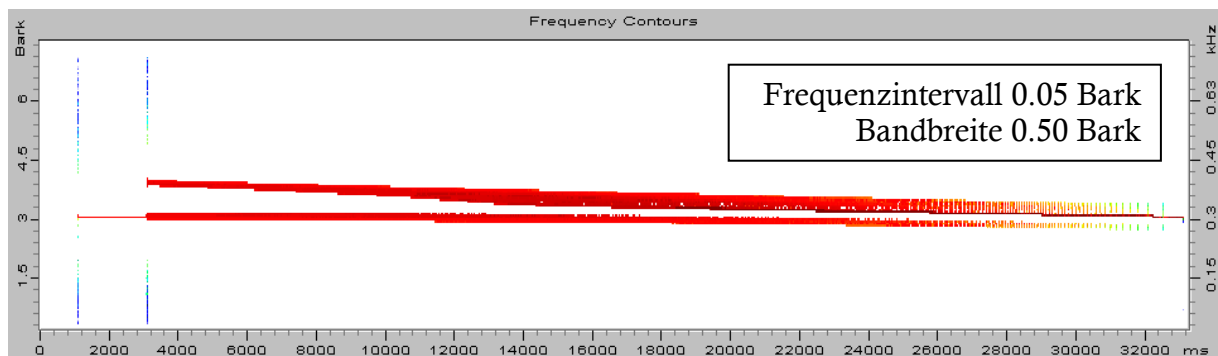


Abb. 5-3b: Analyse mit niedriger Frequenzauflösung

Abbildung 5-3a basiert auf einer deutlich höheren Frequenzauflösung. Aus diesem Grund erscheinen die Linien schärfer und sie verschmelzen später als in der zweiten Abbildung zu einer einzigen tonalen Komponente mit oszillierender Amplitude. Hört man sich das Signal an, so kann man zu Beginn die beiden Töne deutlich getrennt voneinander wahrnehmen. Die Töne werden zunehmend rauer, dann wird ein einziger Ton mit einer zunehmenden Schwankungsstärke wahrgenommen. Am Ende hört man deutlich die Schwebung eines einzigen Tons. Die Grenze der getrennten Wahrnehmbarkeit ist praktisch nicht genau feststellbar, vermutlich würden hier verschiedene Versuchspersonen unterschiedliche Entscheidungen treffen²⁶.

Resynthetisiert man aus den beiden VIPER-Ergebnissen wieder eine Zeitfunktion und vergleicht sie in einem Hörtest mit dem Original, so lassen sich im zweiten Fall (5-3b)

²⁶ Die .wav-Dateien des Schallsignals (original und resynthetisiert) finden sich auf der CD-Rom unter „Beispiele/Kapitel 5 Sweep/“

Unterschiede feststellen, das heißt die Resynthese gelingt nicht optimal. Die Auflösung war somit nicht ausreichend fein!

Die errechnete Zeitfunktion unterscheidet sich auch auf den ersten Blick deutlich vom Original, was zwar nicht im Zusammenhang mit den hörbaren Unterschieden stehen muss, aber in diesem Beispiel mit den zwei Sinustönen ist eine Abweichung bei den Amplituden zu erkennen (siehe Abbildungen 5-4a bis 5-4c).

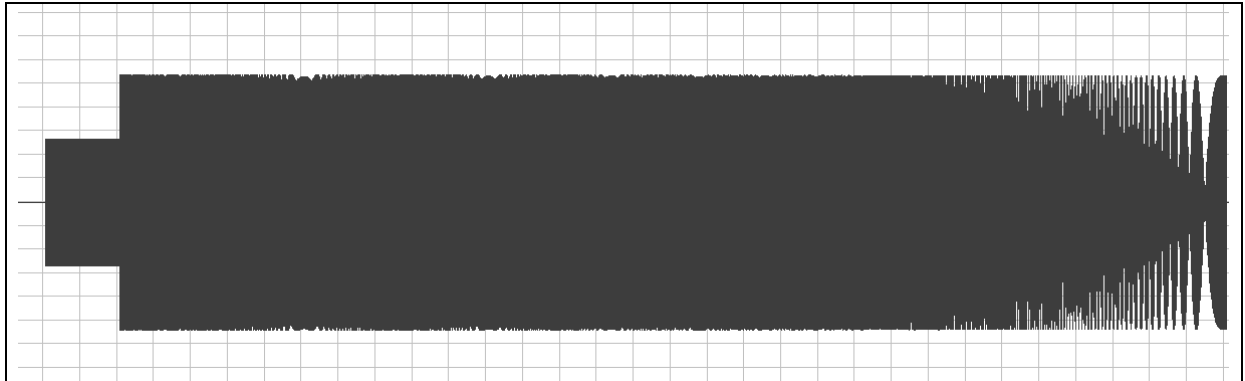


Abb. 5-4c
Original-Zeitfunktion

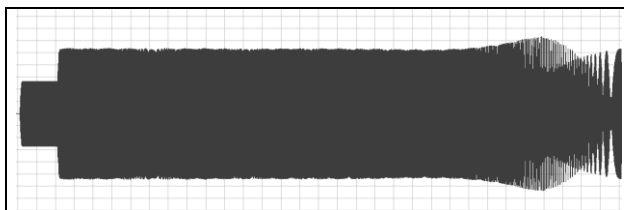


Abb. 5-4a
Resynthetisierte Zeitfunktion nach 5-3a

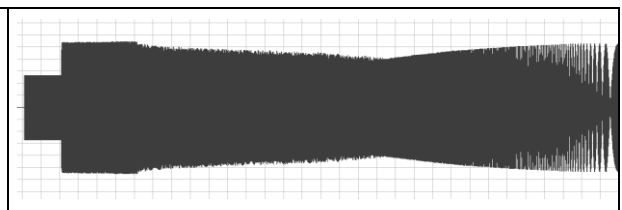


Abb. 5-4b
Resynthetisierte Zeitfunktion nach 5-3b

Es sollte an dieser Stelle auch erwähnt werden, dass die Darstellung 5-3b weitaus weniger Informationen enthält²⁷ als 5-3a. Die Dateigröße der Original-Zeitfunktion beträgt 1,5 MB (als 16 bit; 48 kHz *Wave*-Datei), die der *Frequency-Contours* im unteren Fall beträgt 2,5 MB und im oberen Fall 9 MB.

Obwohl dieses künstliche Beispiel vermutlich einen geringen Praxiswert hat, so verdeutlicht es doch die bei „Gong III“ festgestellte Problematik und zeigt verschiedene psychoakustische Phänomene. Es bleibt bei der menschlichen Wahrnehmung wohl immer eine gewisse Subjektivität bestehen, die eine genaue Bestimmung der „richtigen“ Parameter nur ansatzweise erlaubt.

²⁷ Unterschiedlich ist die Anzahl der „Bildpunkte“, bzw. die Auflösung der Grafiken.

6 Untersuchung verschiedener Gongs mit VIPER

Auf der CD-Rom zu dieser Arbeit (siehe Anhang) befinden sich in den Verzeichnissen „\Gongs\Gong I bis VII\“ und „\Gongs\Gong VIII bis XI\“ die VIPER-Dateien „Gong.wks“ mit den untersuchten Schallsignalen von 11 verschiedenen Gongs.

Zu jedem Gong sind folgende Daten vorhanden:

- *Level vs. Time* (unbewertet);
- *Auditory Spectrogram*²⁸;
- *Time-/Frequency Contours* (mit und ohne *nonlinear masking*)²⁸;
- *Frequency Tracks*²⁸;
- ein (anhand der *Frequency Contours*) resynthetisiertes Schallsignal;
- sowie die originale kalibrierte *.wav*-Datei des Gongs.

Im Anhang finden sich einige Ausdrücke der Grafiken.

Das Verzeichnis kann von der CD-Rom auf eine lokale Festplatte kopiert werden, danach können die Daten weitergehend bearbeitet werden²⁹.

Von den meisten der Gongs ist nur eine Aufnahme analysiert worden. Anhand von Gong X (Abb. 6-1) wird jedoch deutlich, dass viele verschiedene Aufnahmen des selben Instruments zu untersuchen sind, wenn man ein vollständiges Bild erhalten will. In Bezug auf Abstrahlverhalten, Anschlagsart und -lautstärke können große Unterschiede im Klangbild entstehen.

Im Fall von Gong X werden bei stärkeren Anschlägen hochfrequente Schwingungen angeregt als im Fall eines leiseren Anschlags.

²⁸ Je zweimal analysiert mit verschiedenen Parametern. Die Parameter können mit einem Klick mit der rechten Maustaste auf die Namen der Grafiken abgefragt werden.

²⁹ Aufgrund des Schreibschutzes auf der CD-ROM.

6 Untersuchung verschiedener Gongs mit VIPER

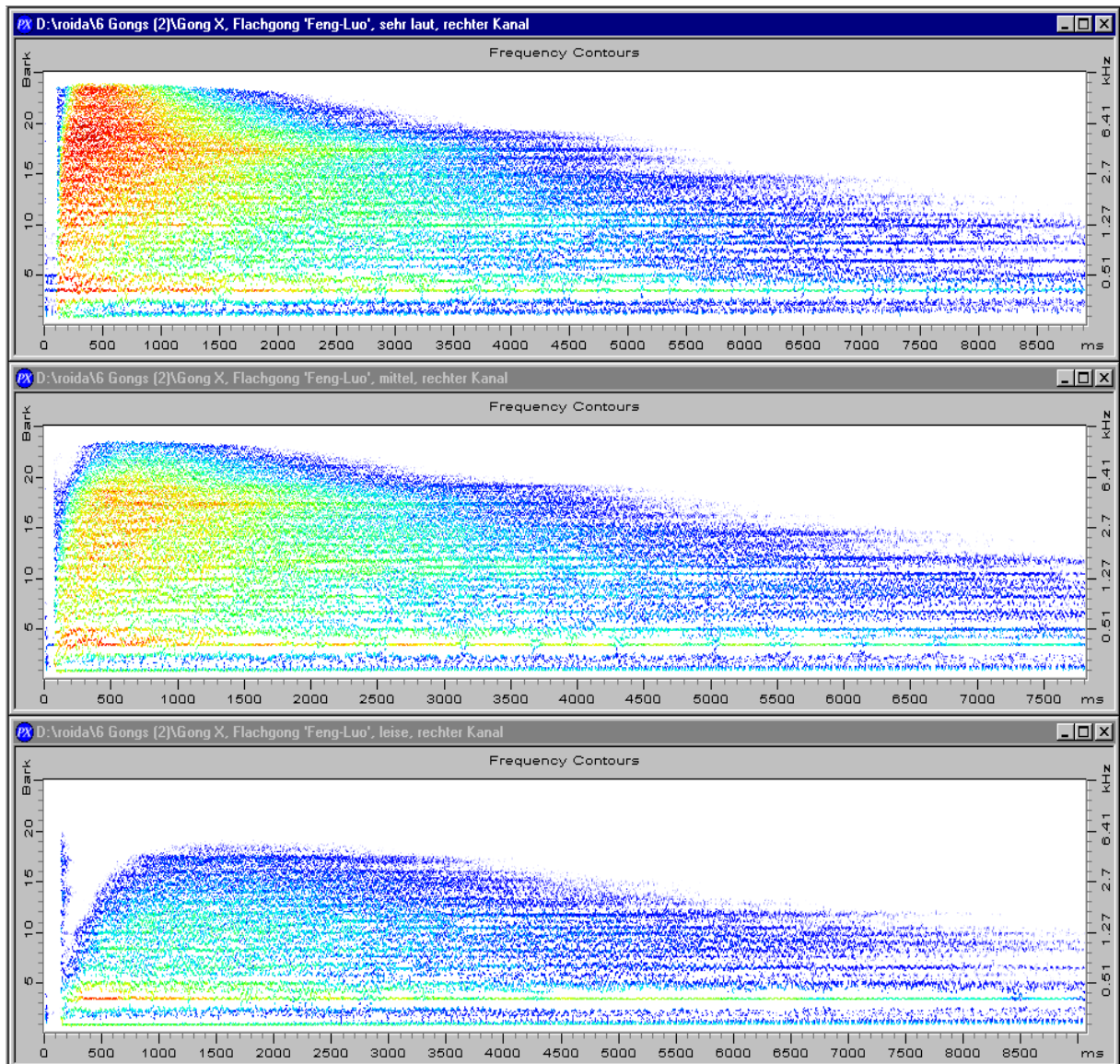


Abb. 6-1: Gong X; laut, mittel und leise angeschlagen.

