

Similarity Search auf Audio Daten

Aufgabensteller:

Betreuer:

Bearbeiter:

Prof. Hans-Peter Kriegel

Peter Kunath, Christian Mahrt

Markus Feil (www.tone2.com)

Gliederung

- Motivation
- Beat-Histogramm (FFT, Autokorrelation, RFNMA)
- Pitch Analyse (verbesserte Autokorrelation, RFNMA)
- Neue Erkenntnisse
- Ausblick

Motivation

Anwendungsbeispiele:

- Ähnlichkeitssuche in Musikstücken
- Automatisierte Klassifizierung nach Genres
- Bestimmung von Doubletten
- Wer klaut von wem die Melodie/Samples?
- Konvertierung von Audiodaten zurück in Noten
- Extraktion von Highlevel-Features – semantische Bedeutung?

Beat-Histogramm

Algorithmus:

Schritt 1: FFT Transformation

Parameter:

Fenstergröße = 2048 Samples

Kaiser Fenster mit beta 5

50% overlap

Beat-Histogramm

Schritt 2: Logarithmieren des Spektrums

Sinn: Gleichverteilung von Energie pro Oktave

Parameter:

12 Halbtöne pro Oktave

20x oversampling

Beat-Histogramm

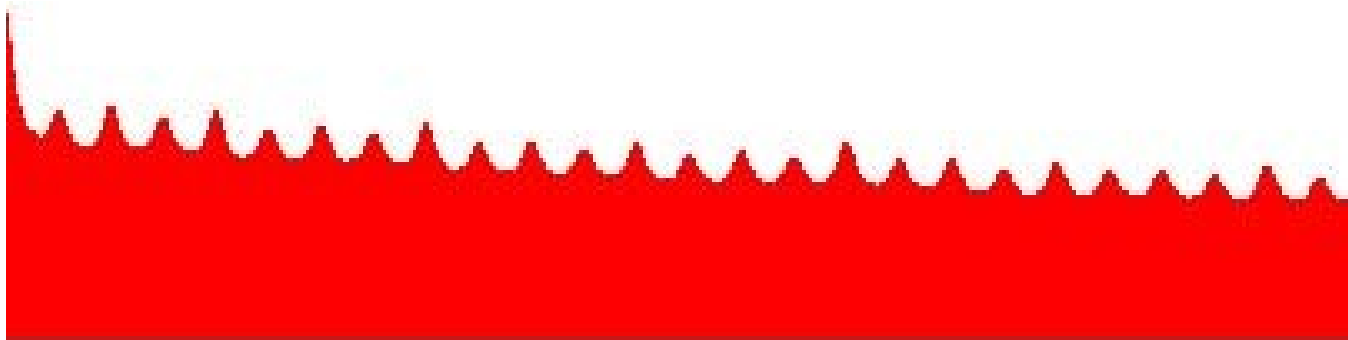


Logarithmiertes FFT Spektrum

Beat-Histogramm

Schritt 3: Autokorrelation des Bildes

Sinn: Verschiebungen, die ähnliche Bildbereiche erfassen resultieren in einem Maximum

