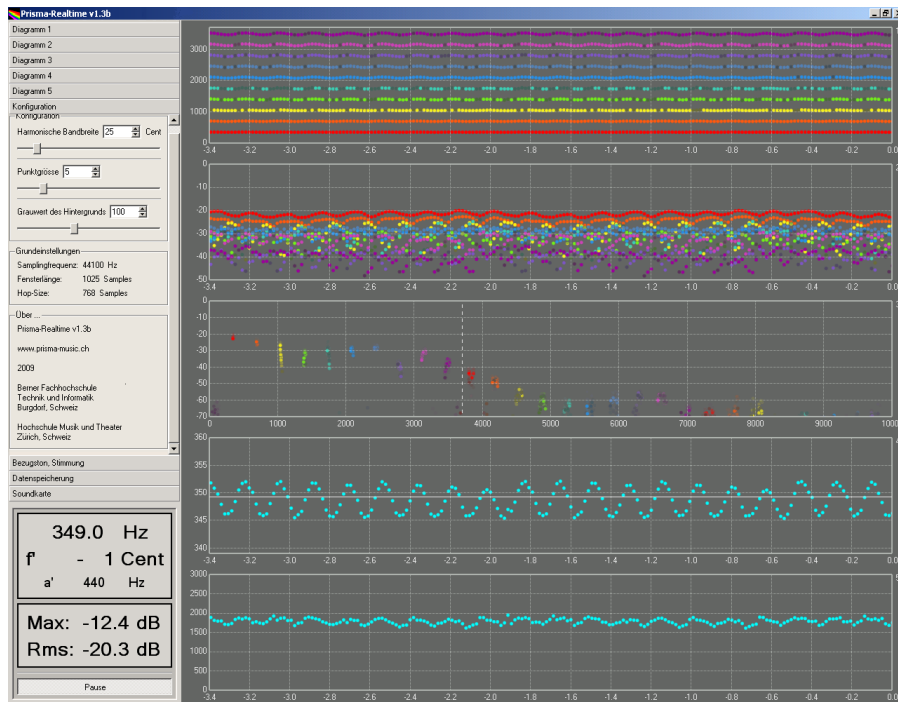


Prisma-Realtime



Version 1.3b

24. März 2010

Berner Fachhochschule (Technik und Informatik), Burgdorf, Schweiz
Hochschule Musik und Theater, Zürich, Schweiz

Franz Bachmann
Michael Bernhard
Hans-Christof Maier

Inhaltsverzeichnis

Prisma-Realtime	1
1 Produktbeschreibung	3
2 Systemvoraussetzungen	4
3 Installation	5
4 Hintergrundinformation	5
5 Programmfenster	6
6 Bedienung	8
6.1 Diagrammtypen	8
6.2 Frequenz- und Amplitudenbereich	13
6.3 Konfiguration	14
6.4 Bezugston, Stimmung	15
6.5 Datenspeicherung	16
6.6 Soundkarte	19
6.7 Standardeinstellungen, Konfigurationsdatei	20
7 Tipps und Tricks	22
7.1 Grafik	22
7.2 Sound	23
8 Support	24

1 Produktbeschreibung

Prisma-Realtime ist eine Applikation zur Anzeige und Analyse von Tönen in Echtzeit auf einem Windows-PC. Das Programm kann mit jeder Soundkarte betrieben werden und benötigt für die Darstellung auf dem Bildschirm eine leistungsfähige Grafikkarte.

Prisma-Realtime funktioniert wie ein Messgerät. Es berechnet in kurzen, aufeinanderfolgenden Zeitfenstern exakte spektrale Daten eines gespielten Tons und stellt diese auf verschiedene Arten auf dem Bildschirm dar. Berechnung und Visualisierung erfolgen in Echtzeit, also praktisch verzögerungsfrei.

Prisma-Realtime enthält einen sehr genauen Algorithmus zur Bestimmung der Grundfrequenz eines Tons und kann deshalb auch als einfaches Stimmgerät verwendet werden.

Prisma-Realtime wurde für die Analyse von Einzeltönen entwickelt. Bei mehrstimmigen Klängen funktioniert die Grundfrequenzerkennung nicht und die Farbgebung der Diagramme wird willkürlich. Es kann jedoch ein Spektrogramm-Modus gewählt werden, der unabhängig von einer Grundfrequenz sinnvolle Resultate liefert.

Prisma-Realtime führt die Tonanalyse in aufeinanderfolgenden, überlappenden Zeitfenstern durch. Die Länge der Zeitfenster und das Ausmass der Überlappung sind wichtige *Systemparameter* mit Auswirkung auf die Genauigkeit und die zeitliche Auflösung der Resultate. Der Benutzer kann einzelne dieser Parameter ändern und damit das Programm seinen Bedürfnissen anpassen .

Prisma-Realtime ist eine Entwicklung der Berner Fachhochschule und der Hochschule Musik und Theater in Zürich. Es ist im Rahmen eines Forschungsprojekts entstanden, das von der Aktion DORE des Schweizerischen Nationalfonds unterstützt wurde. Die Software kann kostenlos von der Website www.prisma-music.ch heruntergeladen werden.

2 Systemvoraussetzungen

Die Systemvoraussetzungen sind nicht einfach zu definieren. Im Rahmen des Forschungsprojekts stand bei Prisma-Realtime die Funktionalität und nicht die Portabilität im Vordergrund. Trotzdem können einige Angaben zu den Anforderungen an die Hardware gemacht werden:

Soundkarte Für den Betrieb von Prisma-Realtime ist eine Soundkarte unabdingbar. Grundsätzlich sind alle Soundkarten möglich, welche in Windows installiert werden können. Insbesondere beim Betrieb mit einem Laptop ist es aber empfehlenswert, eine gute externe Soundkarte zu verwenden. Modelle mit eingebautem Vorverstärker und Phantomspeisung machen die Ausrüstung kompakt und mobil.

Mikrofon Damit Töne in Echtzeit analysiert werden können, braucht es ein Mikrofon. Die Qualität von Mikrofon und Soundkarte ist – neben guten akustischen Umgebungsbedingungen – für die Aussagekraft der Diagramme entscheidend. Trotzdem läuft Prisma-Realtime auch schon mit einem einfachen Mikrofon und der „onboard“ Soundkarte.

Es ist auch möglich, aufgenommene Klänge abzuspielen (z.B. im Windows Media Player) und in Prisma-Realtime einzuspeisen. Jedoch muss die Aufnahme in einem unkomprimierten Datenformat (z.B. wav, aber nicht mp3) gespeichert sein.

Grafikkarte Die Echtzeit-Darstellung der Analyseergebnisse auf dem Bildschirm erfordert eine leistungsfähige Grafikkarte. Tests haben gezeigt, dass aktuelle 3D-Karten von ATI und NVIDIA funktionieren, Intel Grafikkarten dagegen zum Programmabsturz führen.

Bildschirm Je grösser die Bildschirmauflösung, desto leistungsfähiger muss die Grafikkarte sein, da mehr Bildpunkte gezeichnet werden müssen. Alternativ kann das Programmfenster verkleinert werden. Prisma-Realtime versucht, die Diagramme im Takt der Bildwiederholung des Bildschirms zu zeichnen. Gelingt dies, so entsteht ein ruhiges und kontinuierliches Bild. Ein Flachbildschirm mit 60Hz Bildwiederholfrequenz ist optimal.

Prozessor Eine moderne CPU von Intel oder AMD genügt. Da viel Rechenarbeit auf die Grafikkarte ausgelagert wird, ist der Prozessor nicht allzu stark belastet.

RAM Prisma-Realtime benötigt ca. 20 MB RAM. Wird der Buffer zum Speichern erhöht, steigt auch der Platzbedarf im RAM an.

