

Parametrische Audiocodierung bei sehr niedrigen Datenraten

Heiko Purnhagen

Laboratorium für Informationstechnologie
Universität Hannover

Gliederung

- Einleitung
- Quellenmodelle für Audiosignale
- Parametrische Audiocodierung
- MPEG-4 Parametric Audio Coder (HILN)
- Ausblick

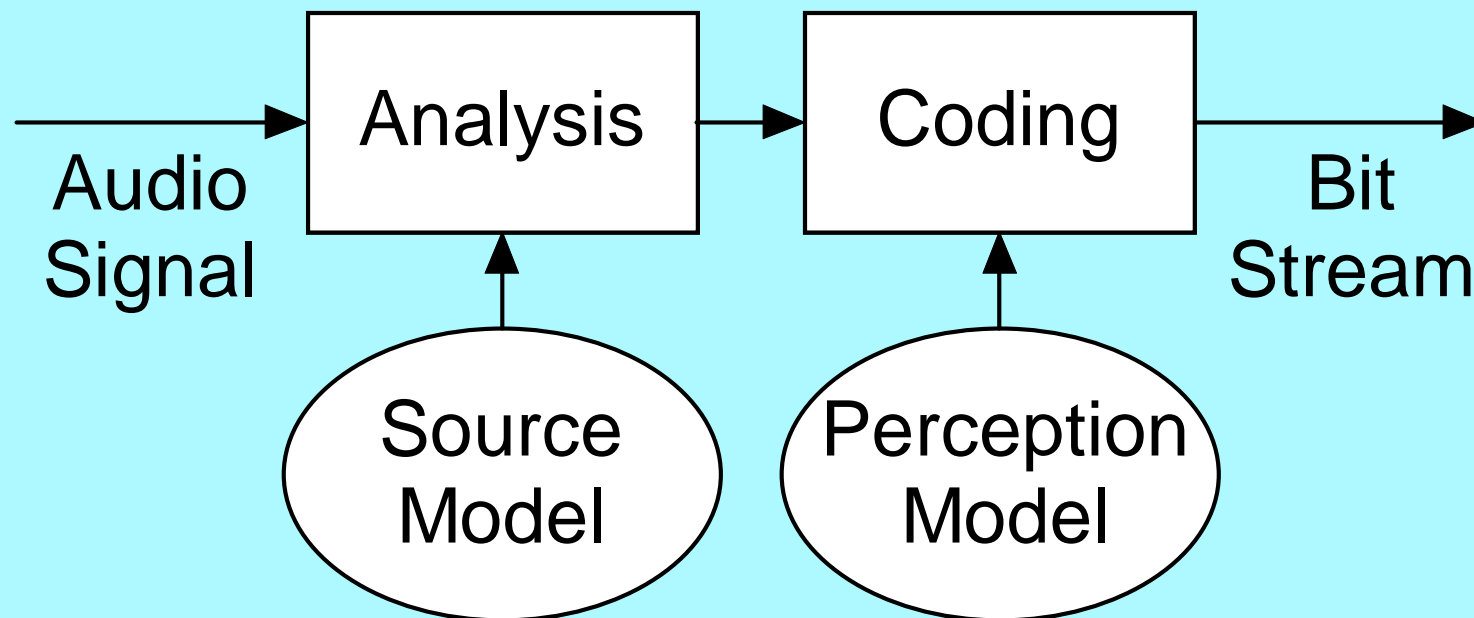
Einleitung

- Darstellung eines Audiosignals x
 - *physikalisch*: Wellenform $x(t)$
 - *abstrakt*: Partitur (kompakt, mehrdeutig)

=> vielversprechender Ansatz für Audiocodierung
- Encodierung: *phys. Darst.* => *abstr. Darst.*
aber: automatische Transkription sehr schwierig !!!
- Audiocodierung:
*Kompakte Darstellung eines Audiosignals
automatisch aus realen Signalen abgeleitet*

Einleitung

- Signaldarstellung in der Audiocodierung:
Quellenmodell + Modellparameter
=> Parametrische Audiocodierung



Quellenmodelle für Audiosignale

Ansätze für Audio-Quellenmodelle:

- Spektrale Zerlegung
 - stationäres Signal innerhalb Zeitrahmen (Dauer T)
=> Zeit/Frequenz-Transformation (z.B. MDCT)
 - signaladaptive Zeit/Frequenz-Auflösung
- Physikalische Modellierung: Erregung + Resonanzen
 - Sprache: periodische/zufällige Erregung + LPC
 - Musiksynthese: z.B. "waveguide"