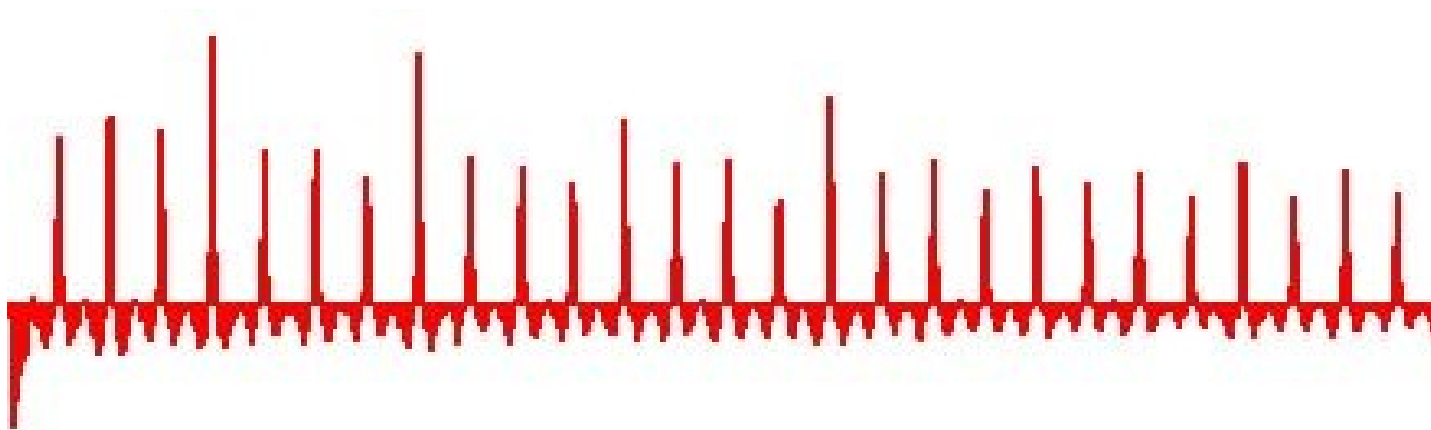


Problem: Maxima auch bei Halbe/Drittel/Viertel
Geschwindigkeit

Beat-Histogramm

Schritt 4: Extraktion der lokalen Maxima

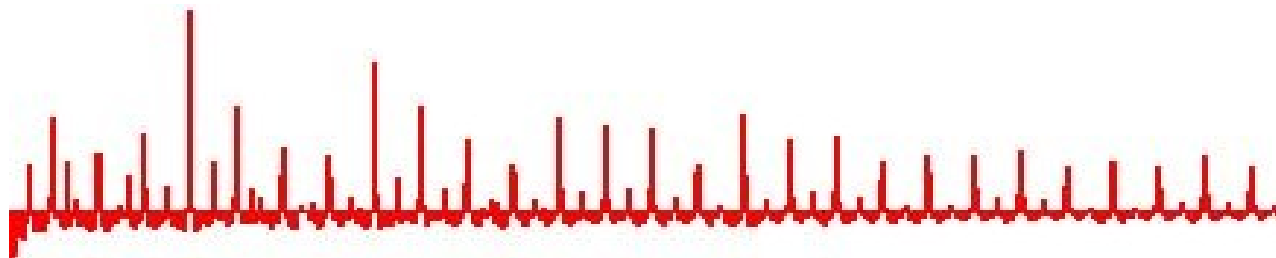
Verfahrensweise: 2. Ableitung oder FIR Hochpass 3. Ordnung über Funktion bilden



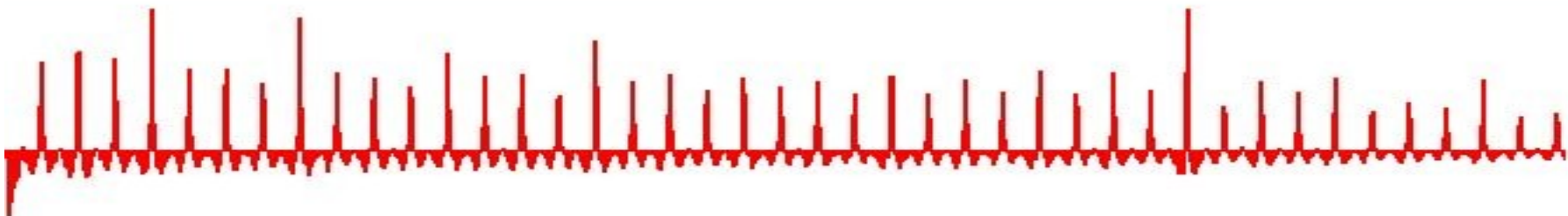
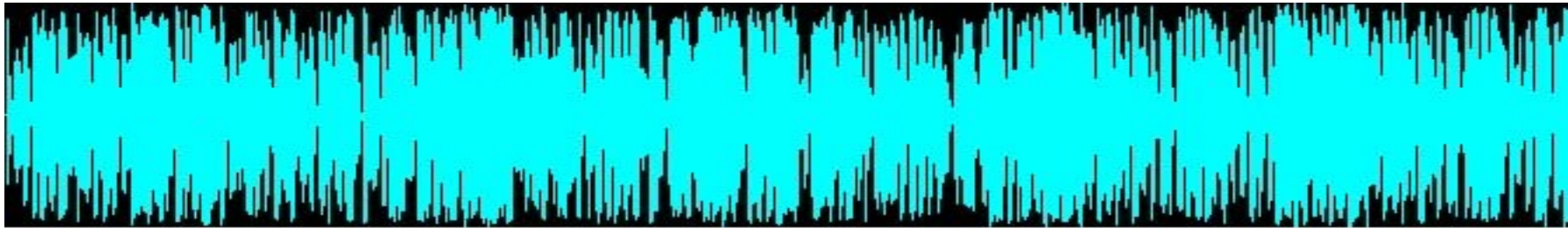
Beat-Histogramm

