Optimal Codec for Wireless Communication Links

W. Bastiaan Kleijn and Minyue Li
School of Electrical Engineering
KTH - Royal Institute of Technology
Stockholm, Sweden
HEARCOM Context

services:
- recognition server
- announcements
- telephone
- VoIP
- radio/television
- theater/church
- audio

WiFi
Bluetooth
3G
Internet

BAN
acoustics

32 kb/s
256 kb/s

Personal Communication Hub

HA

HA
Objectives

• Audio coder for hearing instruments
  – Real-time scalable
  – Delay < 1 ms; facilitates inter-ear communication
  – Sampling-rate 16 kHz
  – “Design” Rate 32 kb/s
  – Low complexity (processor dependent)
  – Quality better than ITU-T G.722 at 64 kb/s
    • Delay 1.625 ms.
Design Philosophy: *Model Everything*

- Statistical models of
  - Source
  - Channel
  - Receiver
  - Encoder
  - Decoder

- Estimate / optimize in real time
Real-Time Design Method Based on High-Rate Theory

- **Rate-distortion theory** (Shannon, 1959)
  - Needs densities; bounds for simple densities only
  - Variable-rate only
  - No direct relation to practical systems

- **Lloyd algorithm** (Lloyd, 1958)
  - Not a model; leads directly to quantizer
  - Iterative / results in codebooks / not scalable
  - Locally optimal / no need density function

- **High-rate theory** (Bennett, 1948)
  - Assumes signal density constant in quantization cell
  - Asymptotically optimal
  - Fixed and variable rate
  - Analytic solutions / scalable
  - Provides centroid density / requires additional step
Coder Design Choices

• Signal modeling
  – Predictive coding
  – Transform coding

• Adaptation of the model
  – Backward adaptation
  – Forward adaptation

• Rate or distortion variation
  – Constrained-entropy coding
  – Constrained-resolution coding
Choices: Prediction vs Transform

- 3-bit/dimension constrained-resolution quantizer
Reverse Waterfilling

- Code only where needed

![Diagram showing frequency/eigenfunction index with spectrum optimal reconstruction highlighted and don't spend bits here noted.](image)
Choices: Modeling

- DCT-based coding: unitary transform
  - Does not affect space-filling
  - Reverse water filling
  - Imperfect decorrelation for fixed transform
  - Long delay

- CELP (analysis-by-synthesis AR coder)
  - AR model functions well
  - Inefficient space filling
  - No inherent reverse water filling (requires post-filter)
  - Nightmare for adaptive coding
Modeling: Forward Adaptive

1. segment
2. weight
3. subtract ringing
4. KLT
5. quantize
6. entropy code

- estimates AR model
- estimates perception
- accounts for inter-block correlations
- quantizes
- entropy code

Slide 10
Choices: Adaptation

• **Forward adaptation**
  – Estimate model parameters on original signal
  – Transmit model parameters
  – Requires long blocks

• **Backward adaptation**
  – Estimate model parameters on reconstructed signal
  – No transmission required
  – Frequent update natural -> **short delay**
  – Estimate not accurate after transition
HEARCOM Coder Architecture

- segment
- weight
- subtract ringing
- KLT
- quantize
- entropy code

- estimate perception
- estimate AR model
- estimate KLT
Constrained-Entropy Coding

• Constrained-resolution coding
  – Fixed rate
  – Outliers in distortion

• Constrained-entropy coding
  – “Fixed” distortion
  – Outliers in rate

leads to error feedback in backward adaptation

No error feedback in backward adaptation!
Illustration: Backward Adapt and CE

- Constrained entropy = fixed distortion
- Same average rate
Constrained-Entropy Coding

• Constrained resolution coding
  – Fixed rate
  – Outliers in distortion

• Constrained entropy coding
  – “Fixed” distortion
  – Outliers in rate

leads to error feedback in backward adaptation

No error feedback in backward adaptation!

generally very bad rate variation
Removal Rate Outliers

Diagram:
- Segment
- Weight
- Subtract ringing
- KLT
- Quantize
- Entropy code
- Estimate perception
- Estimate KLT
- AR model codebook
- Mixture selection
- AR model estimation
Improved Backward Model Estimation

1. segment
2. weight
3. subtract ringing
4. KLT
5. quantize
6. entropy code
7. estimate KLT
8. estimate perception
9. AR model codebook
10. mixture selection
11. AR model estimation
HEARCOM Coder Architecture

• KLT modeling:
  – Better than CELP
  – Better than MDCT

• Backward-adaptive, variable rate
  – Very short delay
  – Constant quality
  – No feedback of distortion
  – Limited rate variation
  – Enhanced model estimation
Comparative Quality and Rate

- MUSHRA (MUltiple Stimuli with Hidden Reference and Anchor) test
- HEARCOM coder is *scalable*!

<table>
<thead>
<tr>
<th>coder</th>
<th>Rate (kb/s)</th>
<th>Delay (ms)</th>
<th>MUSHRA score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITU-T G.722</td>
<td>64</td>
<td>1.625</td>
<td>69.6</td>
</tr>
<tr>
<td>ITU-T-G.722</td>
<td>48</td>
<td>1.625</td>
<td>56.9</td>
</tr>
<tr>
<td>HEARCOM</td>
<td>32*</td>
<td>0.625</td>
<td>71.3</td>
</tr>
</tbody>
</table>
Conclusions

• Architecture determines effectiveness low-D coding
  – CELP not naturally scalable
  – MDCT has long delay, imperfect decorrelation
  – KLT-based architecture performs best

• HEARCOM coder attributes:
  – Scalable: can redesign coder in real time at any time
  – Given quality: reduces rate of G.722 by factor two
  – Very low delay (0.625 ms)
  – Low computational complexity
  – Architecture can also be used for forward-adaptive approach
    • Facilitates multiple-description coding
  – Coder is programmed in C