Windows Die Software wurde unter Windows XP SP2 entwickelt. Windows 2000 und Windows XP werden unterstützt. Einige Tests mit Windows Vista und Windows 7 waren erfolgreich.

Zusammengefasst: Ein moderner PC mit aktueller 3D-Grafikkarte und eine Soundkarte mit Mikrofon reichen aus, um Prisma-Realtime zu betreiben.

Hier noch die technischen Daten einiger erfolgreicher Testsysteme:

- Desktop, Intel Pentium IV HT 3GHz, 1GB RAM, NVIDIA GeForce FX 5700, Windows XP Pro SP2
- Desktop, Intel Pentium IV HT 3GHz, 1GB RAM, ATI Radeon X600 Pro, Windows XP Pro SP2
- Notebook, Intel Pentium M 1.8GHz, 1GB RAM, ATI Mobility Radeon X300, Windows XP Pro SP2
- MacBookPro3.1, Intel Core 2 Duo 2.2GHz, 2GB RAM, NVIDIA GeForce 8600M GT, Windows XP Pro SP3

3 Installation

- Die Datei `prisma_realtime_v1_3b.zip` von <http://www.prisma-music.ch> herunterladen und lokal speichern
- Diese Zip-Datei in einen neuen Ordner entpacken
- Die Datei `prisma.exe` ausführen: Prisma-Realtime startet
- Optional kann eine Verknüpfung zu `prisma.exe` auf dem Desktop erstellt werden.

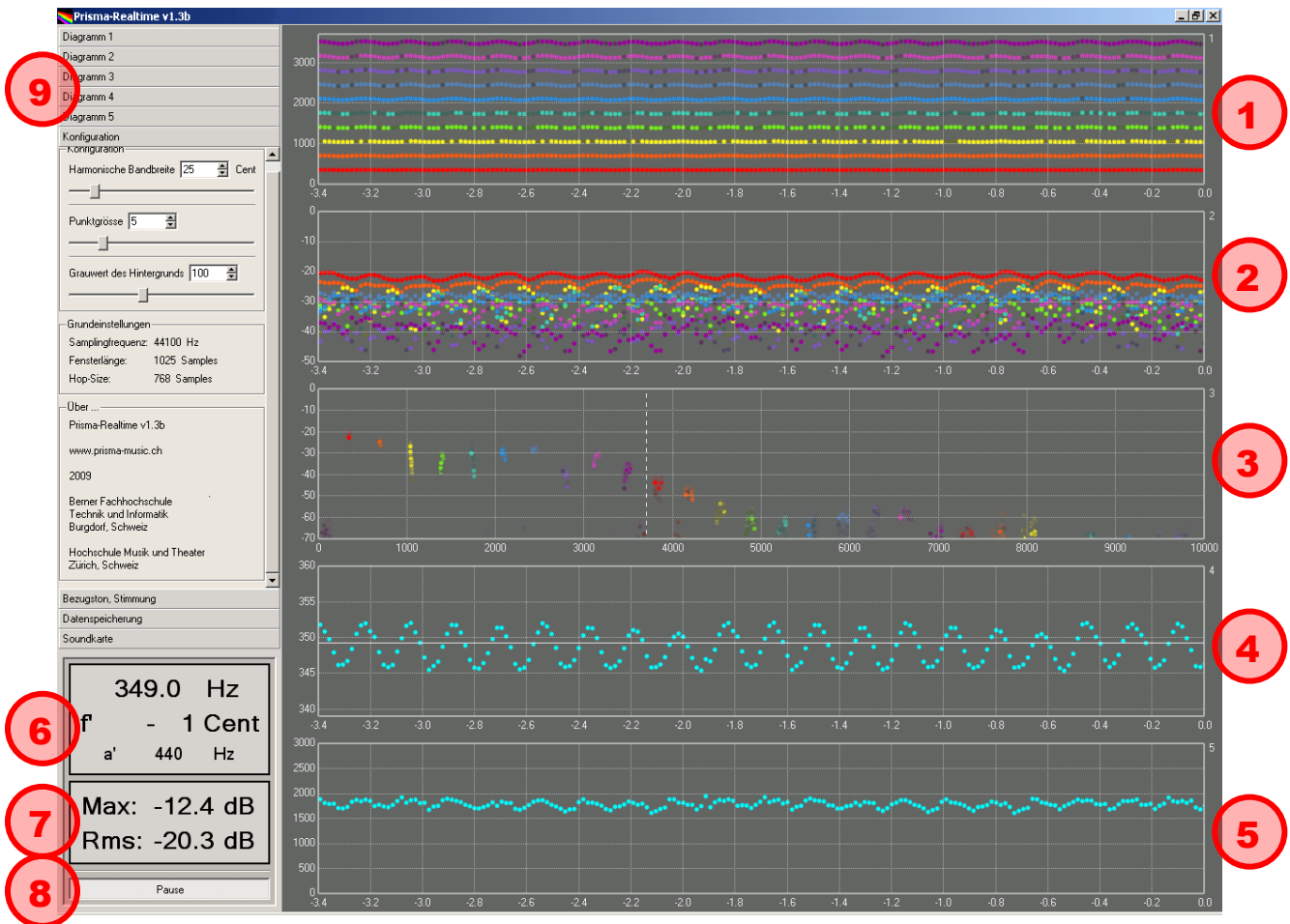
4 Hintergrundinformation

Informationen zu den verwendeten Algorithmen und zur Software-Entwicklung finden sich in der Arbeit

F. Bachmann, M. Bernhard, H.-C. Maier, PRISMA – A TIMBRE TUNING DEVICE FOR MUSICIANS AND INSTRUMENT MAKERS, Proceedings of the International Symposium for Musical Acoustics (ISMA), Barcelona (2007)

Diese Arbeit kann ebenfalls von der oben erwähnten Prisma-Website heruntergeladen werden.

5 Programmfenster



- ① Zeit-Frequenz-Diagramm
- ② Zeit-Amplituden-Diagramm
- ③ Frequenz-Amplituden-Diagramm
- ④ Fenster für Shorttime-Feature (z.B. Grundfrequenz)
- ⑤ Fenster für Shorttime-Feature (z.B. Spectral Centroid)
- ⑥ Grundfrequenz-/Tonhöhen-Anzeige
- ⑦ Level-Meter
- ⑧ Pause-Taste
- ⑨ Bedienfeld

Prisma-Realtime kann maximal fünf Diagramme anzeigen. Sie sind am rechten Rand des Grafikfensters nummeriert. Die Nummer stimmt mit den Tabs des Bedienfeldes ⑨ im linken Teil des Fensters überein. Einzelne Diagramme können ein- und ausgeblendet werden.

Die Diagramme haben eine einheitliche Farbgebung. Punkte in hellen Farben kennzeichnen harmonische Frequenzkomponenten, die in der Nähe von ganzzahligen Vielfachen der Grundfrequenz liegen. Die dem Regenbogen nachempfundene Färbung läuft von der rot gezeichneten Grundschwingung bis zur violetten 10. Teilschwingung, danach wiederholt sich die Abfolge der Farben. Nicht harmonische Anteile des Tons werden in einem dunkleren und transparenten Farbton dargestellt.

Im Diagramm ① kann zusätzlich das *Spektrogramm* eines Klangs betrachtet werden. Jeder Punkt der Zeit-Frequenz-Ebene wird entsprechend seiner Intensität (Amplitude) gefärbt: rot steht für grosse Amplituden, blau für kleine.

Der in den Diagrammen ① und ② gewählte Frequenzbereich erscheint mit schwarzen Marker-Linien im Diagramm ③. Man behält so den Überblick über die angezeigten Daten.