Schritt 5: Revers-Filterung von n-fach Maximas der Autokorrelationsfunktion (RFNMA)

Kreuzkorrelation (Faltung) mit Kammfunktion



Beat-Histogramm



Beat-Histogramm

Ergebnisse:

Das logarithmierte FFT Spektrum bringt gegenüber einer FIR Filterbank (CQT) etwa 90% Performance Vorteil, jedoch um 10% schlechtere Datenqualität.

Erkenntnisse:

Durch die RFNMA konnte das Problem der mehrfach Maxima bei der Autokorrelation gelöst werden!

Pitch-Analyse

Ton = Grundton + Obertöne

Akkord = mehrere Töne, die ein harmonisches Frequenzverhältnis haben

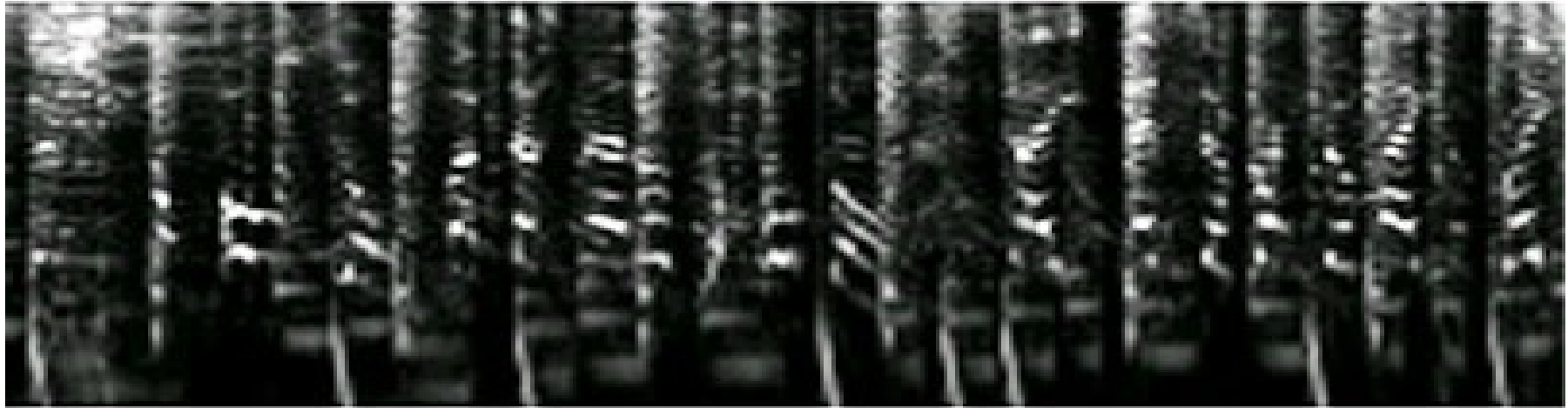
Probleme: Phasenauslöschungen, Obertöne, Schwebungen, Geräusche (z.B. Rhythmus Instrumente)

Ziel der Pitch-Analyse ist es, einen als Sample Daten vorgegebenen Audio-Stream in seine ursprünglichen Noten zurückzurechnen.

