Perceptual Distortion-Rate Optimization of Long Term Prediction in MPEG AAC

Tejaswi Nanjundaswamy, Vinay Melkote, Emmanuel Ravelli and Kenneth Rose

Signal Compression Lab
Department of ECE
UCSB

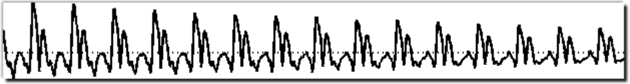
November 7, 2010

Outline

Outline

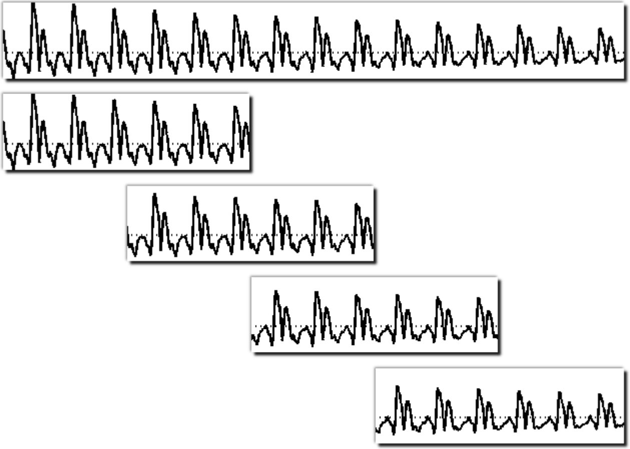
Audio Coding

- Audio Signal



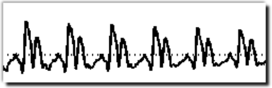
Audio Coding

- Signal divided into overlapping frames



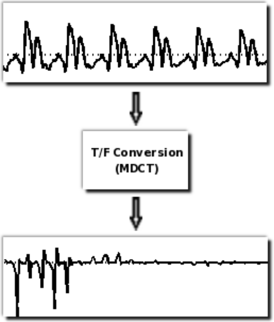
Audio Coding

- Frame transformed



Audio Coding

- Frame transformed



Audio Coding

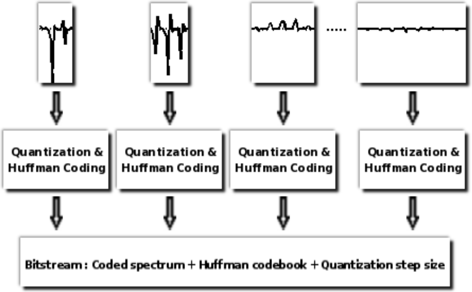
- Transform coefficients split into bands

Audio Coding

- Transform coefficients split into bands

Audio Coding

- Bands quantized and coded to generate bitstream



Audio Coding

- Coding problem definition: Achieve minimum perceptual distortion at a given rate
- Perceptual?
  - Based on content, human brain can tolerate (or mask) variable amount of noise in each band
  - Captured in distortion measure as Maximum Noise to Mask Ratio (MNMR)

$$\text{MNMR} = \max_{\forall \text{ bands}} \frac{\text{Quantization noise energy}}{\text{Masking threshold}}$$

  - Masking threshold estimated via psycho-acoustic analysis of input frame

Audio Coding

- Coding problem definition: Achieve minimum perceptual distortion at a given rate
- Perceptual?
  - Based on content, human brain can tolerate (or mask) variable amount of noise in each band
  - Captured in distortion measure as Maximum Noise to Mask Ratio (MNMR)

$$\text{MNMR} = \max_{\forall \text{ bands}} \frac{\text{Quantization noise energy}}{\text{Masking threshold}}$$

  - Masking threshold estimated via psycho-acoustic analysis of input frame

Audio Coding

- Coding problem definition: Achieve minimum perceptual distortion at a given rate
- Perceptual?
  - Based on content, human brain can tolerate (or mask) variable amount of noise in each band
  - Captured in distortion measure as Maximum Noise to Mask Ratio (MNMR)

$$\text{MNMR} = \max_{\forall \text{ bands}} \frac{\text{Quantization noise energy}}{\text{Masking threshold}}$$