In den Diagrammen ④ und ⑤ kann der zeitliche Verlauf von numerischen Kenngrößen („Shorttime Features“) des Klangs angezeigt werden. Dazu gehören Grundfrequenz, Intonation oder Lautstärke, aber auch weitere, aus den spektralen Größen berechnete Werte, die in 6.2 und 6.5 näher beschrieben sind.

Die Grundfrequenz-/Tonhöhen-Anzeige ⑥ hat drei Komponenten. Auf der ersten Zeile wird die aktuelle Tonhöhe als Grundfrequenz in Hertz dargestellt. Die zweite Zeile zeigt die Tonhöhe als Abweichung vom nächstgelegenen Halbton (Intonation) in Cent, basierend auf der Stimmung des Kammertons a`.

Schliesslich zeigt das Level-Meter ⑦ die aktuelle Intensität des Eingangssignals mit dem Maximal- und RMS-Wert in Dezibel (dB) an.

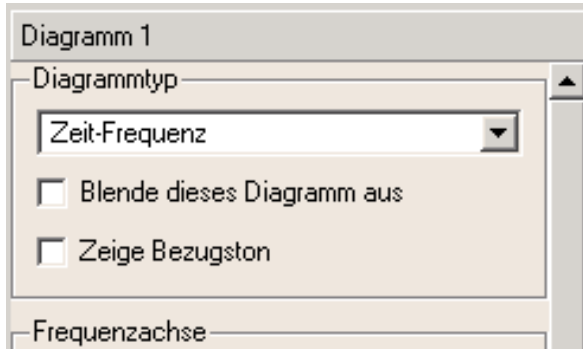
Mit der Pause-Taste ⑧ lässt sich die Anzeige einfrieren. Im Hintergrund laufen die Berechnungen aber weiter, und bei nochmaligem Drücken der Taste werden die dann aktuellen Daten wieder angezeigt. Gewisse Funktionen können nur ausgeführt werden, wenn das System sich im Pause-Modus befindet. Anstatt die Pausen-Taste auf dem Bildschirm anzuklicken kann auch auf der Computertastatur die Funktionstaste F9 gedrückt werden.

Die Konfiguration von Prisma-Realtime geschieht über das Bedienungsfeld ⑨ auf der linken Seite des Fensters. Jede Funktionalität ist in einem eigenen Tab untergebracht. Im folgenden Kapitel werden diese im Detail beschrieben.

6 Bedienung

6.1 Diagrammtypen

6.1.1 Diagramm 1



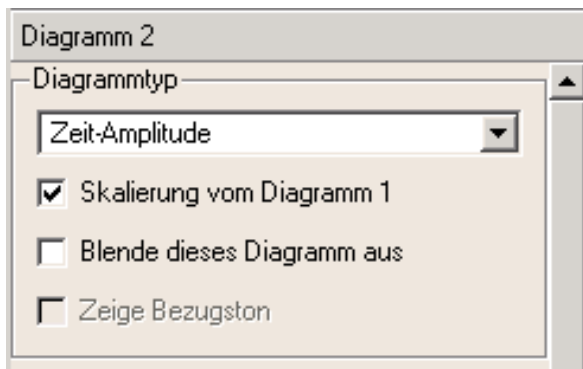
Im Diagramm 1 lassen sich drei Diagrammtypen einstellen:

- Zeit-Frequenz
- Zeit-Amplitude
- Spektrogramm

Mit der Checkbox „Blende dieses Diagramm aus“ kann das Diagramm ein- und ausgeblendet werden. Auf der

Frequenzachse – sofern vorhanden – kann der Bezugston (6.4) angezeigt werden.

6.1.2 Diagramm 2



Im Diagramm 2 lassen sich zwei Diagrammtypen einstellen:

- Zeit-Frequenz
- Zeit-Amplitude

Mit der Checkbox „Blende dieses Diagramm aus“ kann das Diagramm ein- und ausgeblendet werden.

Die zusätzliche Checkbox „Skalierung vom Diagramm 1“ übernimmt die Einstellungen der Frequenz- und Amplituden-Achsen dynamisch vom Diagramm 1. Somit lassen sich zwei Ansichten mit einer einzigen Einstellung steuern.

Achtung: Auch wenn das Diagramm 1 ausgeblendet ist, bleiben seine Einstellungen gültig. Ist eine Frequenzachse sichtbar, kann darauf auch hier der Bezugston angezeigt werden.

6.1.3 Diagramm 3

Das Diagramm 3 bietet nur einen Diagrammtyp:

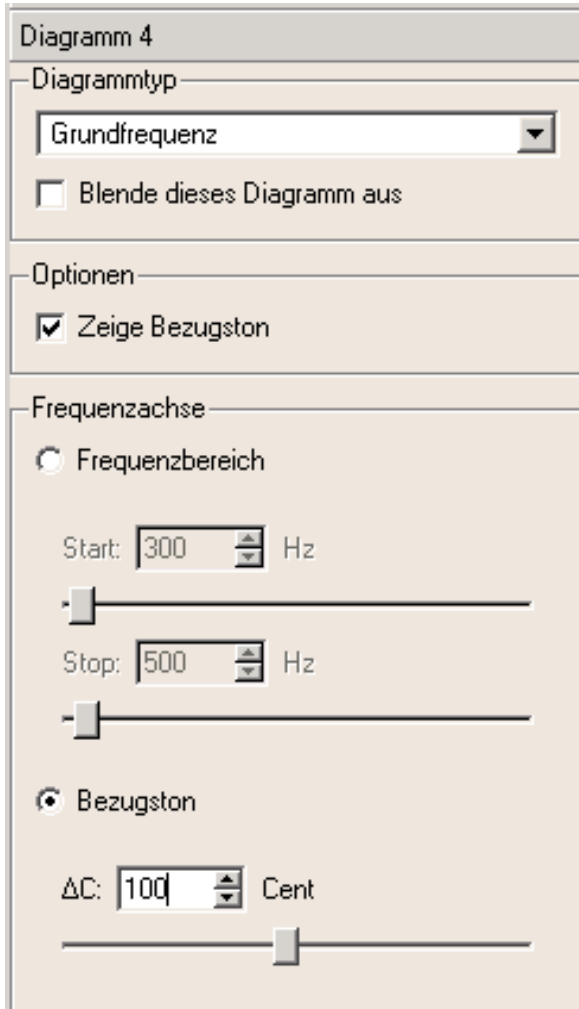
- Frequenz-Amplitude

Auch dieses Diagramm lässt sich mit der Checkbox „Blende dieses Diagramm aus“ ein- und ausschalten, ebenso die Anzeige des Bezugstons auf der Frequenzachse, der als durchgehende Linie dargestellt wird. Mit gestrichelten Linien werden die Frequenzgrenzen der Anzeigebereiche in den Diagrammen 1 und 2 angezeigt.

6.1.4 Diagramme 4 und 5

Die Diagramme 4 und 5 bieten die Möglichkeit, den Verlauf von *Shorttime Features* über die Zeit darzustellen. Das sind Kenngrößen des Klangs, die in Frames (Zeitfensters) definiert sind. Folgende Features können gewählt werden:

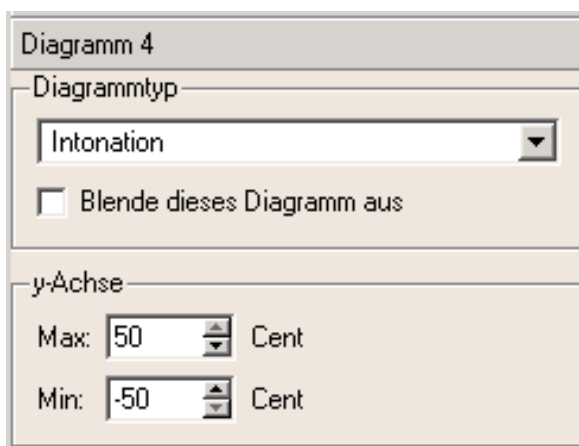
6.1.4.1 Grundfrequenz



Die *Grundfrequenz* (in Hz) wird mit einem hochpräzisen Verfahren ermittelt und ist ein grundlegendes Shorttime Feature.