7 Resümee und Ausblick

Während meiner Arbeit mit VIPER hat sich der Nutzen der gehörorientierten Schallanalyse deutlich gezeigt und mit der Software haben sich dazu viele Möglichkeiten herauskristallisiert. Auch wenn es in meiner Arbeit nicht darum ging, konkrete Ergebnisse zu bestimmten Schallsignalen zu erhalten, so konnten bei experimentellen Arbeiten viele Informationen gewonnen werden.

Die (Re-)synthesefunktionen bieten eine sehr gute Kontrollmöglichkeit für die mit VIPER errechneten Darstellungen. Wie gezeigt wurde bieten sich auch Möglichkeiten zur gezielten Modifikation von Klängen.

Ein komplexes Schallsignal – vor allem Musik – lässt sich allerdings nur mühsam analysieren. Die in dieser Arbeit behandelten Schallsignale bestanden aus diskreten Elementen und waren somit noch gut überschaubar. Eine Folge von mehreren Klängen oder Sprache macht die Analyse weitaus schwieriger. Das Problem liegt hierbei nicht in der Software, sondern betrifft das verwendete Prinzip der visuellen Darstellung an sich. Um aus solch komplexen Grafiken tiefgreifende Informationen zu erhalten ist sehr viel Erfahrung notwendig.

Die Software weist noch viele Windows-typische Probleme auf, bei denen es sich entweder um „Kinderkrankheiten“ oder einen schlechten Programmierstil der Benutzeroberfläche handelt. Die Software stürzte auch in der aktuellsten Programmversion häufig ab (zum Beispiel beim Scrollen oder Zoomen) oder führte Operationen nicht fehlerfrei aus. Beispielsweise trat es mehrfach auf, dass eine verschobene Schallkomponente nicht mehr mit resynthetisiert wurde. Oder aber es kam zu fehlerhaften Bildschirmdarstellungen. Oft steht man als Benutzer auch einfach vor kleinen Ungereimtheiten bei der Bedienung. Der Hersteller scheint hier die Philosophie „Funktionsumfang vor Stabilität“ zu vertreten.

Die Dokumentation zu VIPER (in englischer Sprache) deckt fast nur die Bedienung der Software ab – über Grenzen und Einschränkungen der Software oder über technische Details erhält man keine ausführlichen Informationen. Eine Ausnahme stellen hier die Formeln der Bark-Skala und der Berechnung des *Auditory Spectograms* dar, die abgedruckt sind. Zwar sind in der zweiseitigen Einleitung viele Quellenangaben zu den Verfahren angegeben, aber hieraus lässt sich nur ein ungenügender Rückschluss auf die Funktionsweise der Software ziehen. Einige Optionen sind unzureichend oder gar nicht dokumentiert. Zu der aktuellen Version der Software war keine komplette Dokumentation verfügbar, sondern nur die zu einer alten Version sowie ein Heft über „new features for version 2.11“. In manchen Fällen hilft also nur der schwer erreichbare aber sehr kompetente telefonische Support.

Durch die Lücken in der Dokumentation entsteht meiner Meinung nach neben der Bedienung noch ein weiteres Problem: es fehlt teilweise die Rechtfertigungsgrundlage für die mit VIPER erarbeiteten Ergebnisse.

Dem in der Produktwerbung angepriesenen Verwendungszweck, Audiokompressionsverfahren zu beurteilen, wird die Software meiner Meinung nach nur teilweise gerecht. Wenn bereits (wie in Kapitel 5 gezeigt) bei der Analyse und Resynthese mit VIPER Unterschiede zum Original hörbar werden, so lassen sich moderne, auf hohe Qualität ausgelegte Algorithmen wohl nur eingeschränkt bewerten (beispielsweise nach MP3-Standard umgesetzte Methoden mit hohen Bitraten). Anders sieht es hierbei sicherlich bei hochkomprimierten oder fehlerhaften Sprach- oder Nachrichtenübermittlungen aus – VIPER kann hier (vor allem mit seiner „*Speech Quality Analysis*“) mit Sicherheit bei einer objektiven Bewertung oder Optimierung helfen.

Für die Zukunft steht mit VIPER ein Tool zur Verfügung, das sich als wertvolles Hilfsmittel bei der Modellierung und Simulation von Musikinstrumenten und anderen Schallquellen zeigen könnte. Beispielsweise könnte man die Qualität von FEM-Rechnungen in Bezug auf das menschliche Gehör beurteilen und die Frage beantworten, wie gut eine simulierte mit einer echten Schallquelle übereinstimmt.

Vor allem die Objektivität der Beurteilung spielt hier eine wesentliche Rolle. Ohne aufwendige Hörversuche mit einer großen Anzahl von Testpersonen können reproduzierbare Aussagen gemacht werden.

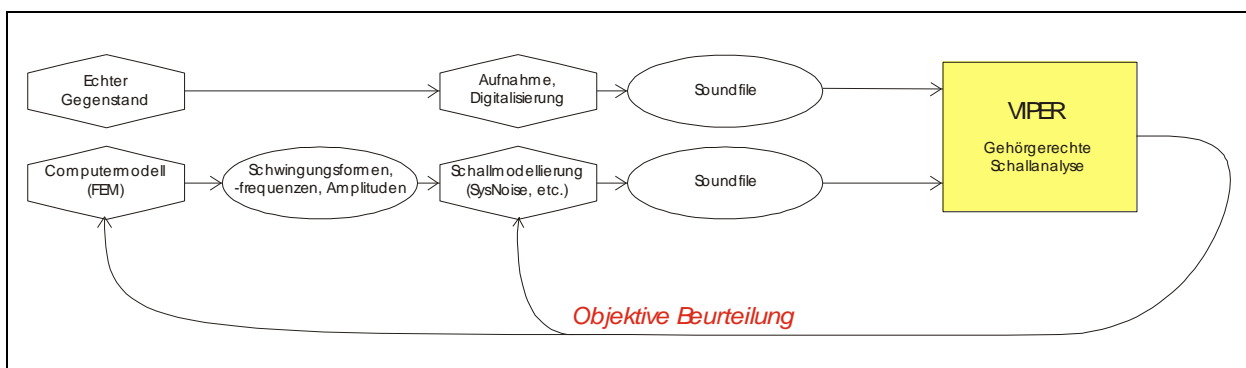


Abb. 7-1: VIPER als neues Bindeglied bei der Modellierung von Schallquellen

Eine umfangreiche Einarbeitung, wie sie zum Beispiel bei FEM-Programmen notwendig ist, ist nicht erforderlich um zu ersten brauchbaren Ergebnissen zu kommen.

Die Fähigkeit, „das Hörbare sichtbar zu machen“, hat VIPER gezeigt. Anhand meiner Kritik sollte nicht der Eindruck entstehen, man könne nicht gut mit der Software arbeiten – hat man einmal die oben geschilderten Probleme überwunden, so lassen sich viele Aufgaben sehr komfortabel und schnell erledigen.

ANHANG

I Die BARK Frequenzskala

Die Frequenzauflösung des menschlichen Gehörs ist bei tiefen Frequenzen feiner als bei hohen Frequenzen. Die Angabe von Frequenzen in Hertz ist aus diesem Grund schlecht geeignet. Deshalb wurde die nichtlineare Bark-Frequenzskala eingeführt: im unteren Bereich verläuft sie annähernd linear, im oberen annähernd logarithmisch (siehe Abb. I-1). Benannt wurde sie nach *Heinrich Barkhausen*. Im Grunde genommen handelt es sich um eine Mittelung vieler Hörversuche, und Zwicker und Fastl³⁰ beschreiben eine Näherungsformel:

$$z / \text{Bark} = 13 \cdot \arctan\left(\frac{0,76 \cdot f}{1000\text{Hz}}\right) + 3,5 \cdot \arctan\left(\frac{f}{7500\text{Hz}}\right)^2 \quad (\text{I-1})$$

$$1\text{Bark} = 100\text{mel} \quad (\text{I-2})$$

Mit:

z : Tonheit in *Bark*

f : Frequenz in *Hz*

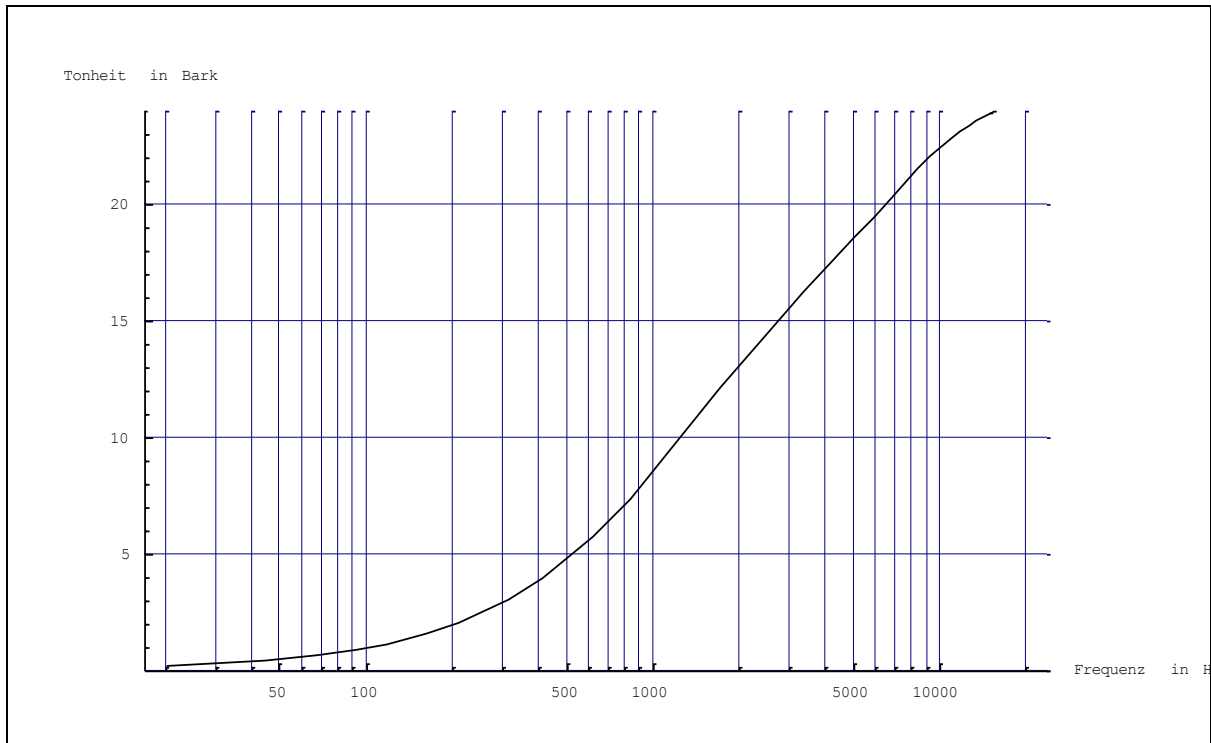


Abb. I-1: Zusammenhang zwischen Frequenz und Tonheit (Bark-Skala) nach Zwicker und Fastl³⁰

³⁰ Zwicker, E.; Fastl, H.: Psychoacoustics, S. 164

In VIPER wird eine andere Formel zur Berechnung verwendet, die von der hier dargestellten um bis zu 0,2 Bark abweicht³¹ (siehe Abbildung I-2). Es handelt sich um eine aus drei Teilfunktionen zusammengesetzte Funktion, die zwar stetig, aber nicht stetig differenzierbar ist.