  - Masking threshold estimated via psycho-acoustic analysis of input frame

Audio Coding

- Coding problem definition: Achieve minimum perceptual distortion at a given rate
- Perceptual?
  - Based on content, human brain can tolerate (or mask) variable amount of noise in each band
  - Captured in distortion measure as Maximum Noise to Mask Ratio (MNMR)

$$\text{MNMR} = \max_{\forall \text{ bands}} \frac{\text{Quantization noise energy}}{\text{Masking threshold}}$$

  - Masking threshold estimated via psycho-acoustic analysis of input frame

Audio Coding

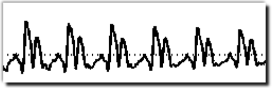
- Coding problem definition: Achieve minimum perceptual distortion at a given rate
- Perceptual?
  - Based on content, human brain can tolerate (or mask) variable amount of noise in each band
  - Captured in distortion measure as Maximum Noise to Mask Ratio (MNMR)

$$\text{MNMR} = \max_{\forall\ \text{bands}} \frac{\text{Quantization noise energy}}{\text{Masking threshold}}$$

  - Masking threshold estimated via psycho-acoustic analysis of input frame

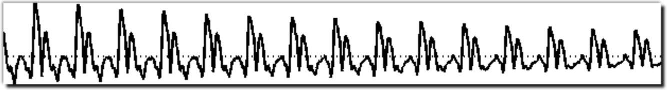
Audio Coding

- MDCT exploits redundancies within the current frame



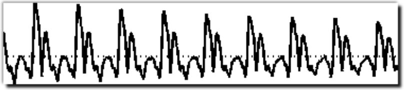
Audio Coding

- MDCT exploits redundancies within the current frame
- Audio signal has a repeating pattern



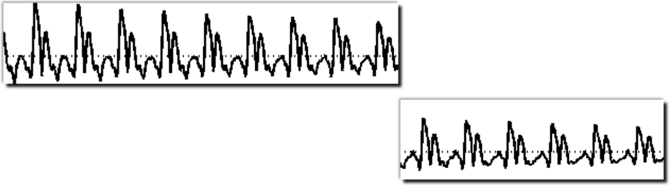
Audio Coding

- MDCT exploits redundancies within the current frame
- Audio signal has a repeating pattern
- Previously reconstructed data available at decoder



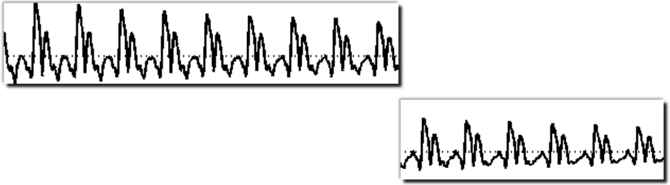
Audio Coding

- MDCT exploits redundancies within the current frame
- Audio signal has a repeating pattern
- Previously reconstructed data available at decoder
- Can we exploit this correlation?



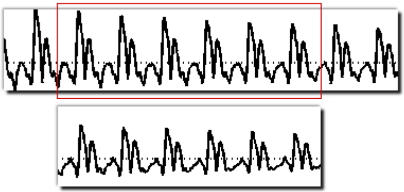
MPEG AAC LTP

- MPEG AAC uses the Long Term Prediction (LTP) tool



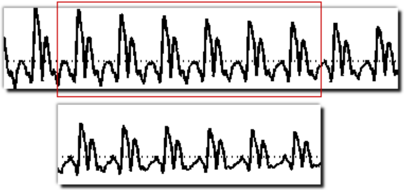
MPEG AAC LTP

- MPEG AAC uses the Long Term Prediction (LTP) tool
- Predicts current frame from history



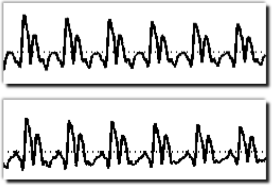
MPEG AAC LTP

- MPEG AAC uses the Long Term Prediction (LTP) tool
- Predicts current frame from history
- Reference position indicated via lag index