Die Frequenzachse kann auf zwei Arten skaliert werden: Entweder wird der Frequenzbereich direkt mit den Schieberegler oder Spinboxes definiert oder es wird ein Bezugston (6.4) gewählt und ein zugehöriges Frequenzband in Cent mit dem Schieberegler oder einer Spinbox eingestellt.

6.1.4.2 Intonation



Im Kontext von Prisma-Realtime ist die *Intonation* die Abweichung der aktuellen Tonhöhe vom nächstgelegenen Halbton einer gleichschwebend temperierten musikalischen Notenskala. Die Intonation wird in Cent, einem logarithmischen Mass, angegeben, wobei einem Halbton 100 Cent entsprechen.

Die y-Achse kann mittels Spinboxes skaliert werden.

6.1.4.3 RMS & MAX Amplitude

The screenshot shows a control panel for 'Diagramm 4'. It is divided into three sections: 'Diagrammtyp', 'Kurve', and 'Amplitudenachse'. In the 'Diagrammtyp' section, a dropdown menu is set to 'RMS & MAX Amplitude' and there is an unchecked checkbox 'Blende dieses Diagramm aus'. In the 'Kurve' section, there are two checked checkboxes: 'Max' and 'RMS'. In the 'Amplitudenachse' section, there are two spinboxes: 'Max' is set to '0 dB' and 'Min' is set to '-70 dB'.

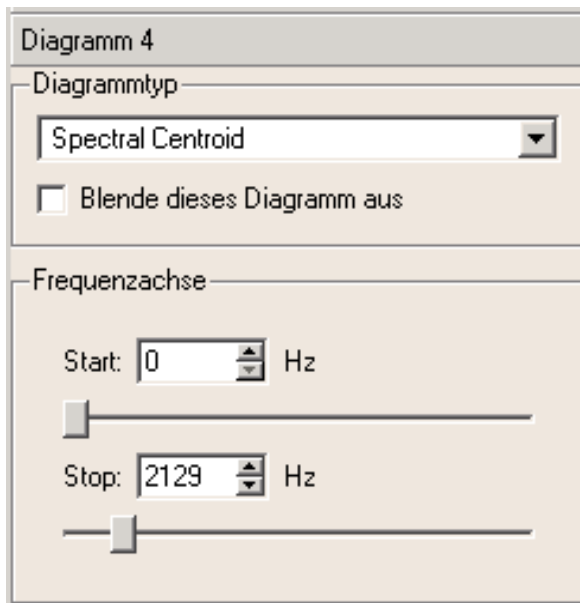
Die *RMS* (Root Mean Square) ist ein Mass für die mittlere Signalstärke im aktuellen Frame, hier also auch für die Lautstärke. *MAX Amplitude* gibt die maximale Signalstärke an. Beide können in den Checkboxes separat gewählt werden. Die Skalierung der Amplitudenachse in dB erfolgt in den darunter stehenden Spinboxes.

6.1.4.4 Partial

The screenshot shows a control panel for 'Diagramm 4'. It is divided into three sections: 'Diagrammtyp', 'Kurve', and 'Amplitudenachse'. In the 'Diagrammtyp' section, a dropdown menu is set to 'Partial' and there is an unchecked checkbox 'Blende dieses Diagramm aus'. In the 'Kurve' section, there are three checked checkboxes: 'Partial 1', 'Partial 2', and 'Partial 3'. In the 'Amplitudenachse' section, there are two spinboxes: 'Max' is set to '0 dB' and 'Min' is set to '-80 dB'.

Hier können die Signalstärken der ersten drei Partial (Teilschwingungen) direkt dargestellt werden. Die Farbcodierung entspricht derjenigen der Diagramme 1 bis 3, also die erste Teilschwingung in rot, die zweite in orange und die dritte in gelb. Wiederum wird die Amplitudenachse in dB mittels Spinboxes skaliert.

6.1.4.5 Spectral Centroid



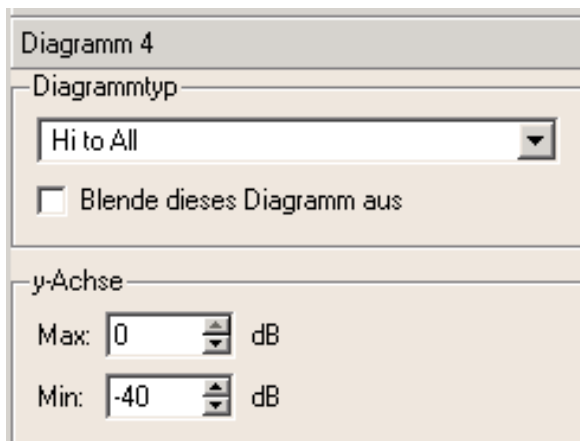
Das *Spectral Centroid* ist eine Kenngröße zur Messung des Obertongehalts eines Klangs. Fasst man das Amplitudenspektrum als Massenverteilung auf, so entspricht das Spectral Centroid dem Schwerpunkt dieser Verteilung.

Eine genauere Beschreibung findet sich in 6.5.

Das Spectral Centroid wird in Hertz gemessen. Ist es markant grösser als die Grundfrequenz, so hat das Signal im aktuellen Zeitfenster meist mehrere starke Obertöne.

Die Frequenzachse kann mit Schiebereglern oder Spinboxes skaliert werden.

6.1.4.6 Hi to All

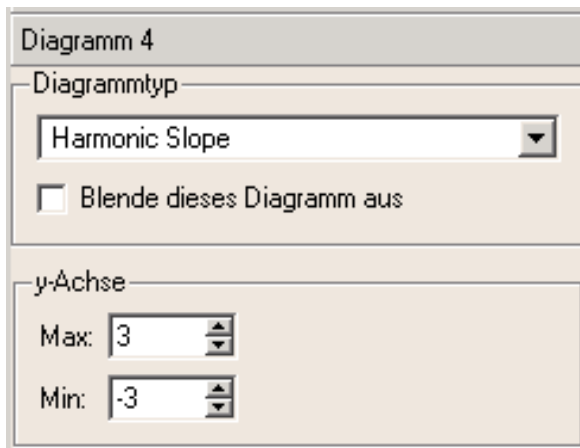


Das Feature *Hi to All* misst den Anteil aller Oberschwingungen des Signals im Verhältnis zum Gesamtsignal und wird in dB gemessen.

Eine genauere Beschreibung findet sich im 6.5.

Der Wertebereich von Hi to All kann mit zwei Spinboxes skaliert werden.

6.1.4.7 Harmonic slope



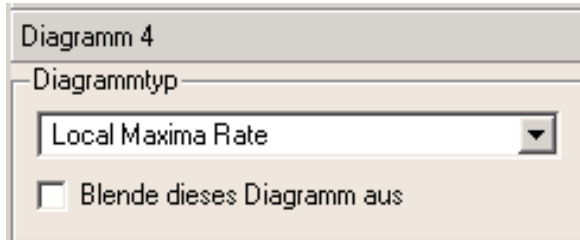
Die *Harmonic Slope* misst das Ausmass des Abfallens der harmonischen Amplituden bei wachsender Frequenz und ist damit ebenfalls ein Mass für den Obertongehalt im aktuellen Frame.