Quellenmodelle für Audiosignale

- Modellierung mit Sinustönen

$$\hat{x}(t) = \sum_{i=1}^N a_i(t) \cdot \sin\left(\varphi_i + 2\pi \int_0^t f_i(\tau) d\tau\right)$$

- Anwendungen: Musikinstrument-Analyse/Synth.
Sprach- und Audiocodierung
- Nachbildung spektraler "Spitzen"
- Verfolgen von Trajektorien / Phasenkontinuität
- Phase φ_i oft nicht wahrnehmungsrelevant

Quellenmodelle für Audiosignale

- Transienten-Modelle
 - Sinustöne mit Amplituden-Hüllkurve (attack&decay)
 - Sinuston-Modellierung des DCT-Spektrums
 - Transformations-Codierung (T/F)
- Rausch-Modelle
 - MA-Modell: DCT des Rauschspektrums
 - AR-Modell: weißes Rauschen + LPC Filter
 - Bark-Band oder ERB Rauschmodell
 - "Bark-warped" LPC

Quellenmodelle für Audiosignale

- Erweiterte Sinuston-Modelle
 - Sinustöne mit gemeinsamer Grundfrequenz
=> harmonischer Ton
 - "bandwidth enhanced sinusoids":
Sinuston => Schmalbandrauschen (AM/FM-Mod.)
- => Problem: Wahl des Quellenmodells ?

