Audio Artifacts in Perceptual Audio Coding

Speaker: Chi-Min Liu,


http://psplab.csie.nctu.edu.tw
Why Artifacts Investigation

Are you satisfied with the audio coding technologies like MP3, AAC, HE-AAC, PS, ...
Outline

- Introduction
- Common Artifacts
- Artifacts in Temporal Noise Shaping
- Artifacts in SBR
- Artifacts in Parametric Stereo Coding
- Artifact Summary
## Introduction

### Audio Technologies & Artifacts

<table>
<thead>
<tr>
<th>Analog Audio</th>
<th>Digital Audio</th>
<th>MP3</th>
</tr>
</thead>
<tbody>
<tr>
<td>WOW&amp;Flutter, Tape Saturation,</td>
<td>Aliasing,                     Pre-echo,</td>
<td></td>
</tr>
<tr>
<td>Crosstalk,...</td>
<td>Quantization Nonlinearity, Group Delay,</td>
<td>Birdies,...</td>
</tr>
<tr>
<td>AAC</td>
<td>Bandlimited,...</td>
<td></td>
</tr>
<tr>
<td>SBR, PS</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>
Outline

- Introduction
- Common Artifacts
- Artifacts in Temporal Noise Shaping
- Artifacts in SBR
- Artifacts in Parametric Stereo Coding
- Artifact Summary
Common Artifacts

**Band-limited Effect**

- Original Spectrum
- Compressed Spectrum with Narrow Bandwidth
- Compressed Spectrum with **High Frequency Reconstruction**
Common Artifacts (c.1)

- **Birdie Effect**

- Original Spectrum
- Compressed Spectrum with Zero Band
- Compressed Spectrum with **Zero Band Dithering**
Common Artifacts (c.2)

Audio Patch Method on HE-AAC Decoder

Audio Patch Method on HE-AAC Decoder

HE-AAC Signal with Zero Bands in LF and Band-limited HF

HE-AAC Signal with Audio Patch Method
Common Artifacts (c.3)

- **Pre-echo Effect**

Pre-echo Phenomenon

Post-masking, Pre-masking and Simultaneous Masking
Outline

- Introduction
- Common Artifacts
  - Artifacts in Temporal Noise Shaping
  - Artifacts in SBR
  - Artifacts in Parametric Stereo Coding
- Artifact Summary
Artifacts in Temporal Noise Shaping

**TNS Overview**

\[ r[k] = \sum_{j=1}^{n} h_j \cdot r[k - j] + q[k] \]

- **Analysis Part**
  - Linear Predictor \( H \)
  - Quantizer \( Q \)

- **Synthesis Part**
  - De-Quantizer \( Q^{-1} \)
  - Linear Predictor \( H \)

Open-loop prediction coding scheme in TNS

**Reconstruction Error**:

\[ r[k] = x[k] - \hat{x}[k] \]

**Quantization Error**:

\[ R[z] = \frac{Q[z]}{1 - \sum_{j=1}^{n} h_j z^{-j}} = \frac{Q[z]}{H[z]} \]
Artifacts in Temporal Noise Shaping (c.1)

TNS Overview (c.1)

Original Signal in Time Domain

Coded Signal without TNS

Coded Signal with TNS
Artifacts in Temporal Noise Shaping (c.2)

- **Noise Amplification around Attack**

![Graph showing time response for a single pole with different values of r (0.7, 0.9, 1) and amplitude values ranging from -30000 to 30000.](image)

- **Time Response for a Single Pole**

- **Noise with TNS**
Artifacts in Temporal Noise Shaping (c.3)

- **Time-Domain Aliasing**

Original Signal in Time Domain

Coded Signal without TNS

Coded Signal with TNS

Pre-aliasing Artifact

Post-aliasing Artifact
Decompositions of MDCT and IMDCT
Artifacts in Temporal Noise Shaping (c.4)

Noise from High-Order Prediction Filter

- Quantization noise without TNS
- Quantization noise with order 3
- Quantization noise with order 12
Outline

- Introduction
- Common Artifacts
- Artifacts in Temporal Noise Shaping
- Artifacts in SBR
- Artifacts in Parametric Stereo Coding
- Artifact Summary
Artifacts in SBR

**SBR Overview**

- Original Spectrum
- Decoded AAC LF Spectrum
- HF Generation by SBR
- HF Adjustment by SBR
- Tone Compensation
Artifacts in SBR (c.1)