Aus der Abweichung von 0,2 Bark resultiert beispielsweise bei einem Auswertintervall von 0,03 Bark ein Unterschied von 7 *frequency channels* in den VIPER-Grafiken.

Meiner Meinung nach ergibt die von VIPER verwendete Approximation keinen Sinn, da der an den Funktionsübergängen „eckige“ Verlauf die Gehöreigenschaften nicht besser nachbilden kann als die glatte Funktion von Zwicker und Fastl, auch wenn sie die tabellierten Werte ihrer Hörversuche besser trifft.

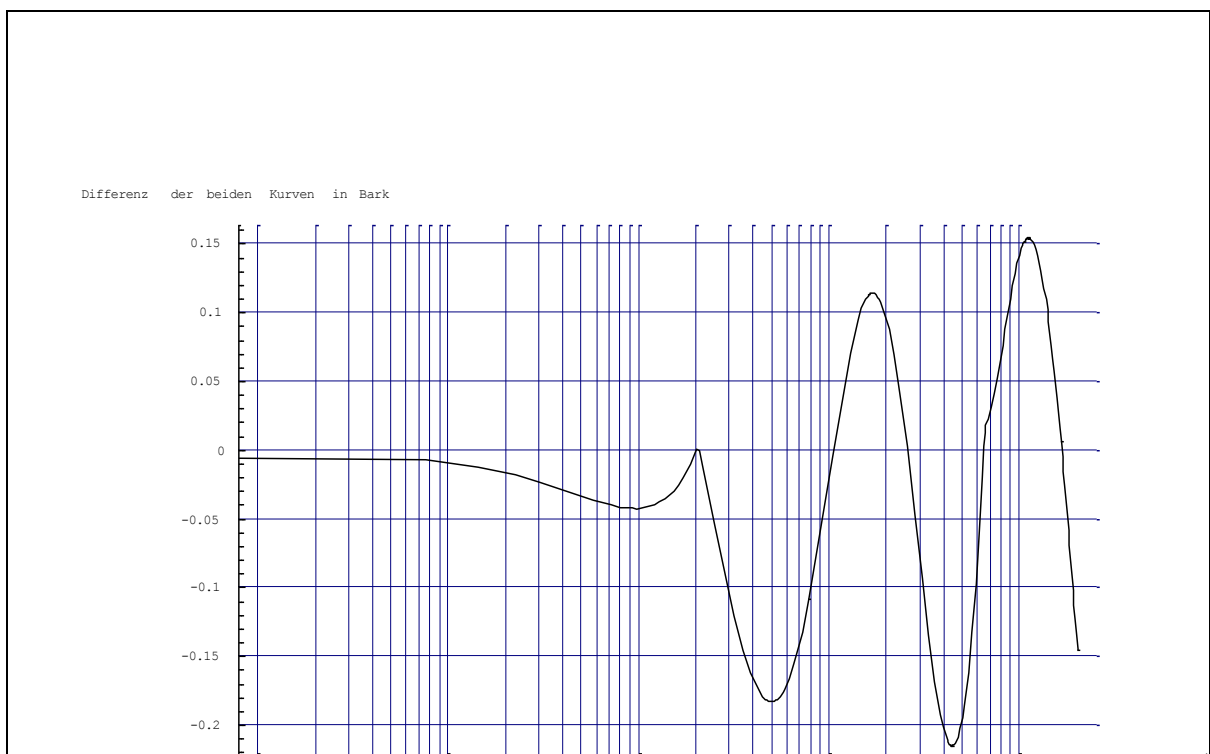


Abb. I-2: Differenz der beiden Funktionsverläufe

³¹ VIPER Manual, Version 1.11, S.42

II Gongs mit VIPER analysiert

Im Folgenden sind Auszüge der mit VIPER gewonnenen Daten über verschiedene Gongs abgedruckt (siehe auch Kapitel 6).

Zu den Pegeln von Gong I bis Gong VII siehe Abbildung II-1.

Für die Gongs VIII bis XI lag kein Bezugspegel vor, sie wurden aber alle bei dem gleichen konstanten Pegel überspielt.

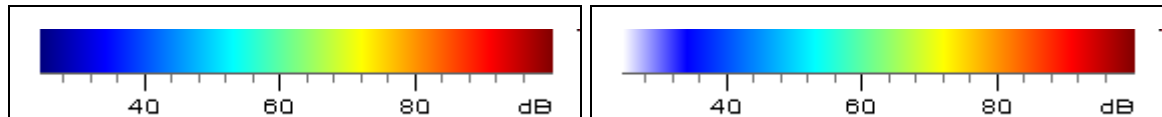


Abb. II-1: Farbskala für den Schallpegel der Gongs I bis VII, 25 – 100 dB
links für die *Auditory Spectograms*, rechts für die *Frequency Contours*

II.1 Gong I, kleiner Kolberg Gong

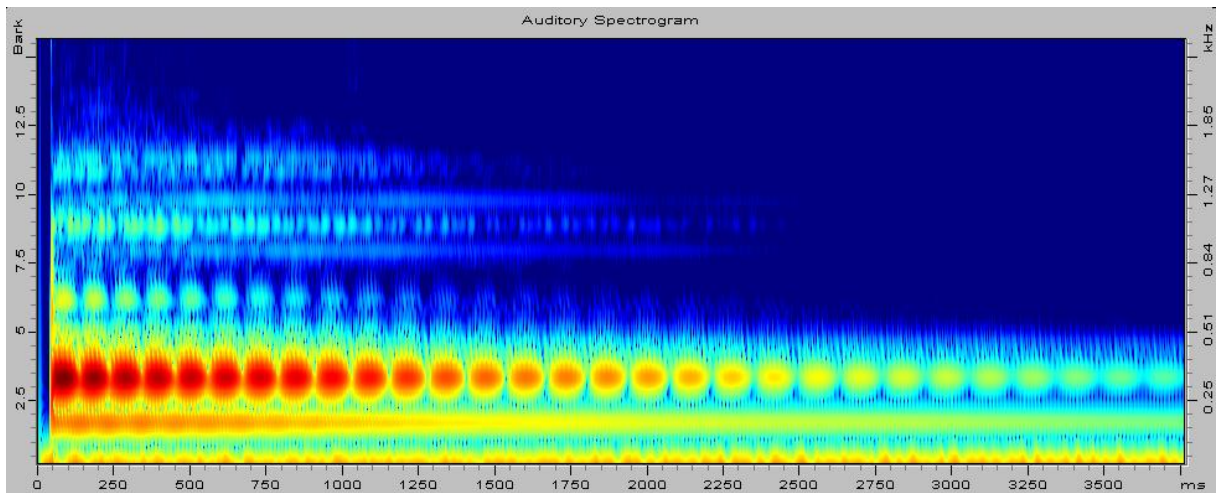


Abb. II.1-1 Auditory spectrogram

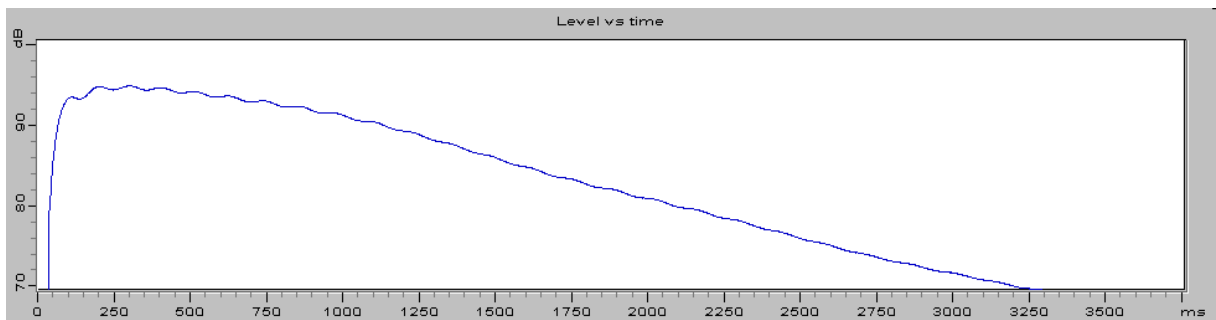


Abb. II.1-2 Level vs. time

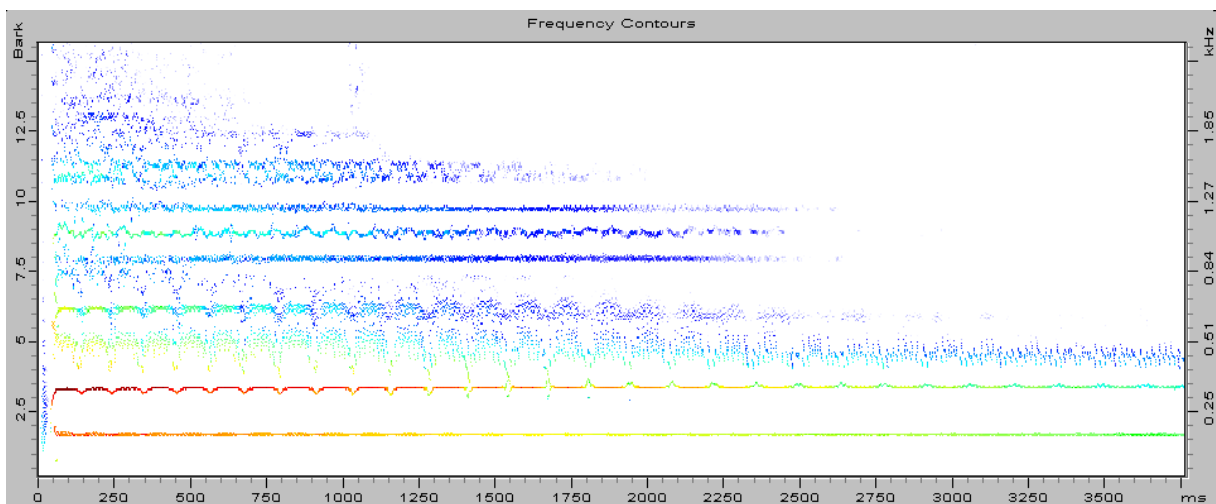


Abb. II.1-3 Frequency contours with nonlinear masking

II.2 Gong II, Kolberg-Gong 'Ewa'