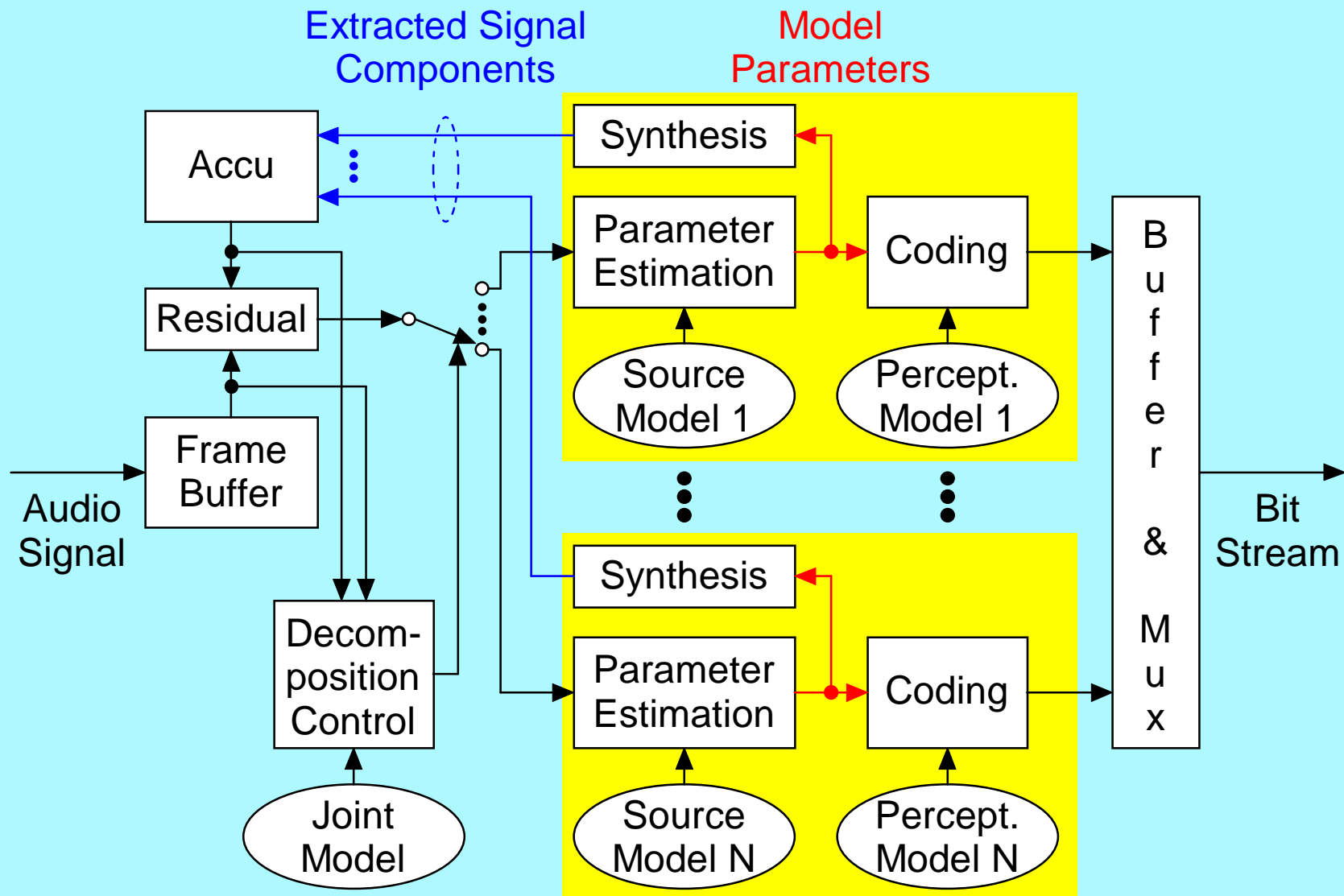
Effizienz vs. Generalität

(spezialisierte Quellenmodelle
nicht für beliebige Signale geeignet)

Parametrische Audiocodierung

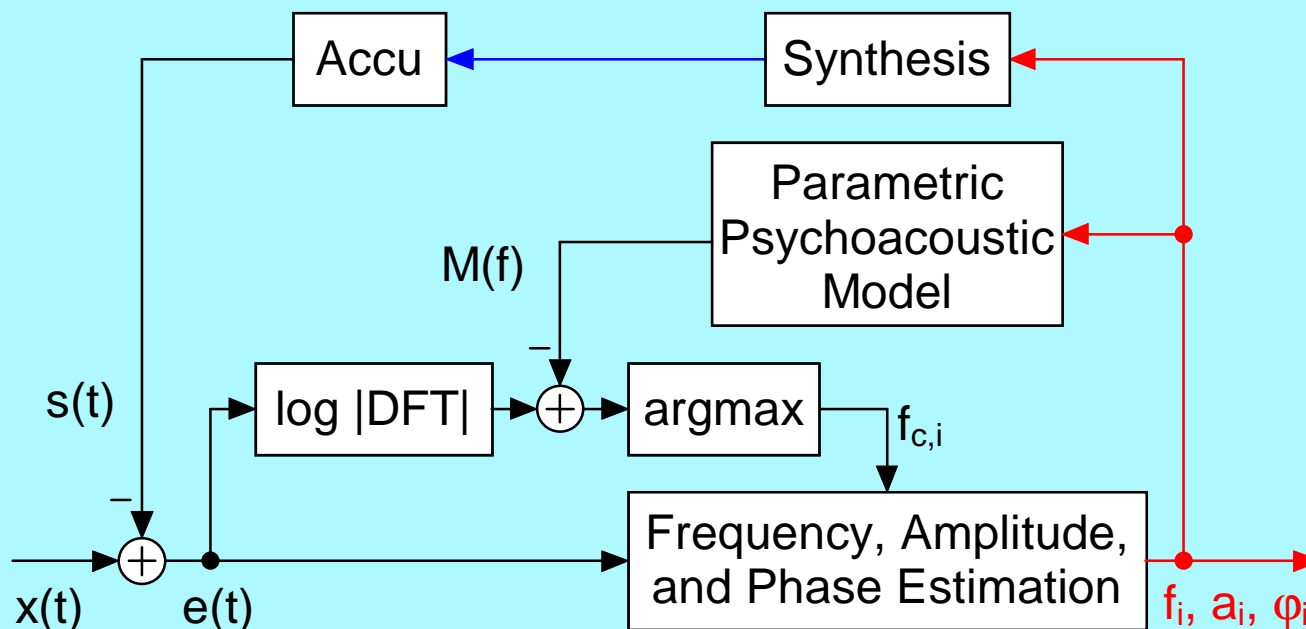
- Parametrische Audiocodierung
 - Kombination von verschiedenen *Quellenmodellen*
=> Zerlegung des Audiosignals in Komponenten
 - Berücksichtigung von *Wahrnehmungsmodellen*
=> "optimale" Zerlegung
(wichtigste Komponenten auswählen)
 - => Analyse/Synthese-Ansatz
- Parameter-Quantisierung und -Codierung
 - Quant.-Stufengröße: "just noticeable differences"
 - Entropiecodierung