Pitch-Analyse

Methoden 1 (FFT):

- FFT (Fenstergröße 2048, Kaiser beta 5, 50% overlap)
- Logarithmieren der FFT (12 Halbtöne pro Oktave)
- Obertonkamm auf Bildfunktion



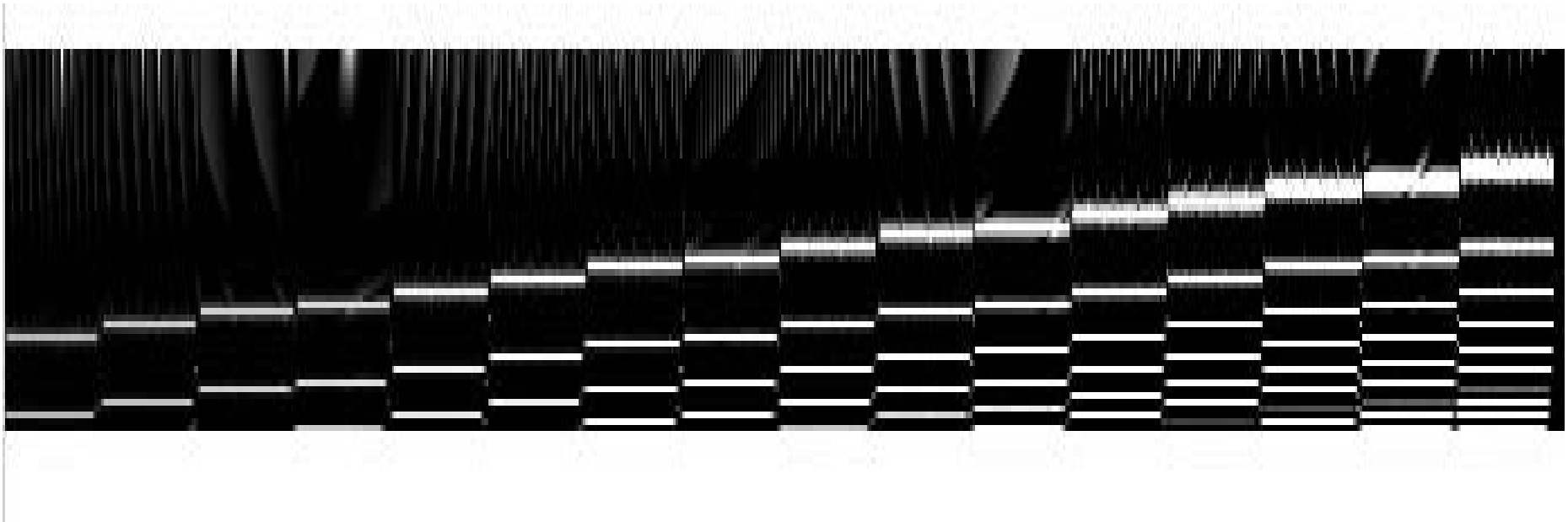
Pitch-Analyse

Probleme:

- Zunehmende Unschärfe bei niederen Frequenzen
- Phasen/Datenverlust durch FFT Methodik „power = $\sqrt{\text{re}^2 + \text{im}^2}$ “

Aufgrund des ernüchternden Ergebnisses wurde die FFT Methodik für die Pitch Analyse verworfen.