Eine genauere Beschreibung findet sich in 6.5.

Im allgemeinen (aber nicht immer!) ist die Harmonic Slope eine negative Zahl. Liegt sie nahe bei Null, so ist dies ein Indiz dafür, dass mehrere starke Oberschwingungen vorhanden sind.

Die Skalierung des Wertebereichs der Harmonic Slope erfolgt wieder mit zwei Spinboxes.

6.1.4.8 Local maxima rate



Die *Local Maxima Rate* ist ein Mass für die Stärke der Schwankung des Zeitsignals im aktuellen Zeitfenster. Diese Grösse hat Werte zwischen 0 (keine Schwankung, ruhiges Signal) und 1 (starke Schwankung, unruhiges Signal).

6.2 Frequenz- und Amplitudenbereich

The screenshot shows a control panel with two main sections: 'Frequenzachse' and 'Amplitudenachse'.
 In the 'Frequenzachse' section:
 - There are two radio buttons: 'Frequenzbereich' (unselected) and 'Teilschwingung' (selected).
 - Under 'Frequenzbereich', there are two spinboxes: 'Start: 0 Hz' and 'Stop: 3700 Hz', each with a corresponding slider below it.
 - Under 'Teilschwingung', there is a spinbox 'ΔC: 54 Cent' with a slider below it.
 - Below that is a spinbox 'Teilschwingung Nr.: 3' with a slider below it.
 - At the bottom of this section is a checkbox 'Zeige von Teilschwingung 1 an' which is currently unchecked.
 In the 'Amplitudenachse' section:
 - There are two spinboxes: 'Max: 0 dB' and 'Min: -50 dB'.

Der Anzeigebereich der Frequenzachsen in den Diagrammen 1, 2 und 3 kann auf zwei Arten ausgewählt werden: „Frequenzbereich“ oder „Teilschwingung“.

Im Modus „Frequenzbereich“ werden die minimale und die maximale Frequenz des Anzeigebereiches mit Hilfe von Schieberegler oder Spinboxes festgelegt.

Der Modus „Teilschwingung“ ist komplexer. Er erlaubt die gezielte Darstellung einzelner Frequenzbereiche und damit die Untersuchung von Teilschwingungen. In diesem Modus wird auf die Information aus 6.4 zugegriffen. Dort werden der *Bezugston* und die Stimmung, d.h. die Frequenz des Kammertons festgelegt. Wir schreiben f_0 für die Frequenz des Bezugstons.

Wählt man einen Wert ΔC in Cent und als „Nummer der Teilschwingung“ eine

Zahl p , so wird eine Nahaufnahme des Frequenzbereichs

$$f_{\min} \leq p \cdot f_0 \leq f_{\max}$$

dargestellt. Dabei sind die Frequenzgrenzen durch

$$f_{\min} = f_0 \cdot p \cdot 2^{-\frac{\Delta C}{1200}} \quad \text{und} \quad f_{\max} = f_0 \cdot p \cdot 2^{\frac{\Delta C}{1200}}$$

gegeben. Ist zum Beispiel $f_0 = 392\text{Hz}$ (entspricht bei der Stimmung 440 Hz dem Bezugston g'), die maximale Abweichung 54 Cent und $p=3$, so erhält man als Bereichsgrenzen

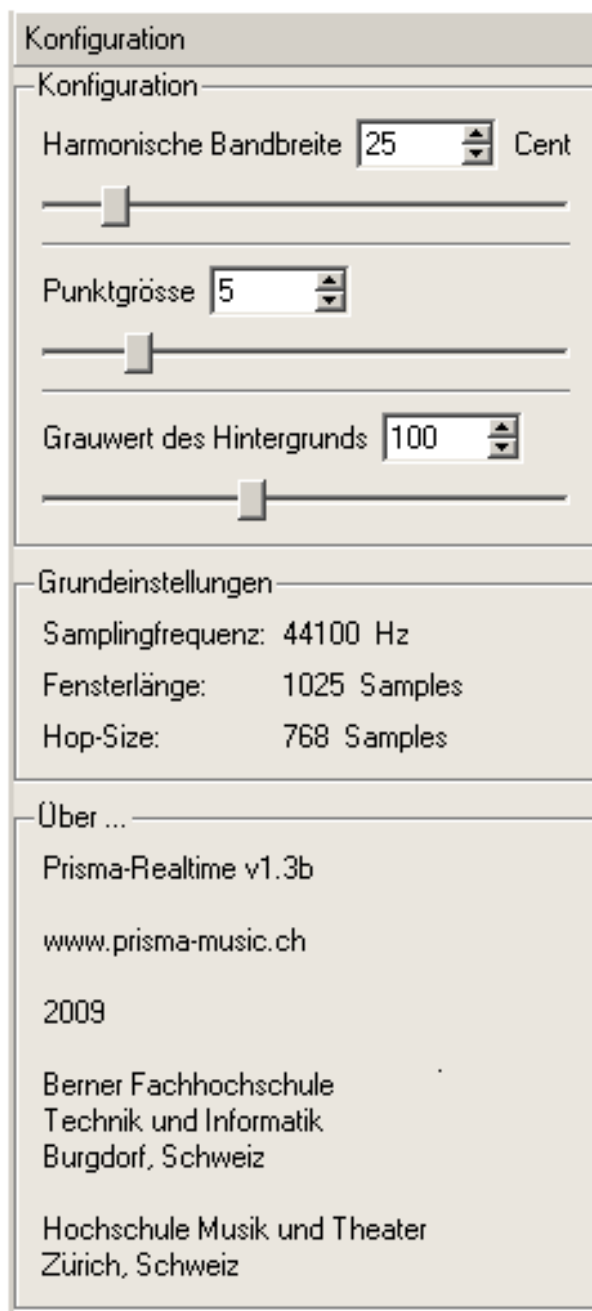
$$f_{\min} = 392 \cdot 3 \cdot 2^{-\frac{54}{1200}} = 1139.9\text{Hz} \quad \text{und} \quad f_{\max} = 392 \cdot 3 \cdot 2^{\frac{54}{1200}} = 1213.3\text{Hz}$$

Aktiviert man die Checkbox „Zeige von Teilschwingung 1 an“ wird f_{\min} auf 0 Hz gesetzt. Man sieht dann den Bereich von 0 Hz bis zur p -ten Teilschwingung.

Der Wert ΔC beeinflusst nur die Achsenskalierung, nicht die Messparameter. Hingegen definiert die *Harmonische Bandbreite* (6.3) den Frequenzbereich um die Teilschwingungen, der als harmonisch betrachtet wird.

Zur Einstellung des Amplitudenbereichs wählt man mit Hilfe zweier Spinboxes die gewünschten Minimal- und Maximalwerte.

6.3 Konfiguration



Im Konfigurations-Tab lassen sich drei Parameterwerte einstellen: „Harmonische Bandbreite“, „Punktgrösse“ und „Grauwert des Hintergrunds“.

Die *Harmonische Bandbreite* legt die Frequenzbereiche der harmonischen Bänder fest. Punkte innerhalb dieser Bänder werden in hellen Regenbogenfarben dargestellt, die restlichen Frequenzkomponenten erscheinen in einem dunkleren und transparenten Farbton. Die *Harmonische Bandbreite* ist nicht zu verwechseln mit der Grösse ΔC (6.2), die allein zur Achsen-skalierung dient.

Der zweite Parameter definiert die Punktgrösse. Zu beachten ist, dass die Leistungsfähigkeit der Grafikhardware ausschlaggebend ist, ob das Bild bei grossen Punkten ruckelt oder nicht.