Parametrische Audiocodierung: Encoder



Parametrische Audiocodierung: Encoder

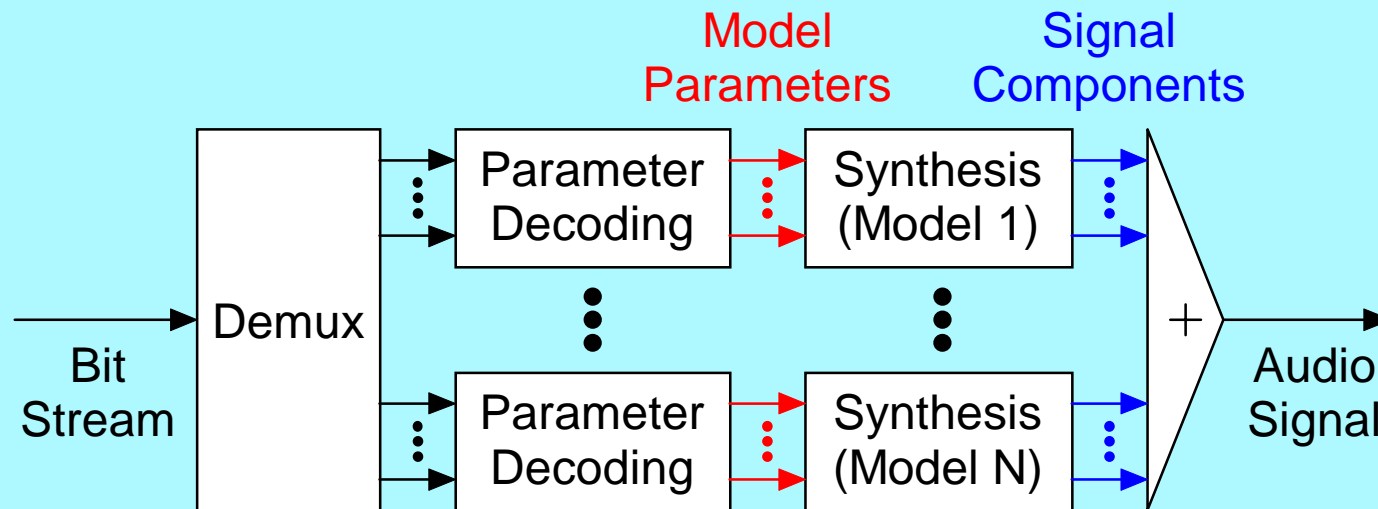
- Beispiel:
Zerlegung in Sinuston-Komponenten



Analyse/Synthese-Schleife

Parametrische Audiocodierung: Decoder

● Parameterdecodierung & Signalsynthese



● Zusätzliche Funktionalitäten

- Skalierbarkeit: Basis- + Ergänzungs-Bitstrom
- Signal-Modifikation: Geschwindigkeit & Tonhöhe

MPEG-4 Parametric Audio Coder

- MPEG-4 Audio Version 2: HILN

"Harmonic and Individual Lines plus Noise"