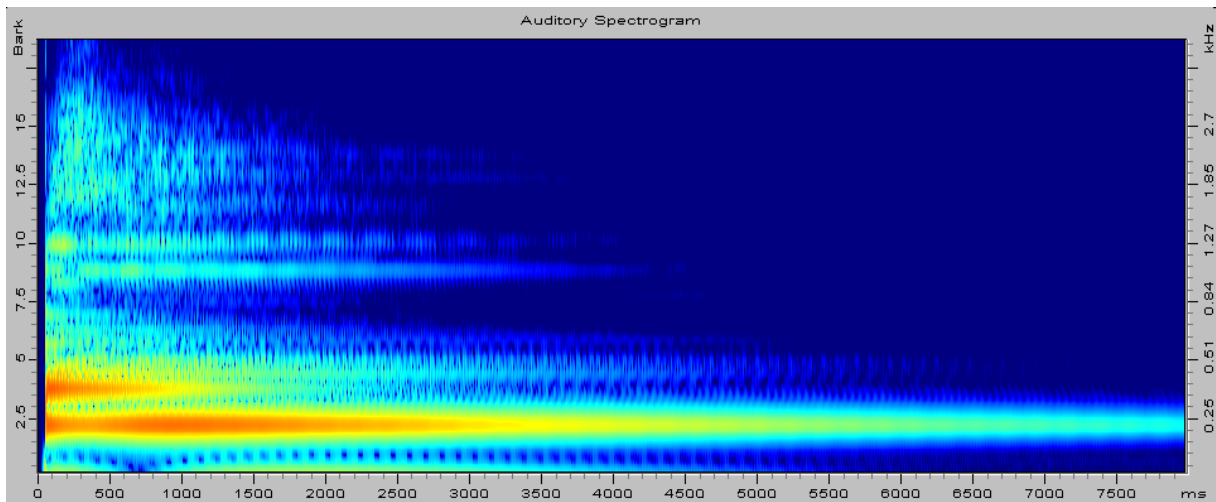


Abb. II.2-1 Auditory spectrogram

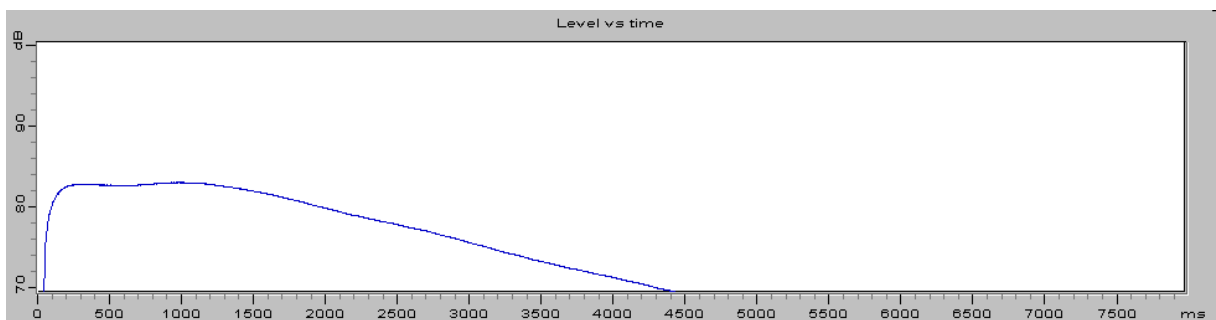


Abb. II.2-2 Level vs. time

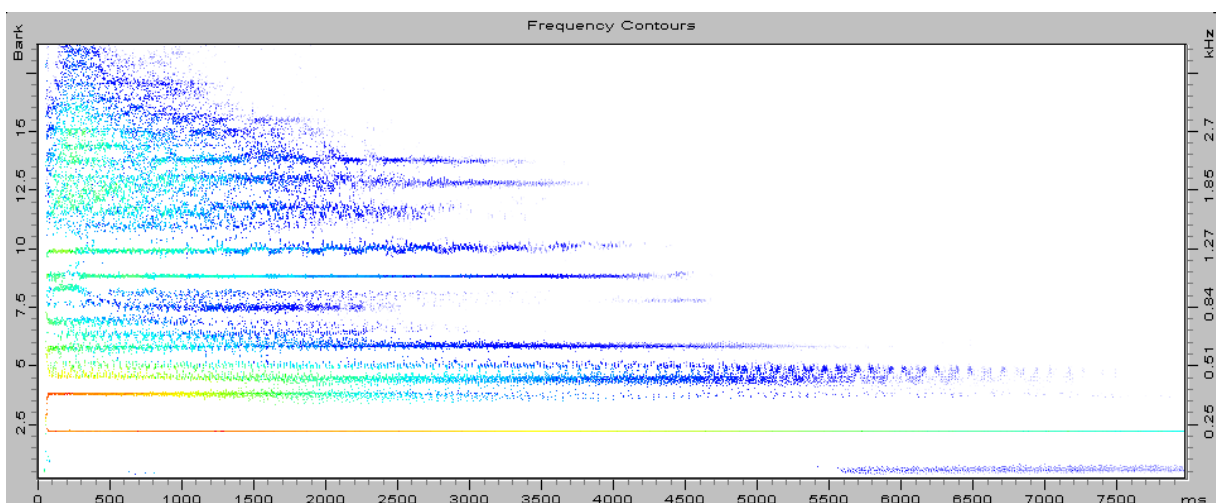


Abb. II.2-3 Frequency contours with nonlinear masking

II.3 Gong III, mittlerer Kolberg-Gong

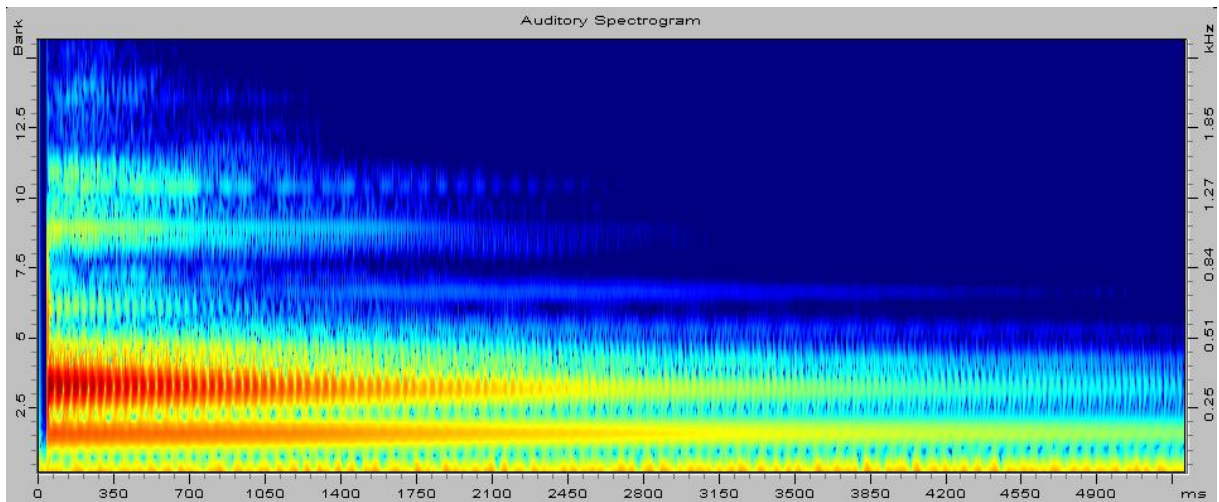


Abb. II.3-1 Auditory spectrogram

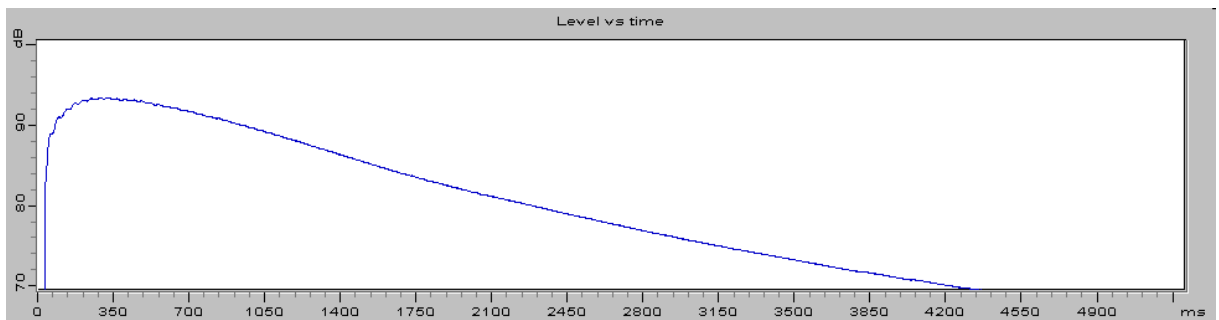


Abb. II.3-2 Level vs. time

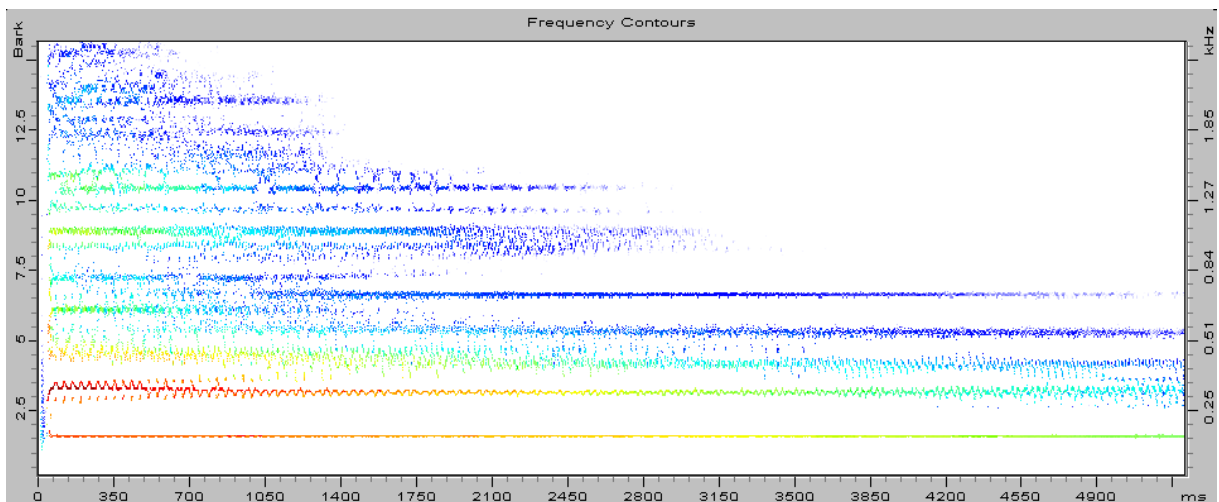


Abb. II.3-3 Frequency contours with nonlinear masking

II.4 Gong IV, mittlerer Kolberg-Gong 'Tobi'