Schliesslich kann – mittels Spinbox oder Schieberegler – auch der Grauwert des Hintergrunds eingestellt werden. Diese Funktion ist hilfreich, wenn Diagramme ausgedruckt oder projiziert werden sollen.

Im unteren Teil des Tabs werden die Werte dreier in der Konfigurationsdatei (6.7) enthaltenen Analyse-Parameter angezeigt:

- Samplingfrequenz
- Fensterlänge (Frame Length)
- Hop Size (Sprungweite).

Ausserdem finden sich hier Angaben zur verwendeten Version des Programms und das Impressum.

6.4 Bezugston, Stimmung

Entscheidet man sich bei der Wahl des Frequenzbereichs (6.2) für die Variante „Teilschwingungen“, so benötigt Prisma-Realtime einen *Bezugston* mit der Frequenz f_0 . Im vorliegenden Tab kann dieser auf drei Arten definiert werden:

- „feste Vorgabe“
- „manuell übernehmen“ oder
- „automatisch übernehmen“.

Bei der Variante „feste Vorgabe“ stellt man ein, welchen Ton man spielen wird, und zwar in musikalischer Notation. Für die Oktavlage gilt die europäische Bezeichnung: -1=grosse, 0=kleine, 1=eingestrichene Oktave usw. f_0 ist dann

gleich der Grundfrequenz dieses Tons. Grundsätzlich ist der Zusammenhang zwischen einem Frequenzwert und einer musikalischen Tonbezeichnung erst definiert, wenn die *Stimmung* (Frequenz des Kammertons a') bekannt ist. Mit einer Spinbox im unteren Teil des Tabs lässt sich die Stimmung definieren.

Wählt man den Modus „manuell übernehmen“, so wartet Prisma-Realtime, bis die Taste „aktuelle Grundfrequenz“ gedrückt wird, nimmt als Frequenz f_0 die im aktuellen Zeitfenster berechnete Grundfrequenz und als Bezugston die nächstliegende Tonbezeichnung. Diese Einstellung wird beibehalten, bis die Taste „aktuelle Grundfrequenz“ erneut gedrückt wird.

Bei der dritten Möglichkeit „automatisch übernehmen“ wird f_0 immer gleich der momentan berechneten Grundfrequenz gesetzt. Bei instabiler Tonhöhe werden sich bei dieser Variante der Bezugston und damit die Frequenzskala laufend verändern.

6.5 Datenspeicherung

Datenspeicherung

Datenspeicherung

Dateiname (ohne Endung):

Testton

.bin .wav

Speicherpfad:

U:\Prisma ...

Bufferfüllung:

36%

Speichern

Bufferlänge

3 Sekunden

Bei Prisma-Realtime geht es um zwei Typen von Daten: Einerseits Audiodaten, welche das musikalische Signal im Zeitbereich beschreiben und andererseits um die von Prisma-Realtime berechneten Daten der Klanganalyse. Letztere bestehen aus einem Datensatz mit den grundlegenden spektralen Größen (Frequenzen und Amplituden) sowie einer Reihe anderer Kenngrößen (Grundfrequenz, Signalstärke und aus den spektralen Größen abgeleitete Shorttime features) für jedes Zeitfenster.

Beide Typen von Daten können auf einem externen Datenträger gespeichert werden, und zwar die Audiodaten im PCM-Format (.wav) und die

numerischen Daten der Klanganalyse in binärer Form (.bin). Mit Hilfe zweier Checkboxes entscheidet man sich für eine oder beide Varianten.

Nach der Wahl eines Dateinamens (ohne Endung) und eines Speicherpfads löst man die Speicherung mit dem Button „Speichern“ aus. Es werden die Daten der letzten n Sekunden vor der Auslösung gespeichert. Unterhalb dieses Buttons erscheint in roter Schrift die Statusmeldung „OK“, wenn der Speichervorgang erfolgreich war. Der aktuelle Füllstand des Buffers wird im Fenster „Bufferfüllung“ laufend angezeigt. Mit einer Spinbox kann die Bufferlänge n eingestellt werden.

Speichern ist nur möglich, wenn das System auf „Pause“ steht.

Die bin-Datei besteht aus einem *Header* und einer Sequenz von *Records*.

Header:

Feld	Anzahl Bytes	Datentyp	Bedeutung
1	2	uint16	Header-Marke, Wert 65535
2	4	uint32	Versionsnummer der bin-Datei, gegenwärtig 2

Jeder Record enthält die Daten für ein Frame und ist wie folgt aufgebaut:

Feld	Anzahl Bytes	Datentyp	Bedeutung
1	4	uint32	Länge des Frame-Datenblocks in Bytes
2	4	uint32	Anzahl nachfolgende Shorttime features (12)
3	8	double	Grundfrequenz in Hz
4	8	double	Qualitätsmass für die Grundfrequenz
5	8	double	Vorgesehen für die Intonation, enthält gegenwärtig den Wert der Grundfrequenz
6	8	double	Mittlere Signalstärke (RMS), linear skaliert
7	8	double	Maximale Signalstärke (MAX Amp), linear skaliert
8	8	double	Stärke von Partial 1, linear skaliert
9	8	double	Stärke von Partial 2, linear skaliert
10	8	double	Stärke von Partial 3, linear skaliert
11	8	double	Spectral Centroid in Hz
12	8	double	Hi to All, linear skaliert
13	8	double	Harmonic Slope
14	8	double	Local Maxima Rate
15	4	uint32	Anzahl m der Frequenz- und Amplitudenwerte
	m x 8	double	m Frequenzwerte in Hz
	m x 8	double	m Amplitudenwerte in dB

Wir beschreiben kurz die Bedeutung dieser Felder:

Feld 3 enthält die im aktuellen Zeitfenster ermittelte *Grundfrequenz* (Hz), Feld 4 einen Parameter r , welcher die Zuverlässigkeit dieses Frequenzwertes angibt. r hat Werte zwischen 0 (kleine Zuverlässigkeit) und 1 (grosse Zuverlässigkeit).

Das Feld 5 ist vorgesehen für den Wert der *Intonation* (6.1.4.2), enthält aber gegenwärtig nur den Wert der Grundfrequenz wie Feld 3.

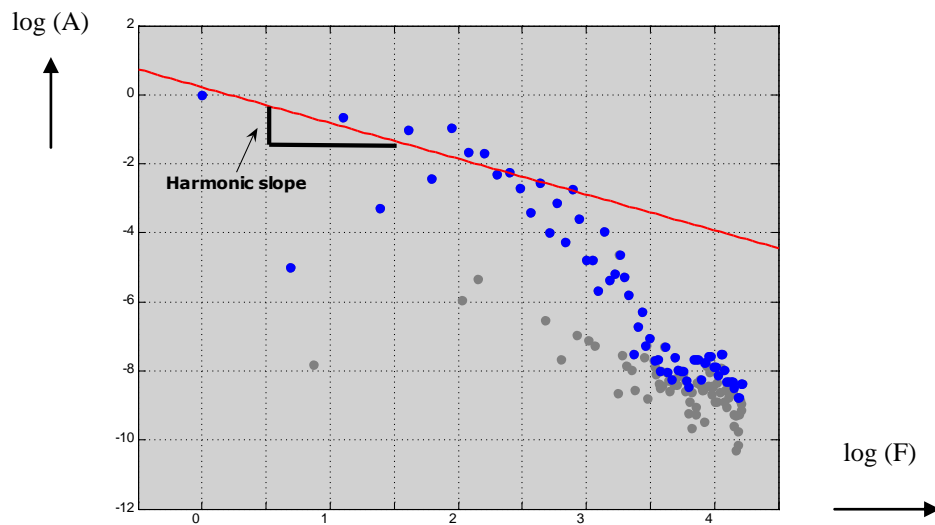
Die Felder 6 bis 10 enthalten Angaben zur *Signalstärke* (volles Signal und Teilschwingungen 1 bis 3) im aktuellen Zeitfenster, und zwar in linearer Skalierung.