harmon. Ton: Grundfrequenz & LPC-Spektrum

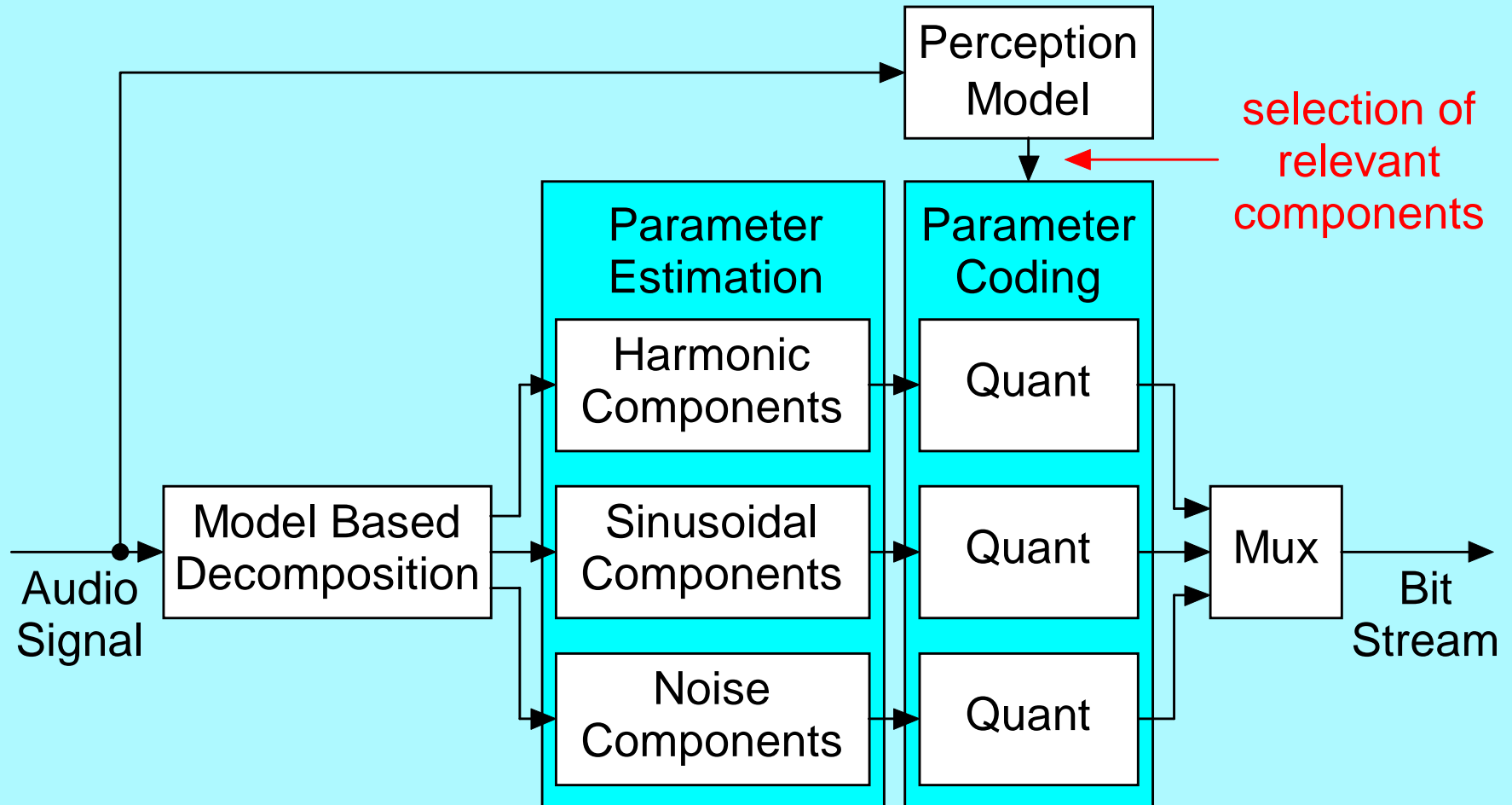
Sinustöne: Frequenz & Amplitude
[opt.: Ampl.-Hüllkurve, Startphase]

Rauschen: LPC-Spektrum

Zeitraumen: 32ms (typ.)

- => 4 .. 16 kbit/s @ 8 kHz Bandbreite (typ.)
 - Skalierbare Datenrate & Signal-Modifikation