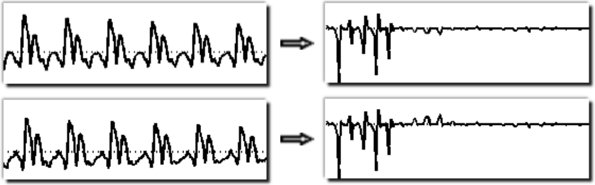
MPEG AAC LTP

- MPEG AAC uses the Long Term Prediction (LTP) tool
- Predicts current frame from history
- Reference position indicated via lag index
- Waveforms matched via gain factor



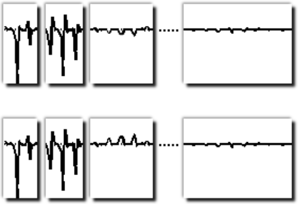
MPEG AAC LTP

- MPEG AAC uses the Long Term Prediction (LTP) tool
- Predicts current frame from history
- Reference position indicated via lag index
- Waveforms matched via gain factor
- Transformed



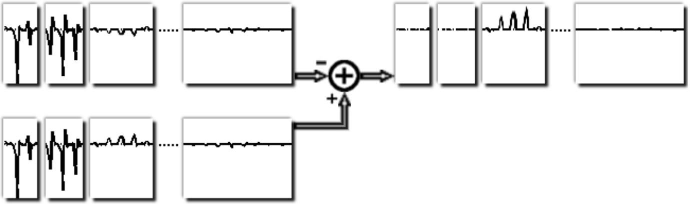
MPEG AAC LTP

- Transformed coefficients split into bands



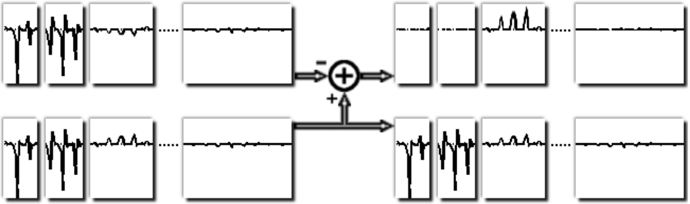
MPEG AAC LTP

- Transformed coefficients split into bands
- Prediction residue generated



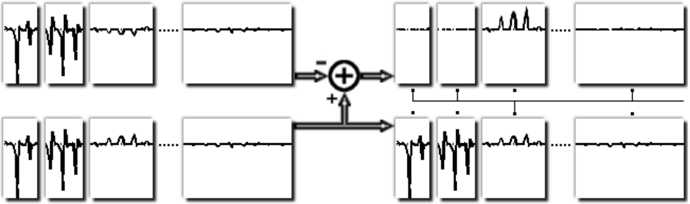
MPEG AAC LTP

- Transformed coefficients split into bands
- Prediction residue generated
- Compared with original



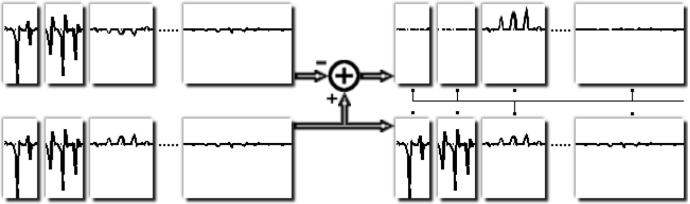
MPEG AAC LTP

- Transformed coefficients split into bands
- Prediction residue generated
- Compared with original
- Per band LTP flag set



MPEG AAC LTP

- Transformed coefficients split into bands
- Prediction residue generated
- Compared with original
- Per band LTP flag set
- Per frame flag indicates if LTP is used at all in current frame



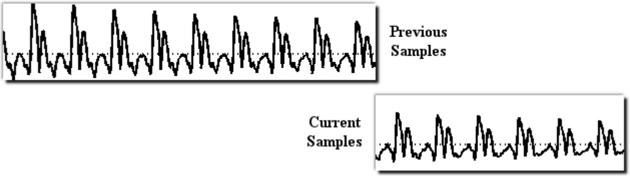
MPEG AAC LTP

- The overall LTP parameter set includes
  - Lag index
  - Gain factor
  - Per band LTP flag
  - Per frame LTP flag