Die Grössen in den Feldern 11 bis 13 sind drei Shorttime Features, welche den Obertongehalt im aktuellen Zeitfenster auf verschiedene Weise messen. Alle drei Grössen lassen sich aus den spektralen Daten (F_k, A_k) , $k=1..m$, berechnen:

a) *Spectral Centroid* = $\sum_{k=1}^m A_k F_k / \sum_{k=1}^m A_k$. Diese Grösse hat die Dimension einer Frequenz und wird in Hertz angegeben.

b) *Hi to All* = $\sum_{p=2}^{np} Ah_p / \sum_{p=1}^{np} Ah_p$. Hier bedeutet Ah_p die Amplitude der p -ten Teilschwingung und np die Anzahl der betrachteten Teilschwingungen. Ah_p wird berechnet aus den Amplitudenwerten im p -ten harmonischen Band, welches durch die Grundfrequenz und die vom Benutzer gewählte Harmonische Bandbreite (6.3) definiert ist. Man beachte, dass *Hi to All* bei der Speicherung linear skaliert angegeben wird, während die graphische Anzeige (Diagramme 4 und 5 in 6.1.4) in dB erfolgt.

c) Die *Harmonic Slope* beruht auf einer linearen Regression der Frequenz-Amplitudenpunkte, deren Frequenzkomponente nahe bei den Vielfachen der Grundfrequenz liegt, also der „harmonischen“ spektralen Daten. Dabei werden Frequenzen und Amplituden in einer logarithmischen Skala betrachtet und die tiefen Frequenzen stärker gewichtet als die hohen. Die Steigung der Regressionsgeraden (in der folgenden Abbildung rot eingezeichnet) bezeichnen wir als *Harmonic Slope*.



Liegt die *Harmonic Slope* nahe bei Null, so bedeutet dies, dass die Regressionsgerade flach verläuft – ein Indiz dafür, dass mehrere starke Oberschwingungen vorhanden sind.

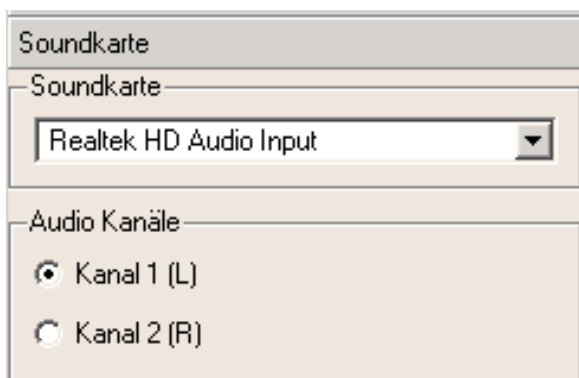
Die im Feld 14 abgelegte *Local Maxima Rate* misst das Ausmass der Schwankung des Zeitsignals im aktuellen Zeitfenster. Diese Grösse hat Werte zwischen 0 (keine Schwankung, ruhiges Signal) und 1 (starke Schwankung, unruhiges Signal).

Im Feld 15 ist die Anzahl m der im aktuellen Zeitfenster gefundenen Frequenz-Amplituden-Paare abgelegt.

Schliesslich enthalten die folgenden $2m$ Felder die Frequenz- und Amplitudenwerte F_k und A_k , wobei hier die Amplituden in dB, also logarithmisch skaliert, angegeben sind.

Erfahrene Programmierer können aufgrund dieser Beschreibung die spektralen Daten aus der bin-Datei lesen und selber weiterverarbeiten.

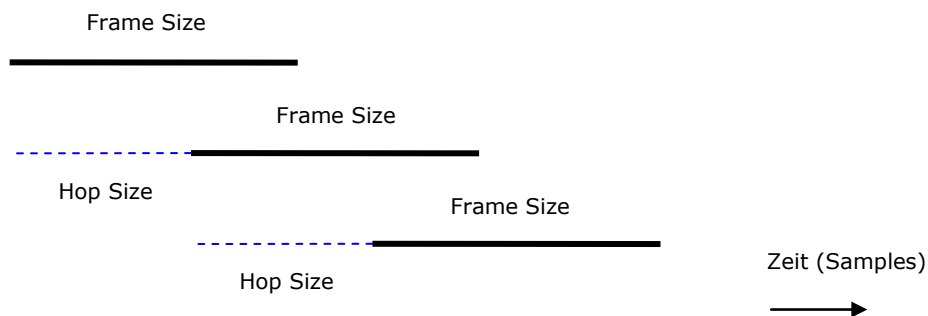
6.6 Soundkarte



In einem Computersystem können verschiedene Soundkarten installiert sein. Mit diesem Tab kann die gewünschte Soundkarte und der Audio-Kanal ausgewählt werden. Dazu muss das System auf „Pause“ stehen. Die Wahl der richtigen Signalquelle sowie das Einstellen des Eingangspegels muss im Windows-Mixer der entsprechenden Soundkarte vorgenommen werden.

6.7 Standardeinstellungen, Konfigurationsdatei

Prisma-Realtime verwendet für Tonanalyse die *Kurzzeit-Fouriertransformation*. Der von Mikrofon und Soundkarte kommende digitale Datenstrom wird in kurze Blöcke (*Frames*, Zeitfenster) einer festen Länge (*Frame Size*) zerlegt. In diesen Frames werden der Reihe nach spektrale Daten (Grundfrequenz, Frequenzen und Amplituden der Teilschwingungen) berechnet. Diese Werte sind umso genauer, je länger das Frame ist – vorausgesetzt, das Signal ändert sich während der Dauer des Frames nicht wesentlich. Andererseits kann man die zeitliche Auflösung beeinflussen, indem man die Frames überlappen lässt. Das Ausmass der Überlappung wird durch die *Hop Size* bestimmt:



In der Standardeinstellung von Prisma-Realtime wurde $\text{Frame Size} = 1025$ Samples und $\text{Hop Size} = 768$ Samples gesetzt. Diese Wahl hat dreierlei Konsequenzen:

- Die Grundfrequenzbestimmung ist wegen der obigen Wahl von Frame Size nur für Frequenzen oberhalb 130 Hz (musikalisch c_0 , d.h. kleines c) zuverlässig. Zur Analyse von tieferen Tönen muss die Frame Size vergrössert werden.
- Bei der oben gewählten Hop Size und einer Samplingfrequenz von 44100 Hz fallen alle 17.4 ms neue Daten an. Prisma-Realtime kann so "normalen" Melodien gut folgen. Zur Untersuchung von Kurzzeitphänomenen wie Einschwingvorgängen oder Artikulationen muss Hop Size verkleinert werden.
- Für eine gut ablesbare und die Grafikkarte nicht überfordernde Bildschirm-Anzeige empfehlen wir, rund 200 Bildpunkte anzuzeigen. Wenn alle 17.4 ms ein Bildpunkt gezeichnet wird, sind 3.5 s als Länge des Anzeigefensters (Display Length) optimal.