- **SBR Overview (c.1)**
  - Time/Frequency Grid
    - Envelope adjustment unit

Choose one frequency table

Determine the time borders

Decide the resolution of time envelopes

```
\begin{array}{c|c|c|c}
  & \text{High} & \text{Low} & \text{High} \\
\hline
  F & \text{High} & \text{Low} & \text{High} \\
\hline
\end{array}
```

SBR Frame
Artifacts in SBR (c.2)

- **SBR Overview (c.2)**

---

**Diagram of SBR Encoder**

**Diagram of SBR Decoder**
Artifacts in SBR (c.3)

Tone Trembling Effect

Normal Spectrogram  Abnormal Spectrogram

22 kHz  0 kHz

Trembling Tones in HF

sound like trembling, or sparkling
Original Signal

Decoded Signal with Tone Trembling Effect
Artifacts in SBR (c.4)

Tone Trembling Effect (c.1)

- Patching Module
  - Find corresponding relation between high bands and replicated low bands

Frequency Tables ➔ Patch Algorithm ➔ Corresponding Low Bands

SBR Range ➔ Original High Bands
Artifacts in SBR (c.5)

- **Tone Trembling Effect (c.2)**

\[
\hat{s}[n] = A[n] \exp\left(i(\omega[n] \cdot n + \Theta)\right)
\]

\(\omega[n]\) \(: Center Frequency

Viewed as frequency modulation signal
Artifacts in SBR (c.6)

**Tone Shift Effect**
- Inherent artifact of direct band replication
- Hard to detect for human hearing

![Original Stereo Signal vs. Decoded Signal](image-url)
Artifacts in SBR (c.7)

• Tone Shift Effect (c.1)

Nero Digital Audio (1.0.0.2)

Original Signal vs. Decoded Signal

Coding Technologies  aacPlus v2 (7.0.5)
Artifacts in SBR (c.8)

- Noise Overflow

Causes
- Tone loosing
- Interpolation mode
Original Signal

Decoded Signal with Noise Overflow
Artifacts in SBR (c.9)

Noise-Floor Overflow (c.1)

Original Signal vs. Decoded Signal

Nero Digital Audio (1.0.0.2)
Artifacts in SBR (c.10)

Noise-Floor Overflow (c.2)

Original Signal vs. Decoded Signal

Coding Technologies  aacPlus v2 (7.0.5)
Artifacts in SBR (c.11)

- **Noise-Floor Overflow (c.3)**
  - Interpolation mode

**Envelope Adjustment**

- Energies of Original HF Bands in a Grid
- Energies of Replicated LF Bands in a Grid
- Adjusted Energies of Replicated LF Bands at Interpolation Mode
- Adjusted Energies of Replicated LF Bands at Non-Interpolation Mode

- Flat Envelope
- Keep Original Envelope
Artifacts in SBR (c.12)

- Noise-Floor Overflow (c.4)

Noise-floor Overflow without Tone Compensation at Interpolation Mode

Noise-floor Overflow with Tone Compensation at Interpolation Mode
Artifacts in SBR (c.13)

**Tonal Spike Effect**
- Due to false alarm of tonality detection
- Overestimated tonal energy
- Underestimated noise energy
Artifacts in SBR (c.14)

- Tonal Spike Effect (c.1)

Nero Digital Audio (1.0.0.2)

Coding Technologies  aacPlus v2 (7.0.5)
Artifacts in SBR (c.15)

- Sawtooth

Original Spectrum

Sawtooth Spectrum
Artifacts in SBR (c.16)

- **Sawtooth (c.1)**
  - **Limiter Gain** Mechanism
    - Upper bound of gain value
    - Avoid excess noise substitution

Decoded Spectrum with Limited Gain Mechanism

Decoded Spectrum without Limited Gain Mechanism
Artifacts in SBR (c.17)

**Beat Effect**

- **Original spectrum containing two tones with long distance.**
- **Original waveform**

- **AAC**
- **SBR**

- **Original spectrum containing two tones with short distance.**

- **Beat Effect**
Artifacts in SBR (c.18)

Beat Effect (c.1)

Shaping of Cosine Function with low frequency

\[ x(t) = \sin(\omega_1 t) + \sin(\omega_2 t + \phi) \]

\[ = 2 \cos\left(\frac{\omega_2 - \omega_1}{2} \cdot t + \frac{\phi}{2}\right) \sin\left(\frac{\omega_2 + \omega_1}{2} \cdot t + \frac{\phi}{2}\right) \]
Artifacts in SBR (c.19)