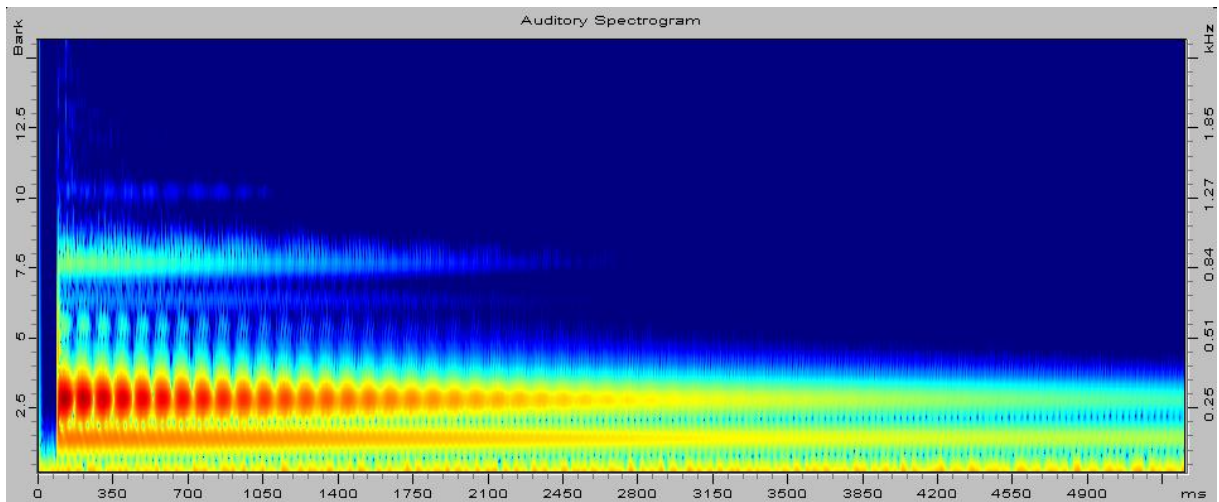


Abb. II.4-1 Auditory spectrogram

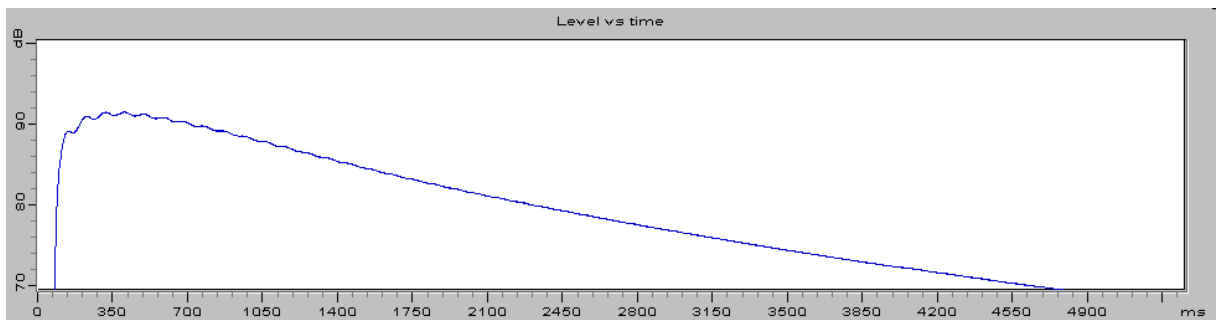


Abb. II.4-2 Level vs. time

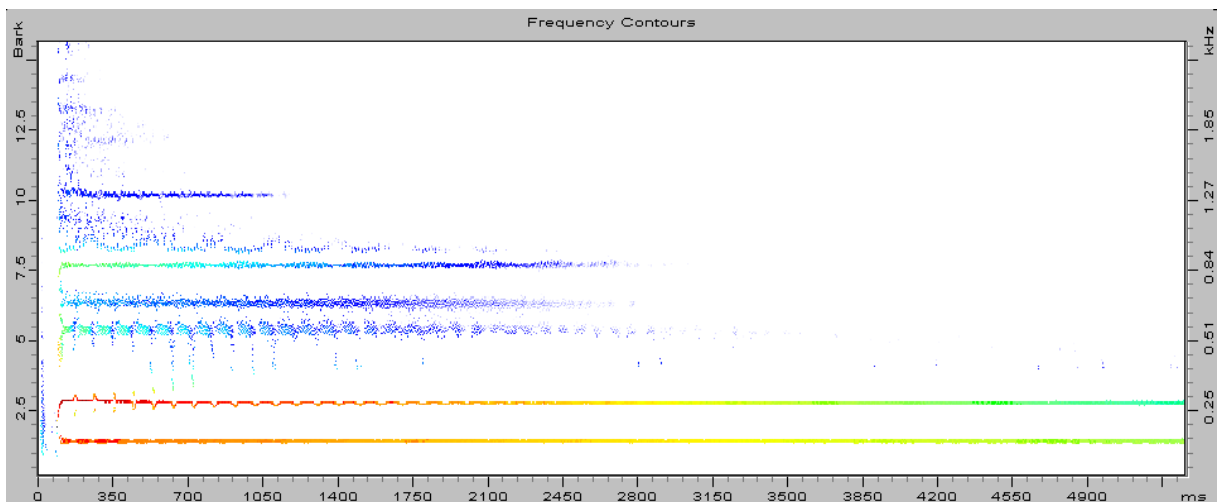


Abb. II.4-3 Frequency contours with nonlinear masking

II.5 Gong V, großer Kolberg-Gong 53 cm

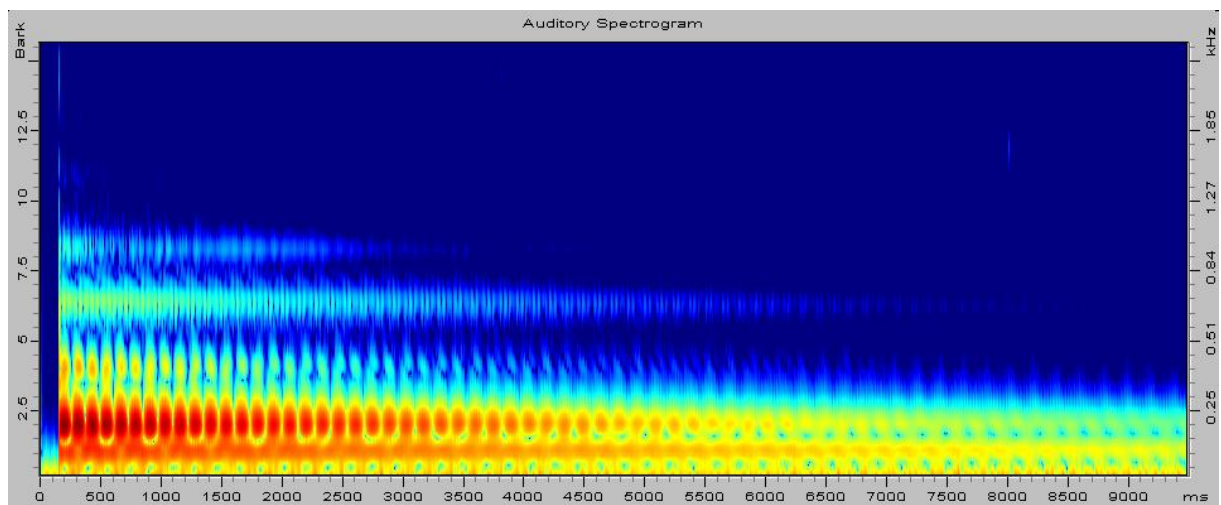


Abb. II.5-1 Auditory spectrogram

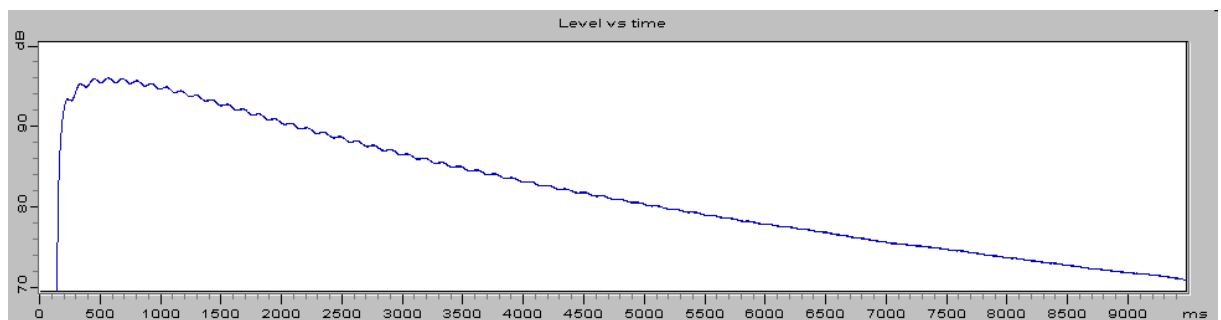


Abb. II.5-2 Level vs. time

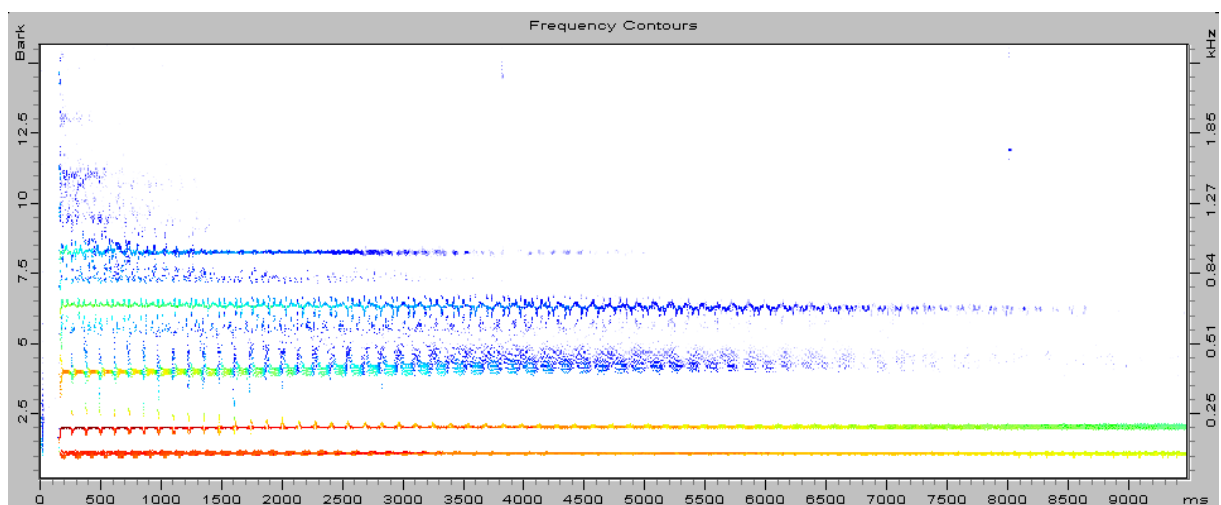


Abb. II.5-3 Frequency contours with nonlinear masking

II.6 Gong VI, großer Kolberg-Gong GIS 56

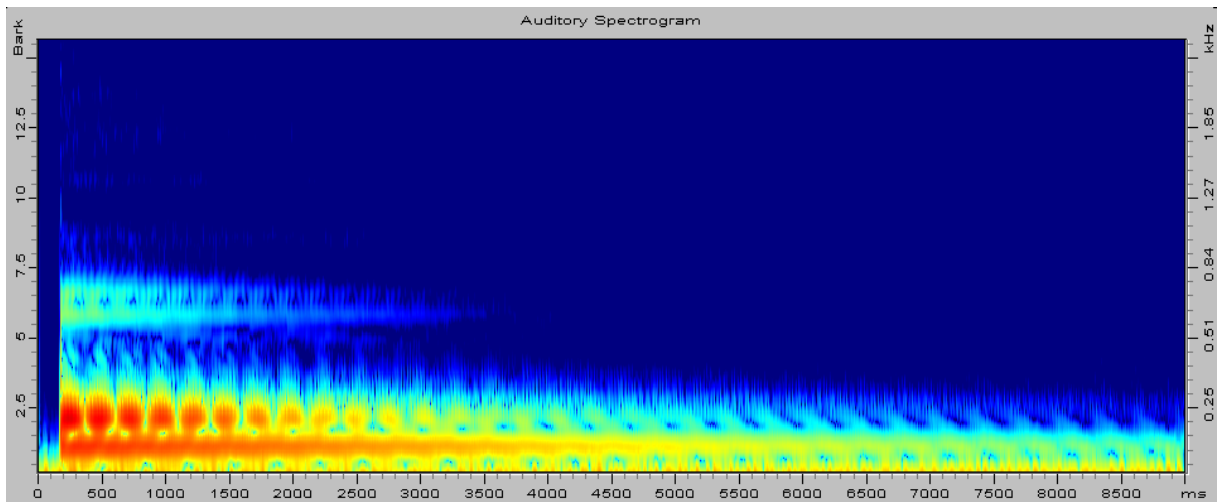


Abb. II.6-1 Auditory spectrogram

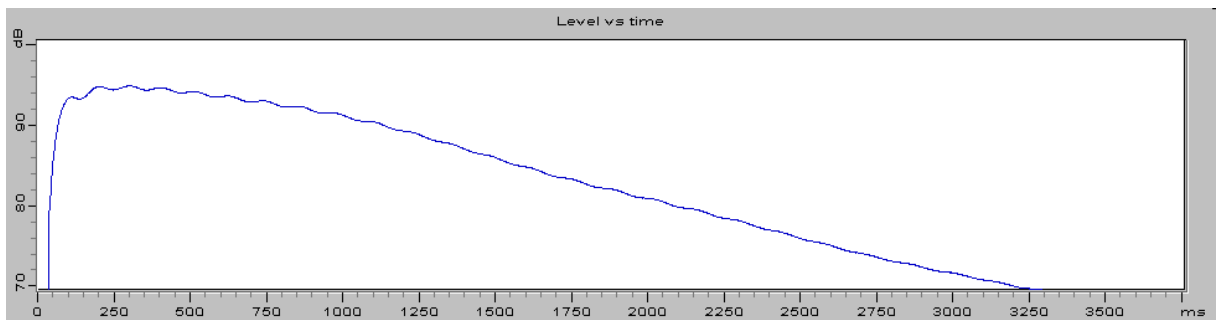


Abb. II.6-2 Level vs. time