Outline

Current approach

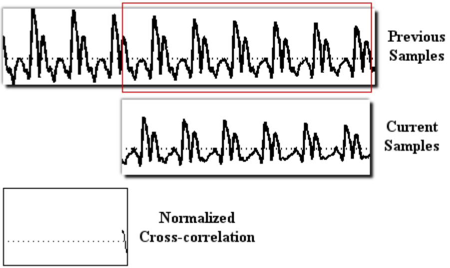
- Lag and gain selected to minimize a mean squared prediction error cost



Current approach

- Lag and gain selected to minimize a mean squared prediction error cost
- The solution results in following choice of lag ($\mathbf{L}$)

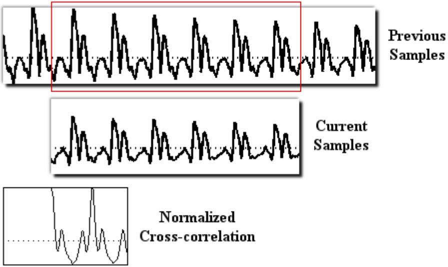
$$\mathbf{L} = \underset{\mathbf{L}' \in [0, 2K)}{\arg\max} \frac{\sum_{m=0}^{2K-1} x[m] z[m + 2K - \mathbf{L}']}{\sqrt{\sum_{m=0}^{2K-1} z^2[m + 2K - \mathbf{L}']}}$$



Current approach

- Lag and gain selected to minimize a mean squared prediction error cost
- The solution results in following choice of lag (**L**)

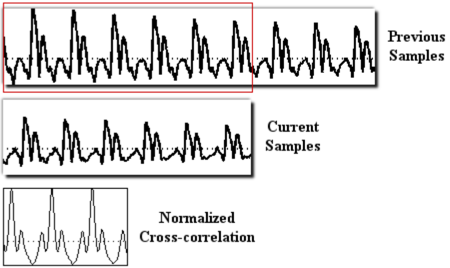
$$\mathbf{L} = \arg\max_{\mathbf{L}' \in [0, 2K)} \frac{\sum_{m=0}^{2K-1} x[m] z[m + 2K - \mathbf{L}']}{\sqrt{\sum_{m=0}^{2K-1} z^2[m + 2K - \mathbf{L}']}}$$



Current approach

- Lag and gain selected to minimize a mean squared prediction error cost
- The solution results in following choice of lag (**L**)

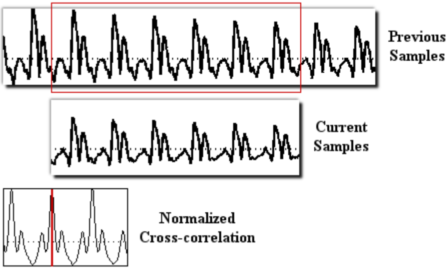
$$\mathbf{L} = \arg\max_{\mathbf{L}' \in [0,2K)} \frac{\sum_{m=0}^{2K-1} x[m] z[m+2K-\mathbf{L}']}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}']}}$$



Current approach

- Lag and gain selected to minimize a mean squared prediction error cost
- The solution results in following choice of lag (**L**)

$$\mathbf{L} = \underset{\mathbf{L}' \in [0, 2K)}{\arg\max} \frac{\sum_{m=0}^{2K-1} x[m] z[m + 2K - \mathbf{L}']}{\sqrt{\sum_{m=0}^{2K-1} z^2[m + 2K - \mathbf{L}']}}$$

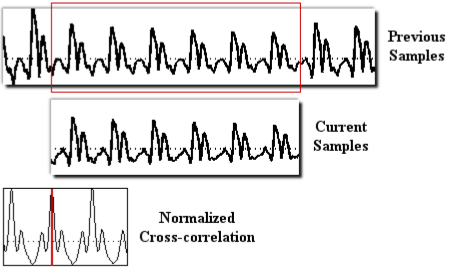


Current approach

- Gain (**G**) calculated as

$$\mathbf{G} = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}$$

- Gain further quantized to one of the 8 levels



Current approach

- For each band, LTP flag chosen as

$$\begin{cases} 1, & \text{if Energy of prediction residue} < \text{Energy of original coefficients} \\ 0, & \text{otherwise} \end{cases}$$

- The per frame flag is set if heuristic bit savings due to LTP > side-information rate of LTP