Beat Effect (c.2)

Coding Technologies  aacPlus v2 (7.0.5)

Original Waveform

Beat Effect

AAC  SBR
Outline

- Introduction
- Common Artifacts
- Artifacts in Temporal Noise Shaping
- Artifacts in SBR
- Artifacts in Parametric Stereo Coding
- Artifact Summary
Artifacts in Parametric Stereo Coding

**PSC Overview**
Artifacts in Parametric Stereo Coding (c.1)

PSC Overview (c.1)

PS in MPEG-4 HE-AAC Version 2 Encoder
Artifacts in Parametric Stereo Coding (c.2)

- **Downmix by Average Method**
  
  \[ d[n] = \frac{l[n] + r[n]}{2} \]

- **Energy cancellation problem**

**Diagram:**

- **Original Binaural Signal**
- **Extracted Monaural Signal**
- **Signal Vanishing Effect**
Artifacts in Parametric Stereo Coding (c.3)

- **Downmix by KLT Method**
  
  \[ d[n] = \lambda_1 \cdot l[n] + \lambda_2 \cdot R[n] \]

- Best energy compactness

Original Binaural Signal

Extracted Monaural Signal

Average Method

KLT Method
Artifacts in Parametric Stereo Coding (c.4)

- **Downmix by KLT Method (c.1)**
  - Inherent disadvantage
    - Weaker component discard
    - Variable combination coefficient
  - Artifact
    - Tone leakage effect
    - Tone modulation effect
Artifacts in Parametric Stereo Coding (c.5)

**Type-I Tone Leakage Effect**
- One tone in some channel leaks to another channel.
- Inherent artifact of all downmix method

![Linear-scaled Spectrum of Original Stereo Signal](image1)

![Linear-scaled Spectrum of Reconstructed Stereo Signal by Average Method](image2)
Artifacts in Parametric Stereo Coding (c.6)

- Type-I Tone Leakage Effect (c.1)

Linear-scaled Spectrum of Original Stereo Signal

Linear-scaled Spectrum of Reconstructed Stereo Signal

Nero Digital Audio (1.0.0.2)
Artifacts in Parametric Stereo Coding (c.7)

- **Type-I Tone Leakage Effect (c.2)**

Linear-scaled Spectrum of Original Stereo Signal

Linear-scaled Spectrum of Reconstructed Stereo Signal

Coding Technologies  aacPlus v2 (7.0.5)
Artifacts in Parametric Stereo Coding (c.8)

- **Type-II Tone Leakage Effect**
  - Weaker tone is easily suppressed.

![Linear-scaled Spectrum of Original Stereo Signal](image1)

![Linear-scaled Spectrum of Reconstructed Stereo Signal by KLT Method](image2)
Artifacts in Parametric Stereo Coding (c.9)

- Type-II Tone Leakage Effect (c.1)

![Graph showing Linear-scaled Spectrum of Original Stereo Signal and Linear-scaled Spectrum of Reconstructed Stereo Signal]

Nero Digital Audio (1.0.0.2)
Artifacts in Parametric Stereo Coding (c.10)

- **Type-II Tone Leakage Effect (c.2)**

![Graphs showing linear-scaled spectrum of original and reconstructed stereo signals, comparing Type-I and Type-II artifacts.](image)

Coding Technologies: aacPlus v2 (7.0.5)
Artifacts in Parametric Stereo Coding (c.11)

- Tone Modulation Effect
- Adaptive coefficient vectors between frames

Red: Original Signal  Blue: Decoded Signal by KLT Method
Artifacts in Parametric Stereo Coding (c.12)

- Tone Modulation Effect (c.1)
  - Downmix subband signal
    \[ d[n] = \lambda_1[n] \exp(i\theta_1(n))l[n] + \lambda_2[n] \exp(i\theta_2(n))r[n] \]
  - \( \lambda_k[n] \exp(i\theta_k[n]) \) causes modulations in both amplitude and phase.
  - Example
    \[ s[n] = A \exp(i(\omega n + \Theta)) \]
    \[ \hat{s}[n] = (A \cdot \lambda[n]) \exp(i(\omega n + \Theta + \theta[n])) \]
Artifacts in Parametric Stereo Coding (c.13)