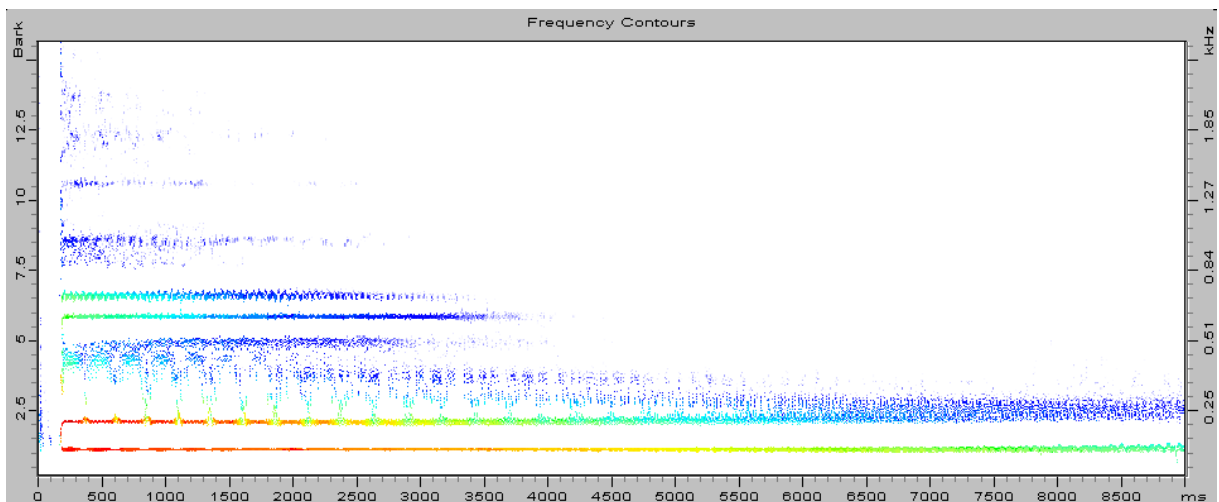


Abb. II.6-3 Frequency contours with nonlinear masking

II.7 Gong VII, großer Kolberg-Gong 'Pappelstrasse'

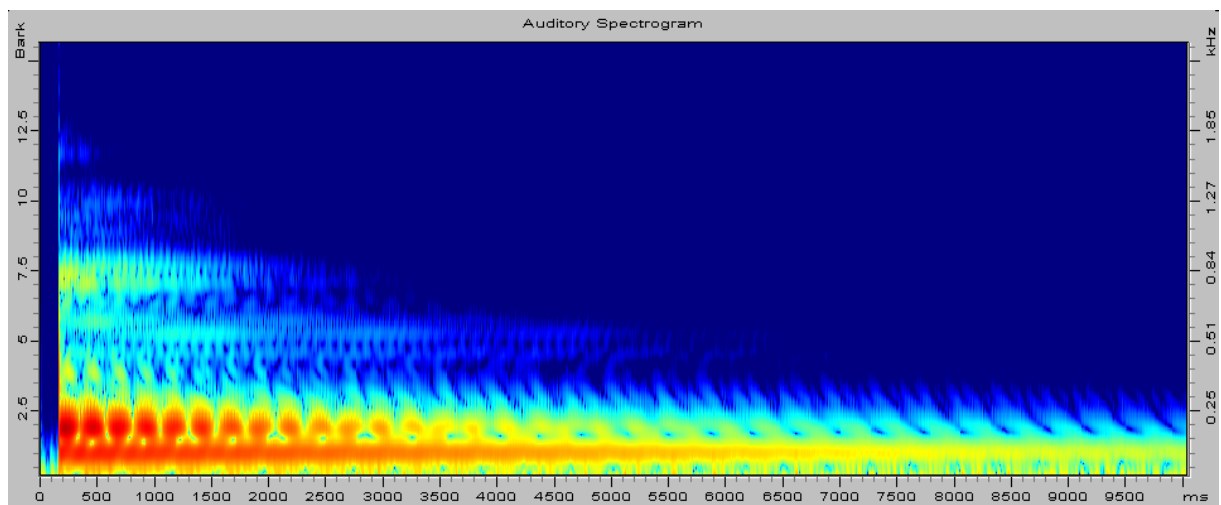


Abb. II.7-1 Auditory spectrogram

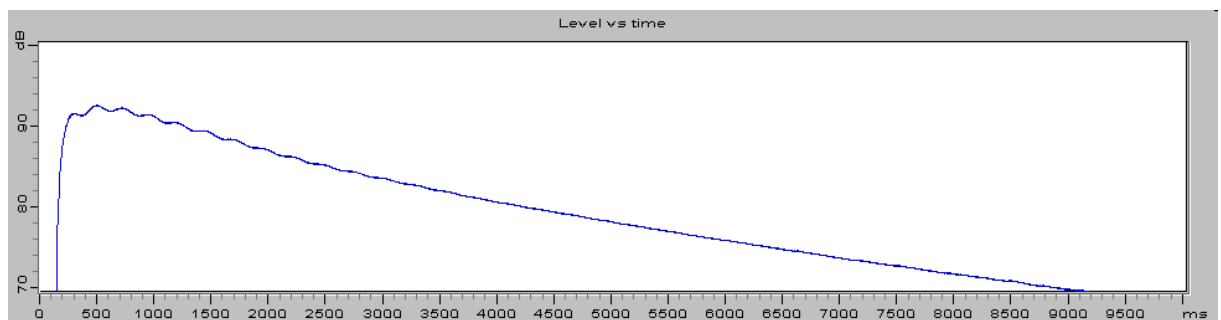


Abb. II.7-2 Level vs. time

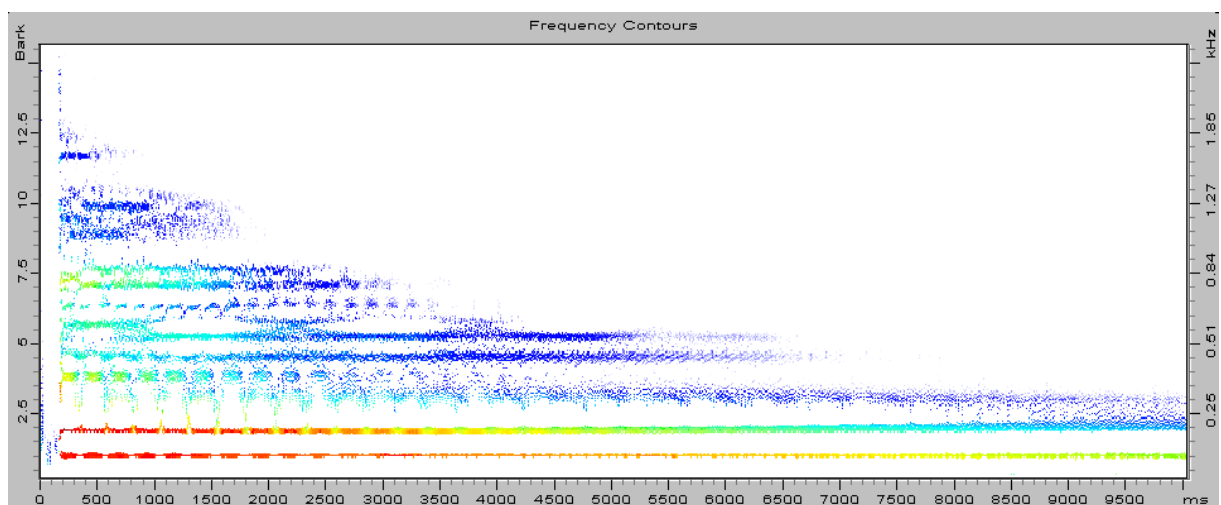


Abb. II.7-3 Frequency contours with nonlinear masking

II.8 Gong VIII, Java-Gong 'Suwokan'

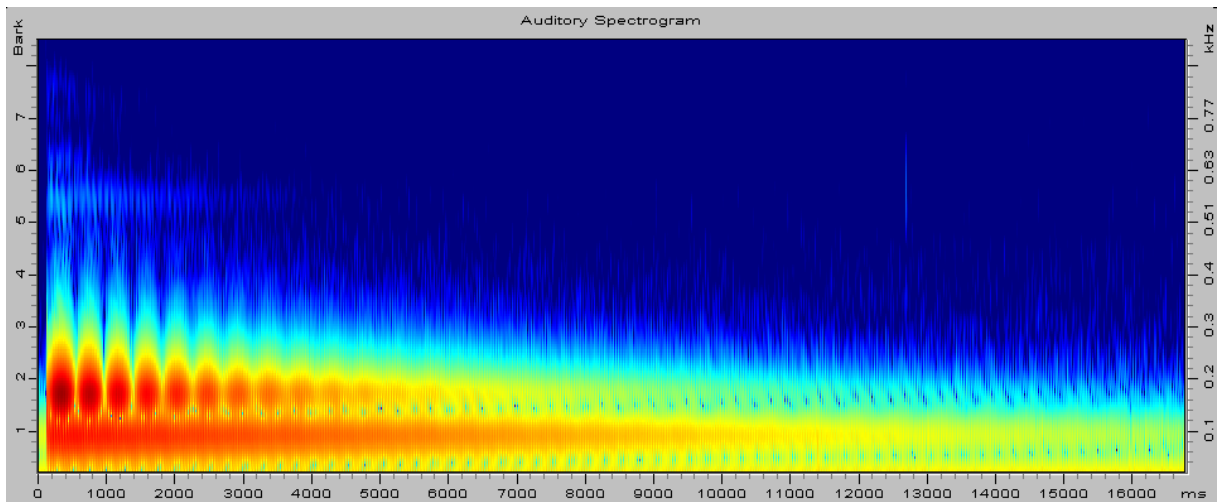


Abb. II.8-1 Auditory spectrogram

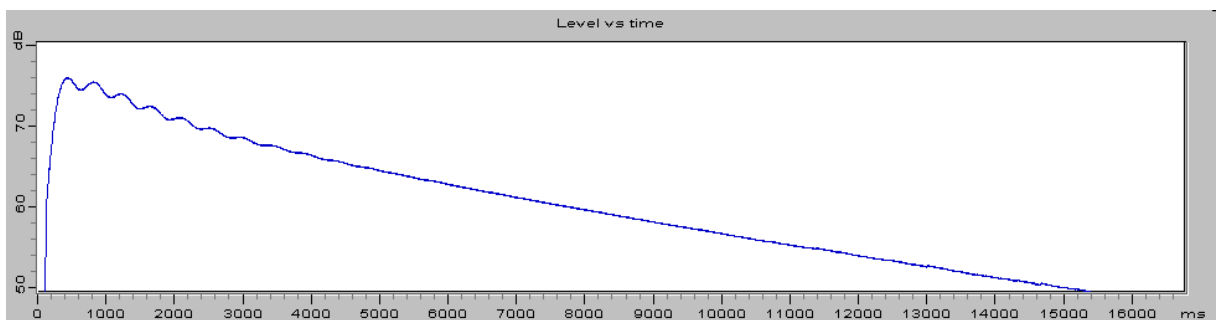


Abb. II.8-2 Level vs. time

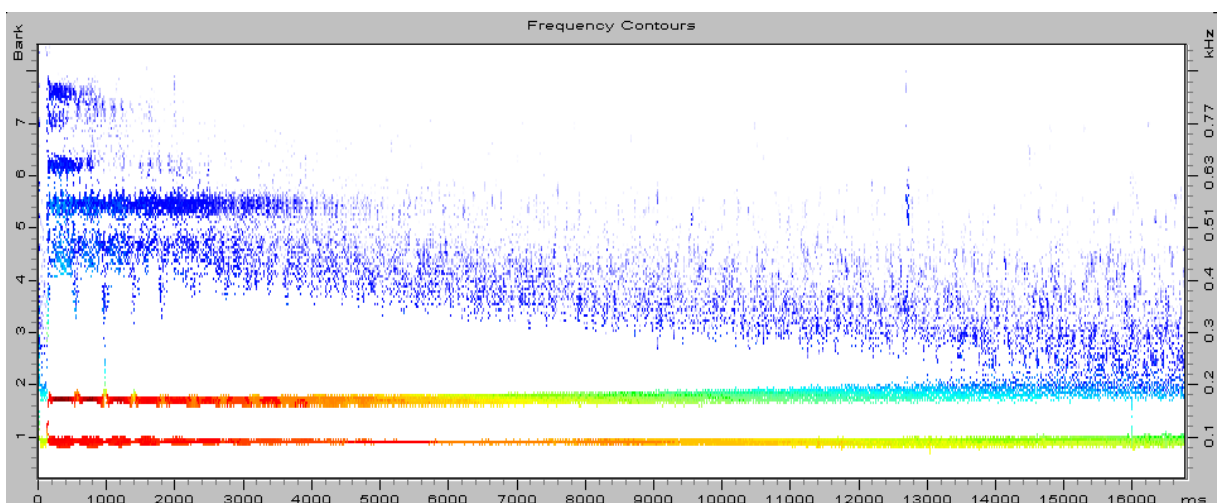


Abb. II.8-3 Frequency contours with nonlinear masking

II.9 Gong IX, Java-Gong 'Ageng'