MPEG-4 Parametric Audio Coder

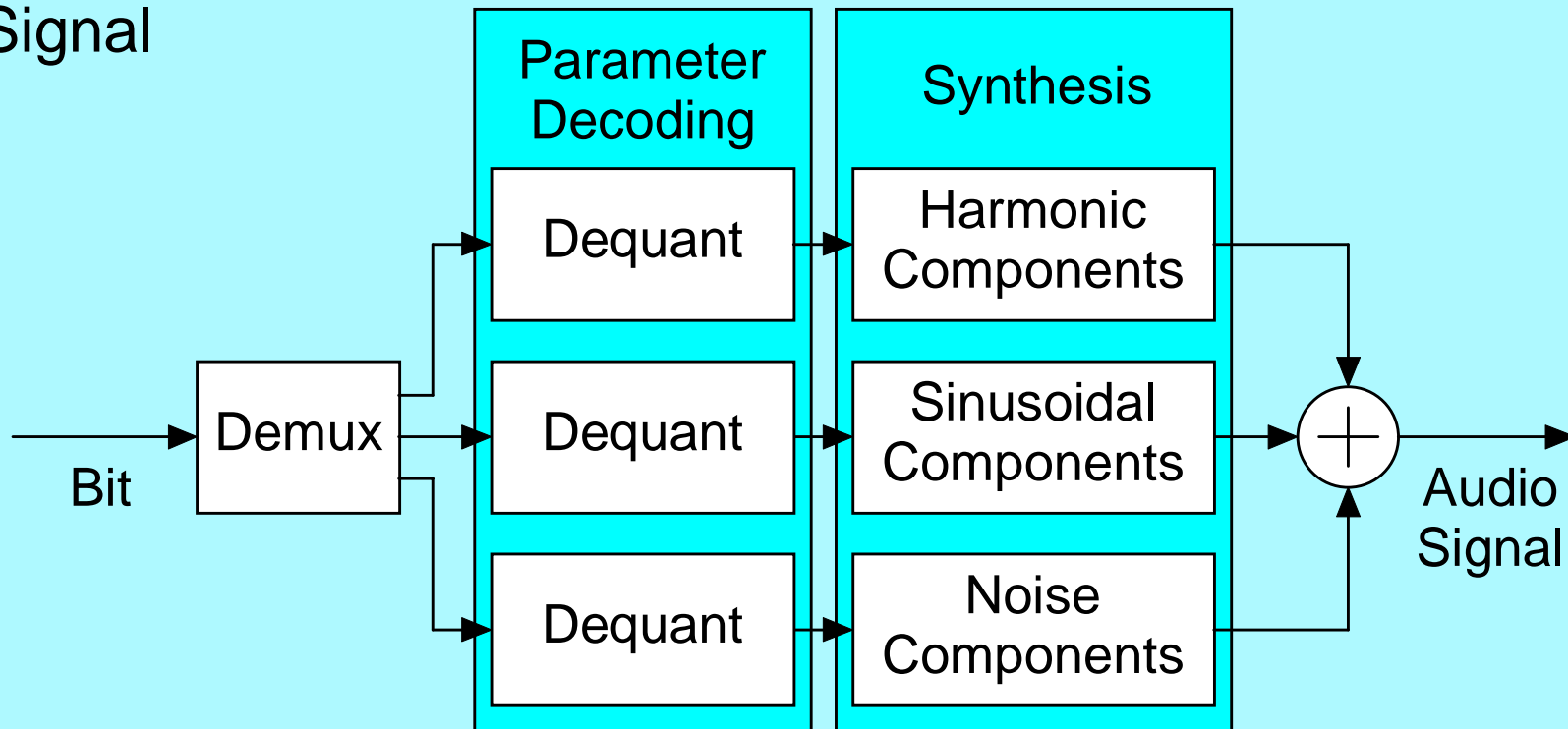


Parametric Audio Encoder (HILN)

MPEG-4 Parametric Audio Coder

- **Demo:** HILN 16 kHz mono @ 6 kbit/s

Original
Signal



Parametric Audio Decoder (HILN)

MPEG-4 Parametric Audio Coder

- **Demo:** HILN 16 kHz mono @ 6 kbit/s
- **Signal-Modifikation:**
 - Tonhöhen-Änderung +20%
 - Geschwindigkeits-Änderung +20%
- **Datenraten-Skalierbarkeit**
 - Basis-Bitstrom 6 kbit/s
 - Basis- + Enhancement-Bitstrom 6+10 kbit/s
 - nicht-skalierbarer Bitstrom 16 kbit/s

MPEG-4 Parametric Audio Coder

- **Demo:** HILN 16 kHz mono @ 6 kbit/s
- Vergleich mit anderen Codierungsverfahren:
 - Original (8 kHz Bandbreite)
 - Sprachcoder (MPEG-4 CELP) 6 kbit/s
 - T/F-Coder (MPEG-4 TwinVQ) 6 kbit/s
 - Parametrischer Audiocoder (HILN) 6 kbit/s

Ausblick

- Verbesserung der Quellenmodelle
 - z.B. "Bark-warped" LPC für Rauschen
- Verbesserung der Signal-Zerlegung (Encoder)
 - Trennung von tonalen/rauschartigen Komponenten
 - automatische Segmentierung von Sprache/Musik (Audio-Objekte sind transparent ...)
- "General Audio Coding" [Scheirer 1999]
 - Quellenmodell als Bestandteil des Bitstroms

Literatur

- *Official* MPEG Home Page

<http://www.cseit.it/mpeg/>

- MPEG Audio Web Page

<http://www.tnt.uni-hannover.de/project/mpeg/audio/>

- Parametrische Audiocodierung

<http://www.tnt.uni-hannover.de/~purnhage/>