- Tone Modulation Effect (c.2)
  - Cosine smooth connection of coefficient vector

\[ \Psi[n] = \frac{\gamma_i - \gamma_{i+1}}{2} \cdot \cos \left( \frac{\pi n}{k} \right) + \frac{\gamma_i + \gamma_{i+1}}{2} \]
Artifacts in Parametric Stereo Coding (c.13)

- Tone Modulation Effect

Original Signal
Artifacts in Parametric Stereo Coding (c.14)

- Tone Modulation Effect

PS Signal under KLT Method
Artifacts in Parametric Stereo Coding (c.15)

- **Tone Modulation Effect**

PS Signal under KLT Method with Smooth Connection
Outline

- Introduction
- Common Artifacts
- Artifacts in Temporal Noise Shaping
- Artifacts in SBR
- Artifacts in Parametric Stereo Coding
- Artifact Summary
## Artifact Summary

### Common Artifact

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Perceptual</th>
<th>Generation Sources</th>
<th>Relief Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band Limited Effect</td>
<td>Muffled</td>
<td>(1) Reduction of sampling rate</td>
<td>High frequency reconstruction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Bit rate constraint</td>
<td></td>
</tr>
<tr>
<td>Birdie Effect</td>
<td>Fishy</td>
<td>(1) Unsuitable bit allocation policies</td>
<td>Zero band dithering</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Excessive masking energy estimation</td>
<td></td>
</tr>
<tr>
<td>Pre-echo</td>
<td>Annoying noise</td>
<td>(1) Transient signal</td>
<td>(1) Window switch</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Inappropriate size of coding block</td>
<td>(2) TNS</td>
</tr>
</tbody>
</table>
### TNS Artifact

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Perceptual</th>
<th>Generation Sources</th>
<th>Relief Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise Amplification</td>
<td>Not Sensitive</td>
<td>TNS</td>
<td>Window switch</td>
</tr>
<tr>
<td>Time-Domain Aliasing</td>
<td>Annoying noise</td>
<td>Shaping effect of the inverse filter $I[\zeta]$ in TNS</td>
<td>A joint method by TNS and window switch</td>
</tr>
<tr>
<td>Noise from High-Order Prediction Filter</td>
<td>Annoying noise</td>
<td>High-Order Prediction Filter in TNS</td>
<td></td>
</tr>
</tbody>
</table>
## Artifact Summary (c.2)

### SBR Artifact

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Perceptual</th>
<th>Generation Sources</th>
<th>Relief Methods</th>
</tr>
</thead>
</table>
| Tone Trembling     | Trembling       | (1) Tone-rich signal in SBR  
(2) Adaptive frequency table and SBR range in SBR | Fixed tables and SBR range       |
| Tone Shift         | Not sensitive   | (1) Harmonic signal in SBR  
(2) SBR replication in SBR |                                  |
| Noise Overflow     | Dull and Noisy  | (1) Tone losing in T/F Grid in SBR  
(2) Envelope adjustment with interpolation mode in SBR | (1) Noise floor correction  
(2) Non-interpolation mode |
| Tonal Spike        | Metallic        | (1) False alarm of tone detection in SBR  
(2) Overestimation of tonal component in SBR |                                  |
| Sawtooth           | Not Sensitive   | Limiter gain mechanism in SBR | Limiter gain turns off          |
## Artifact Summary (c.3)

### PS Artifact

<table>
<thead>
<tr>
<th>Artifacts</th>
<th>Perceptual</th>
<th>Generation Sources</th>
<th>Relief Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tone Leakage Type 1</td>
<td>Blurred Spatial position</td>
<td>Down-mixing procedure in PS</td>
<td></td>
</tr>
<tr>
<td>Tone Leakage Type 2</td>
<td>Blurred Spatial position</td>
<td>Down-mixing procedure of the KLT in PS</td>
<td>Energy normalize</td>
</tr>
<tr>
<td>Tone Modulation</td>
<td>Click</td>
<td>Down-mixing procedure of the KLT in PS</td>
<td>Coefficient smooth</td>
</tr>
</tbody>
</table>
Why Artifacts Investigation

- J. S. Mill "It is better to be Socrates dissatisfied than a pig satisfied."
- Are you satisfied with the audio coding technologies like MP3, AAC, HE-AAC, PS, …
- Can we Improved?
Q & A

http://psplab.csie.nctu.edu.tw