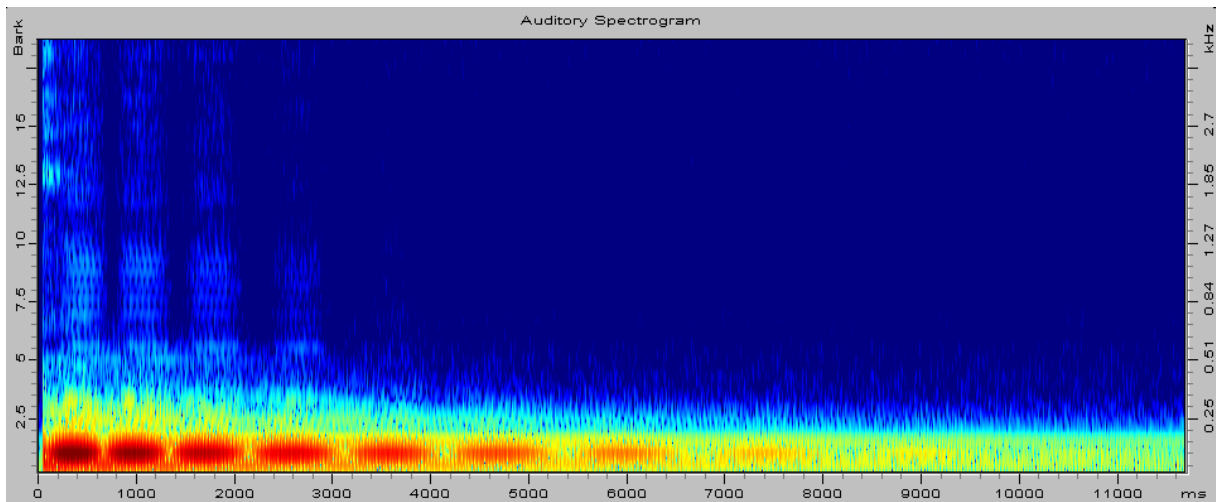


Abb. II.9-1 Auditory spectrogram

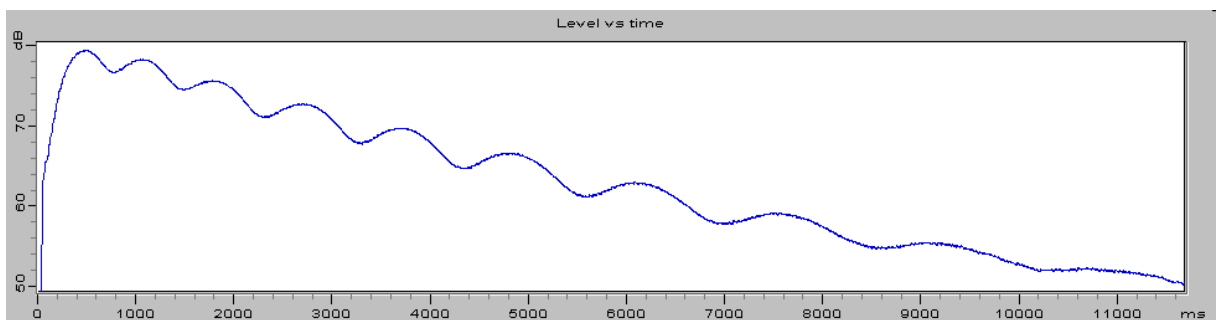


Abb. II.9-2 Level vs. time

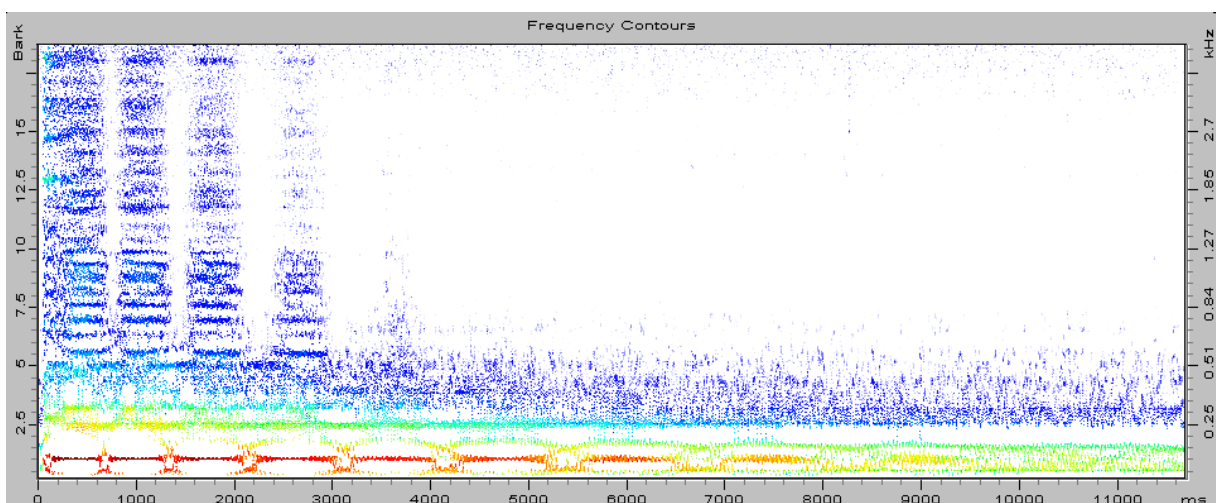


Abb. II.9-3 Frequency contours with nonlinear masking

II.10 Gong X, Flachgong 'Feng-Luo', sehr laut angeschlagen

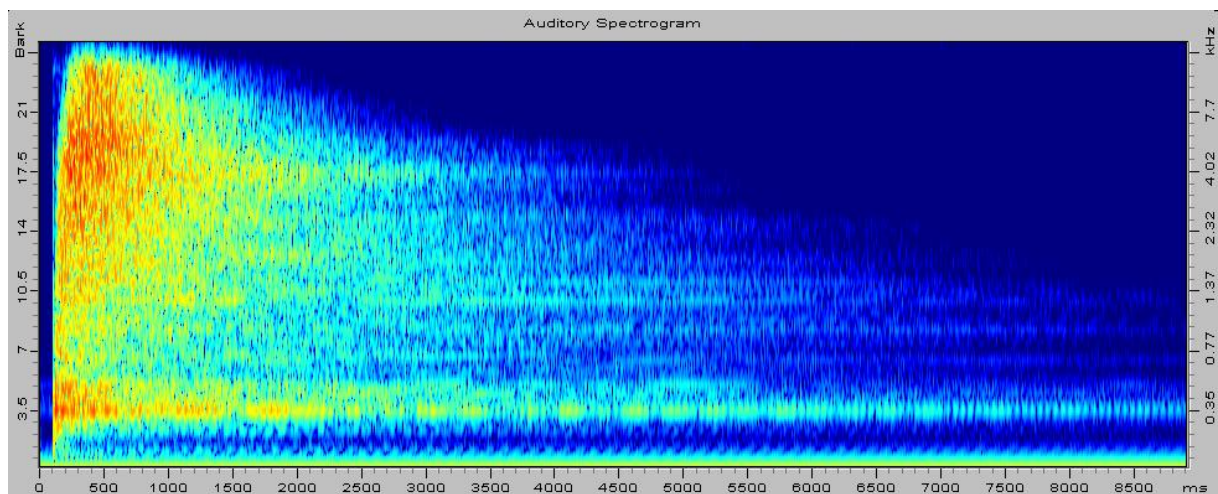


Abb. II.10-1 Auditory spectrogram

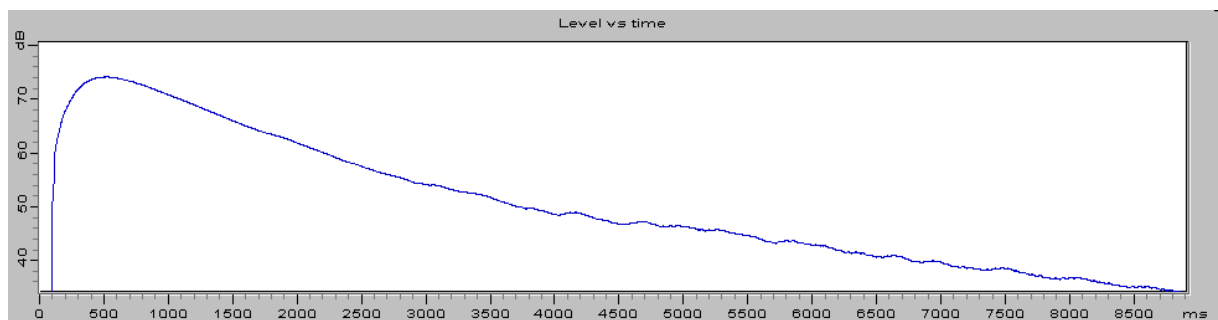


Abb. II.10-2 Level vs. time

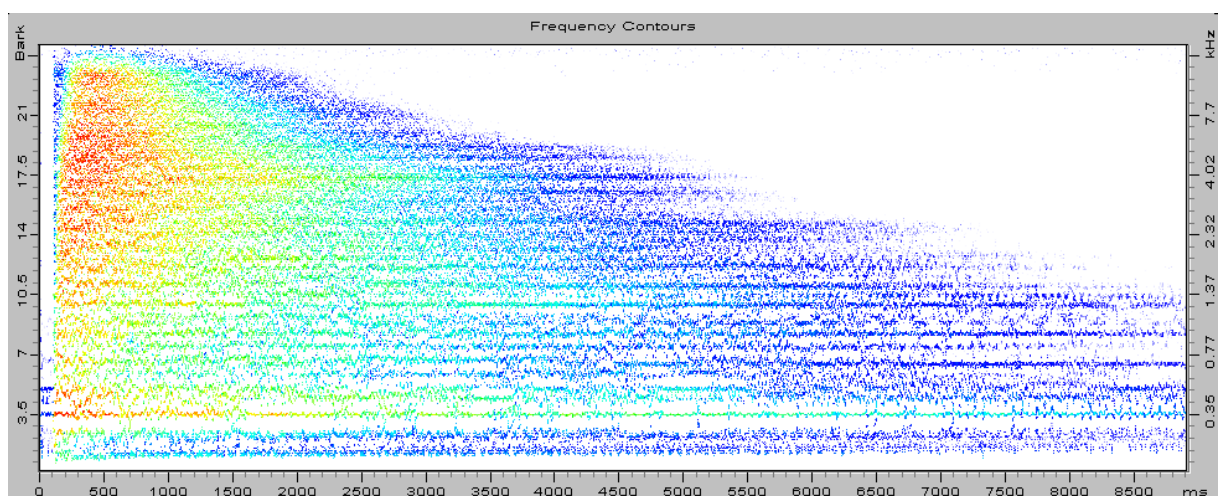


Abb. II.10-3 Frequency contours with nonlinear masking

II.11 Gong XI, Tam-Tam 'Firma Paiste'