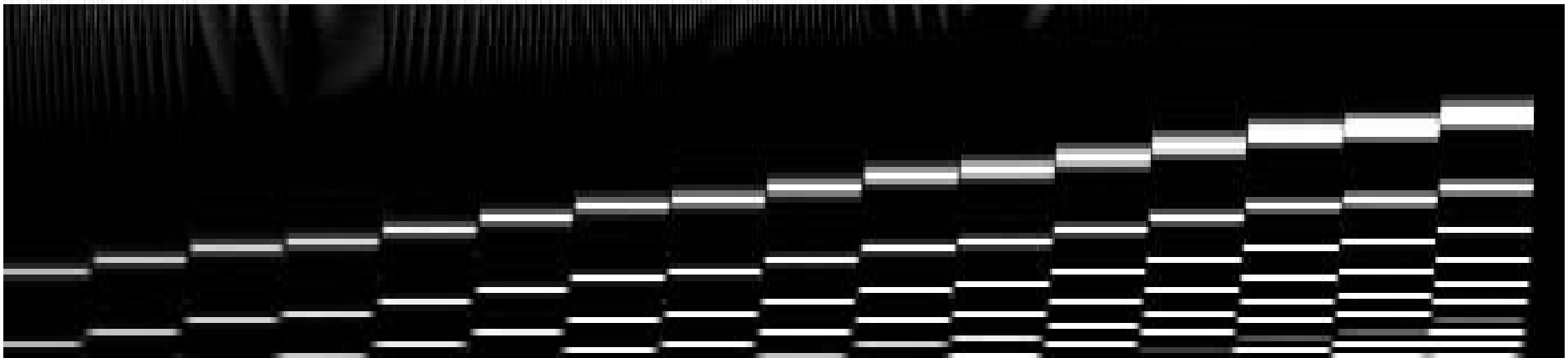
Pitch-Analyse



Tonleiter mit Sägezahnschwingung mit klassischer Autokorrelationsfunktion

Pitch-Analyse

Verbesserung:
Anwendung einer Fensterfunktion

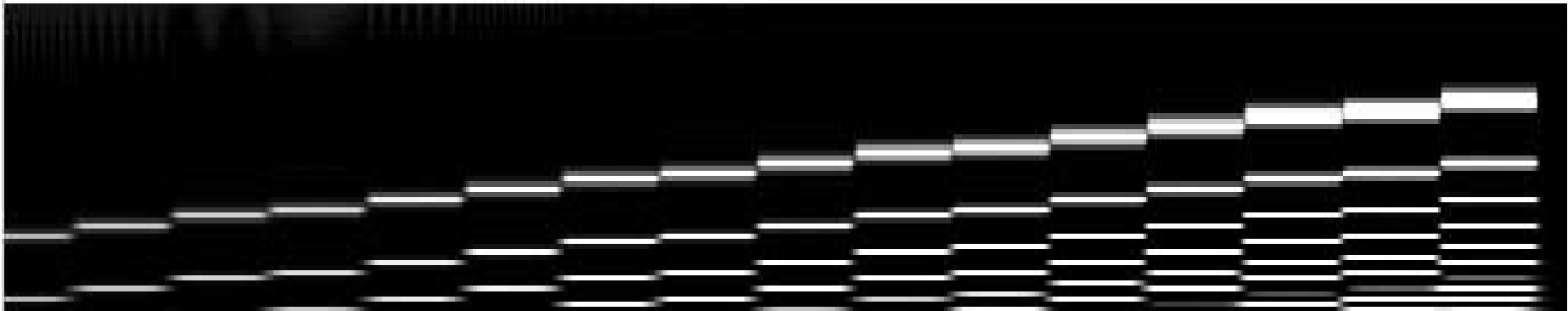


Tonleiter mit Sägezahnschwingung mit gefensterter Autokorrelationsfunktion (Kaiser beta 5, 4 Zyklen)

Pitch-Analyse

Verbesserung 2:

Offsetkorrektur und vorgeschaltetes Hochpassfilter

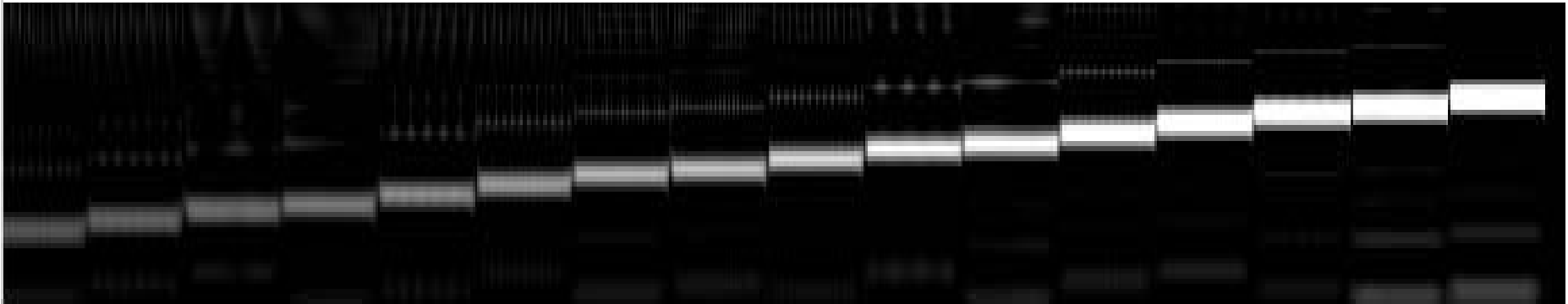


Tonleiter mit Sägezahnschwingung mit gefensterter Autokorrelationsfunktion (Kaiser beta 5, 8 Zyklen), Offsetkorrektur und Hochpass (6dB, Cutoff 0.1)

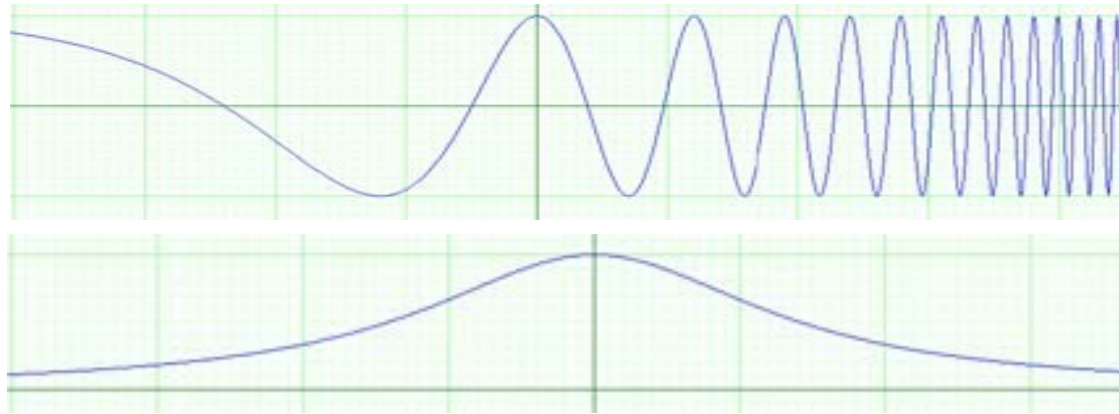
Pitch-Analyse

Verbesserung 3:

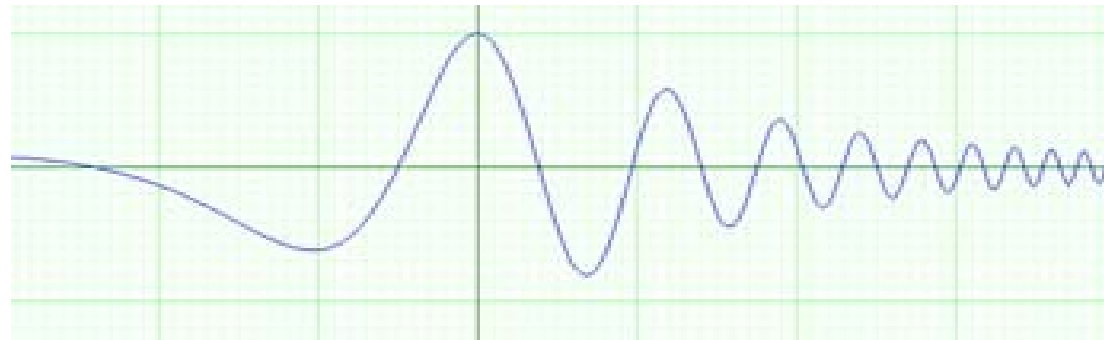
Revers-Filterung von n-fach Maximas der Autokorrelationsfunktion (RFNMA) = Kreuzkorrelation (Faltung) mit Kammfunktion