Versierte Benutzer haben die Möglichkeit, Frame Size, Hop Size, Display Length und nötigenfalls die Samplingfrequenz fs selber zu wählen und damit den Anwendungsbereich von Prisma-Realtime auszuweiten. Alle wichtigen Systemparameter (darunter auch die oben genannten) sind in der *Konfigurationsdatei* `config.xml` definiert. Diese Datei befindet sich im gleichen Ordner wie die Applikation `prisma.exe`. Sie wird von dieser beim Programmstart gelesen und bei Beendigung des Programms mit den aktuellen, im Bedienfeld vorgenommenen Skalierungseinstellungen überschrieben. Bei der Standardeinstellung sieht der Anfang der Konfigurationsdatei so aus:

```
<prismaconfig version="2" description="Standardkonfiguration">
  <global>
    <value id="display_length" type="double">3.5</value>
    <value id="fs" type="Q_UINT32">44100</value>
    <value id="frame_size" type="Q_UINT32">1025</value>
    <value id="hop_size" type="Q_UINT32">768</value>
  </global>
```

Um einen Parameterwert zu ändern, öffnet man die Konfigurationsdatei mit einem Texteditor (zum Beispiel Notepad oder Textpad), führt die Änderungen durch und speichert die Datei unter neuem Namen, z.B. `config_[spez].xml`, wobei der Zusatz `[spez]` durch eine eigene Spezifikation ersetzt werden kann, die den eingestellten Parametersatz kennzeichnet.

Sind mehrere Konfigurationsdateien vorhanden, so werden sie beim Programmstart zur Auswahl angeboten. Im Auswahlfenster erscheinen dann diejenigen Bezeichnungen, die in der ersten Zeile bei „description=“ stehen, im obigen Beispiel also „Standardkonfiguration“.

Grundsätzlich können die Werte für Frame Size, Hop Size und Display Length frei kombiniert werden. Trotzdem sind ein paar Grundsätze zu beachten:

- a) Für den Wert von Frame Size wählt man mit Vorteil Werte vom Typ $2^n + 1$ ($n=10, 11, 12, \dots$), also zum Beispiel 1025, 2049 oder 4097. Dieser Parameter ist entscheidend für die tiefste bestimmbare Grundfrequenz, sowie für die Genauigkeit der Grundfrequenzbestimmung.
- b) Hop Size muss kleiner sein als Frame Size. In der Standardkonfiguration überlappen sich die Frames um ein Viertel.
- c) Bei Kombinationen von grosser Frame Size, kleiner Hop Size und grosser Display Length besteht die Gefahr, dass der Computer an Kapazitätsgrenzen stösst und das Programm abstürzt.

Die folgende Tabelle zeigt Beispiele bewährter Konfigurationen für drei Anwendungsfälle (Samplingfrequenz 44100 Hz):

Bezeichnung	Frame Size (Samples)	Tiefste Grundfrequenz	Hop Size (Samples)	Zeit pro Bildpunkt	Display Length
Standard	1025	130 Hz (c)	768	17.4 ms	3.5 s
Kurzzeitanalyse	1025	130 Hz (c)	384	8.7 ms	1.7 s
Tiefe Töne	2049	92 Hz (FIS)	1536	34.8 ms	7.0 s

Drei Konfigurationsfiles mit diesen Parametersätzen sind im Download-Paket von Prisma-Realtime enthalten.

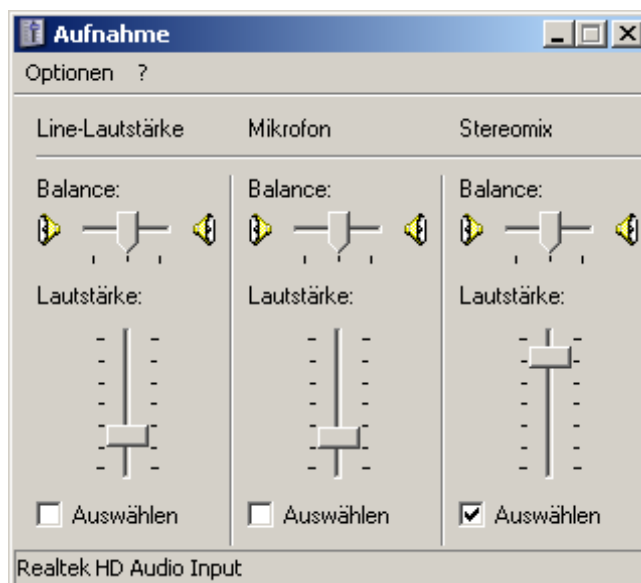
7 Tipps und Tricks

7.1 Grafik

- Wenn der Prozessor zu 100% ausgelastet ist und die Daten ruckartig vorbeifließen, kann es sein, dass der „Vertical Sync“ bei der Grafikkarte ausgeschaltet ist. Unter den OpenGL-Einstellungen der Grafikkarte muss der Wert „Vertical Sync“ (kann auch anders heißen - gemeint ist das Synchronisieren auf die Bildwiederholfrequenz des Bildschirms) so eingestellt werden, dass die Applikation den Modus selbst wählen kann.
- Besonders bei Notebooks kann es vorkommen, dass das Powermanagement die Leistung von CPU und Grafikkarte drosselt. Damit Prisma-Realtime richtig läuft, müssen die Leistungswerte auf maximal gestellt werden.
- Prisma-Realtime synchronisiert sich am Bildschirmaufbau. Wenn die Bildwiederholfrequenz zu hoch eingestellt ist (besonders bei CRT-Bildschirmen) kann es zu Ruckeln oder Stottern kommen. Eine gute Wahl für Flachbildschirme ist 60Hz.
- Bei schwacher Leistung des Computers kann es hilfreich sein, nur ein einziges Diagramm anzuzeigen. Es müssen so weniger Datenpunkte gezeichnet werden.

7.2 Sound

- Prisma-Realtime bietet nur die Möglichkeit zur Auswahl der Soundkarte. Es ist jedoch nicht möglich, direkt im Programm ein bestimmtes Eingangssignal zu wählen. Diese Einstellungen müssen im Windows-Mixer der entsprechenden Soundkarte vorgenommen werden. Dort lässt sich auch der Eingangspegel einstellen.
- Alle während der Entwicklung getesteten Soundkarten bieten die Möglichkeit, das Ausgangssignal, also vom PC abgespielte Töne, direkt aufzunehmen. Das unten abgebildete Fenster erscheint, wenn man im Windows Start-Menü → Einstellungen → Systemsteuerung: Sounds und Audiogeräte: Audio → Soundaufnahme → Lautstärke wählt. Will man einen mit einem Mikrofon aufgenommenen Ton untersuchen, so wählt man die Checkbox „Auswählen“ in der Spalte „Mikrofon“. Möchte man eine wav-Datei untersuchen, so spielt man sie in einem Media-Player ab und markiert die Checkbox „Auswählen“ in der Spalte „Stereomix“. Mit diesem Trick kann man mit Prisma-Realtime auch ohne Mikrofon Audiodaten untersuchen.



- Wird das Eingangssignal übersteuert, so ist das Signal verzerrt und die Prisma-Analyse wertlos. Die Aussteuerung lässt sich mit dem Level-Meter von Prisma-Realtime einfach überprüfen: Der Maximalwert muss kleiner als 0dB und der RMS-Wert kleiner als -3dB sein.

8 Support

Prisma-Realtime wird als Produkt einer Forschungsarbeit kostenlos zur Verfügung gestellt. Die Entwickler übernehmen keinerlei Haftung für das Produkt und irgendwelche Folgen seiner Benützung. Sie geben auch keine Garantie für das korrekte Funktionieren der Software.

An Verbesserungen und Erweiterungen wird weiter gearbeitet. Es ist geplant, in unregelmässigen Abständen neue Versionen von Prisma-Realtime auf der Projekt-Website zum Download anzubieten.

Fragen zu Prisma-Realtime können über das Kontaktformular auf

www.prisma-music.ch

an die Entwickler gestellt werden. Wir werden uns bemühen, darauf zu antworten, bitten jedoch um Verständnis, wenn dies vielleicht etwas Zeit braucht.