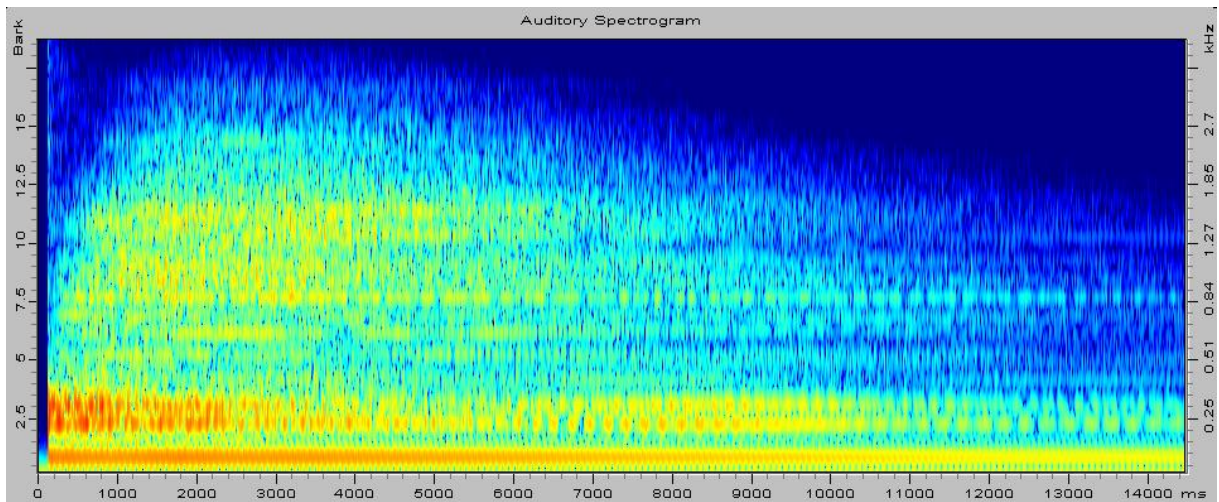


Abb. II.11-1 Auditory spectrogram

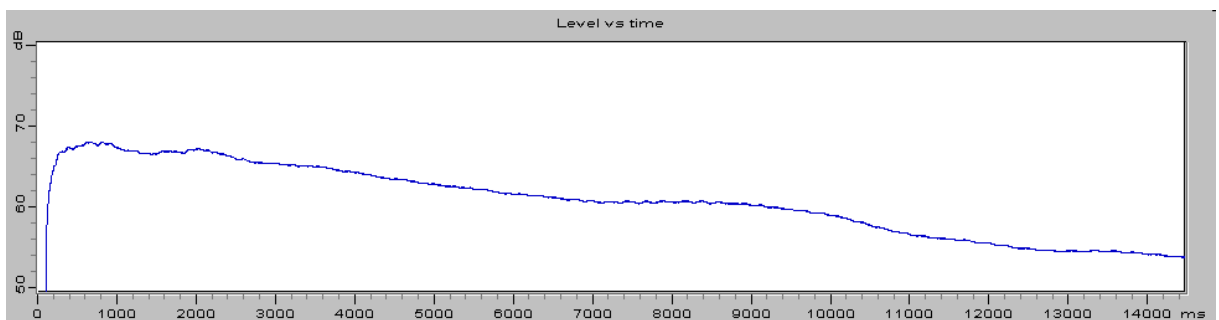


Abb. II.11-2 Level vs. time

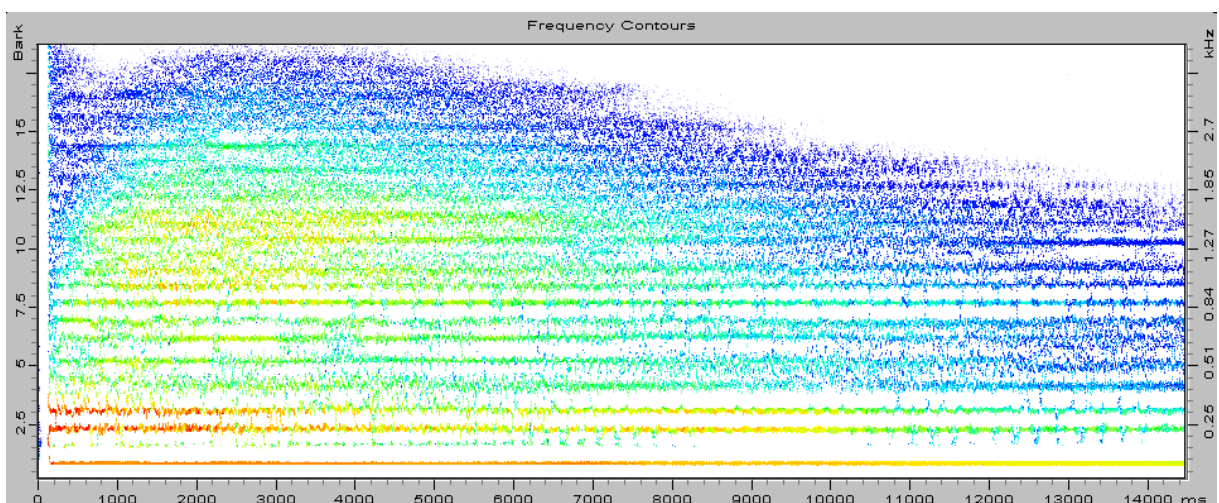


Abb. II.11-3 Frequency contours with nonlinear masking

III Literaturverzeichnis

Valenzuela, Miriam Noemi; Fleischer, Helmut (Herausg.); Fastl, Hugo (Herausg.): Zur Rolle des Gehörs bei akustischen Untersuchungen an Musikinstrumenten; Beiträge zur Vibro- und Psychoakustik. Neubiberg, Heft 1/99

VIPER Manual “New Features”, Neutrik Cortex Instruments, Version 2.12

VIPER PerceptualXplorer Manual, Neutrik Cortex Instruments, Version 1.11. July 2000

Zwicker, E.; Fastl, H.: Psychoacoustics – Facts and Models. Berlin, 1990 (zweite überarbeitete Auflage)

Zwicker, E.; Feldtkeller, R.: Das Ohr als Nachrichtenempfänger. Stuttgart, 1967 (zweite überarbeitete Auflage)

Zwicker, E.: Psychoakustik. Berlin, 1982

IV CD-Rom

Inhalt:

\Archiv	Archiv mit verschiedenen Schallsignalen im Microsoft® Windows™ <i>Wave</i> -Format (siehe nächste Seite)
\Beispiele\	VIPER- und <i>.wav</i> -Dateien zu den in dieser Arbeit behandelten Themen
\Gongs\	VIPER- und (kalibrierte) <i>.wav</i> -Dateien mit der Analyse von 11 verschiedenen Gongs (<i>Gongs.WKS</i>)

Dateien im Archiv:

Name	In Ordner
01 Kalbriersinuston 1 kHz, 90 dB.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
02 Kolberg-Pauke D2 (73 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
03 Kolberg-Pauke A2 (110 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
04 Kolberg-Pauke D3 (147 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
05 Kolberg-Gong Nr 1 (164 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
06 Kolberg-Gong Nr 3 (154 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
07 Kolberg-Gong Nr 4 (101 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
08 Gong Gis (103 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
09 Gong Pappelstr (93 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
10 Perner-Glocke 34kg (513 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
11 Perner-Glocke Klein (639 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
12 Tischglocke (1,67 kHz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
13 Kalbriersinuston 1 kHz, 94 dB.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
14 Winklein2 'vglocke' (491 Hz).wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
15 Kalbriersinuston 1 kHz, 94 dB.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
16 Hoefner Nr 1, alte Saiten, E.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
17 Hoefner Nr 1, alte Saiten, A.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
18 Hoefner Nr 1, alte Saiten, D.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
19 Hoefner Nr 1, alte Saiten, G.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
20 Hoefner Nr 1, alte Saiten, H.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
21 Hoefner Nr 1, alte Saiten, e.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
22 Wandgitarre 2, alte Saiten, E.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
23 Wandgitarre 2, alte Saiten, A.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
24 Wandgitarre 2, alte Saiten, D.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
25 Wandgitarre 2, alte Saiten, G.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
26 Wandgitarre 2, alte Saiten, H.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
27 Wandgitarre 2, alte Saiten, e.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
28 Hoefner Nr 1, neue Saiten, E.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
29 Hoefner Nr 1, neue Saiten, A.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
30 Hoefner Nr 1, neue Saiten, D.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
31 Hoefner Nr 1, neue Saiten, G.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
32 Hoefner Nr 1, neue Saiten, H.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
33 Hoefner Nr 1, neue Saiten, e.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
34 Kalbriersinuston 1 kHz, 94 dB.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
35 Applause, alte Saiten, E.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
36 Applause, alte Saiten, A.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
37 Applause, alte Saiten, D.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
38 Applause, alte Saiten, G.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
39 Applause, alte Saiten, H.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
40 Applause, alte Saiten, e.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
41 Applause, neue Saiten, E.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
42 Applause, neue Saiten, A.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
43 Applause, neue Saiten, D.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
44 Applause, neue Saiten, G.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
45 Applause, neue Saiten, H.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
46 Applause, neue Saiten, e.wav	Archiv\Fleischer & Valenzuela Maerz 1998, 'Analysierte Schalle', 00..46
50 94-dB Pegelton.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
51 Gong VII, grosser Kolberg-Gong Pappelstrasse, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
52 Gong VII, grosser Kolberg-Gong Pappelstrasse, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
53 Gong VI, grosser Kolberg-Gong GIS 56, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
54 Gong VI, grosser Kolberg-Gong GIS 56, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
55 Gong V, grosser Kolberg-Gong 53 cm, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
56 Gong V, grosser Kolberg-Gong 53 cm, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
57 Gong IV, mittlerer Kolberg-Gong Tobi, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
58 Gong IV, mittlerer Kolberg-Gong Tobi, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
59 Gong III, mittlerer Kolberg-Gong, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
60 Gong III, mittlerer Kolberg-Gong, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
61 Gong I, kleiner Kolberg-Gong, frontal.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
62 Gong I, kleiner Kolberg-Gong, seitlich.wav	Archiv\Burma-Gongs im RAR, Dezember 2000, 50..62
71 Gong IX, Java-Gong 'Ageng', leise, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
72 Gong IX, Java-Gong 'Ageng', laut, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
73 Gong X, Flachgong 'Feng-Luo', leise, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
74 Gong X, Flachgong 'Feng-Luo', mittel, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
75 Gong X, Flachgong 'Feng-Luo', sehr laut, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
76 Gong XI, Tam-Tam 'Firma Paiste', leise, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
77 Gong XI, Tam-Tam 'Firma Paiste', laut, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
78 Gong VIII, Java-Gong 'Suwokan', leise, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
79 Gong VIII, Java-Gong 'Suwokan', laut, rechter Kanal.wav	Archiv\Fleischer & Warschani Juni 1996, 'Java-Gongs und Tam-Tams', 71.
81 Gong II, Kolberg-Gong 216 Hz (Ewa).wav	Archiv\Fleischer Juni 1997, 'Kolberg-Gongs', 81