Pitch-Analyse



```
a=cos((1.05946^x-1)*2*pi)  
y = 1/(x*x/200+1)*a
```

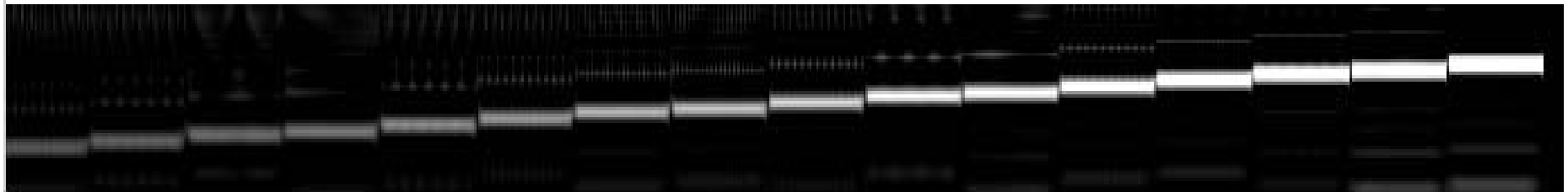


Pitch-Analyse

Vergleich (Sägezahn Tonleiter):



Autokorrelation



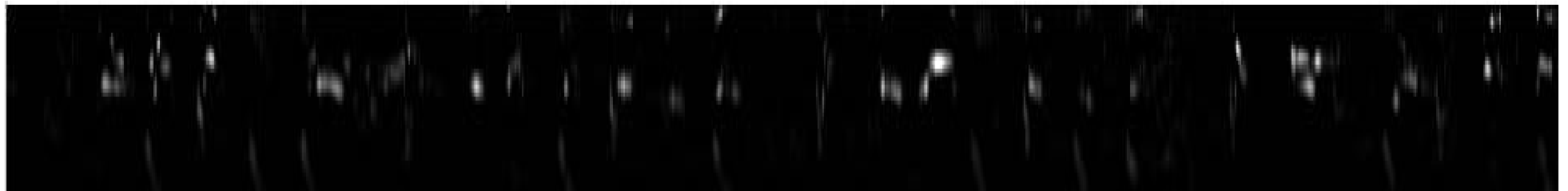
RFNMA

Pitch-Analyse

Vergleich (The Power of Love - Huey Lewis & The News)



Autokorrelation



RFNMA

Pitch-Analyse

Ergebnisse:

Die oben vorgestellte Methodik der Pitch-Analyse macht es möglich, Wahrscheinlichkeiten für einzelne Noten zu gegebenen Zeitpunkten in einem Musikstück zu bestimmen.

Der Algorithmus ist robust für Schwebungen, Phasenauslöschungen, Akkorde und Gesang.

Erkenntnisse

- Durch Anwendung eines Hochpassfilters (erste Ableitung?) vor der Berechnung der Autokorrelation kann das Ergebnis erheblich verbessert werden.
- Durch die Anwendung einer Fensterfunktion und einer Faltung, die sich über mehrere Wellenzyklen (2x,4x,8x) erstreckt, kann das Ergebnis der Autokorrelation zusätzlich deutlich verbessert werden.

Erkenntnisse

- Durch die Anwendung einer Kreuzkorrelation mit Kammfunktion auf das Ergebnis der Autokorrelation kann das Problem der mehrfach Maxima gelöst werden (RFNMA).
- Mithilfe der neu entwickelten RFNMA Transformation ($O(n) = n^3$) lassen sich die Wahrscheinlichkeiten der Grundfrequenzen beliebiger komplexer Wellenformen aus 1-dimensionalen Datenstreams bestimmen.

Ausblick

- Detaillierte Ausarbeitung und Evaluation der RFNMA
- Resynthese der Pitch Analyse
- QFIRB (Quadraturische FIR Filterbank) als Alternative zu FFT, CQT und Autokorrelation