Current approach

- For each band, LTP flag chosen as

$$\begin{cases} 1, & \text{if Energy of prediction residue} < \text{Energy of original coefficients} \\ 0, & \text{otherwise} \end{cases}$$

- The per frame flag is set if heuristic bit savings due to LTP $>$ side-information rate of LTP

Current approach

- Given all LTP parameters, core AAC parameters are selected via a two-loop search (TLS)
- For every band, an inner loop finds quantization step size for a target distortion criterion
- The outer loop then finds Huffman code books that minimize the bits to encode and if this doesn't meet the rate constraint for the frame, the target distortion is changed and inner loop repeated

Current approach

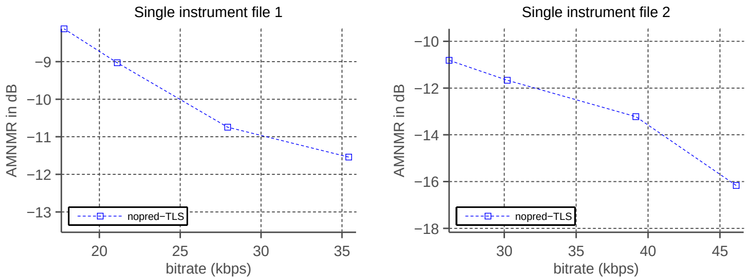
- Given all LTP parameters, core AAC parameters are selected via a two-loop search (TLS)
- For every band, an inner loop finds quantization step size for a target distortion criterion
- The outer loop then finds Huffman code books that minimize the bits to encode and if this doesn't meet the rate constraint for the frame, the target distortion is changed and inner loop repeated

Current approach

- Given all LTP parameters, core AAC parameters are selected via a two-loop search (TLS)
- For every band, an inner loop finds quantization step size for a target distortion criterion
- The outer loop then finds Huffman code books that minimize the bits to encode and if this doesn't meet the rate constraint for the frame, the target distortion is changed and inner loop repeated

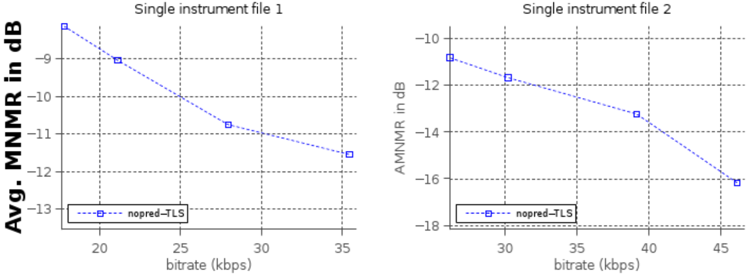
Motivation

- Objective results for reference AAC coder without LTP



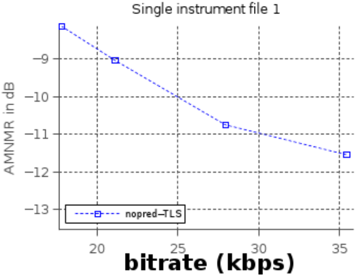
Motivation

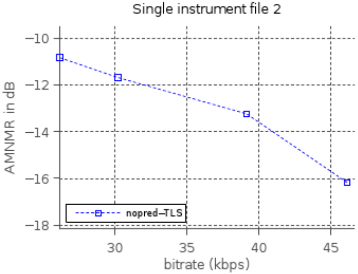
- Objective results for reference AAC coder without LTP



Motivation

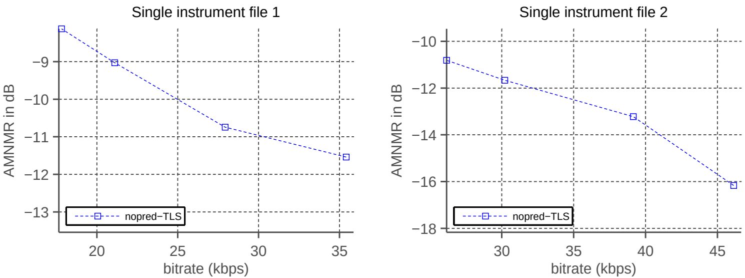
- Objective results for reference AAC coder without LTP





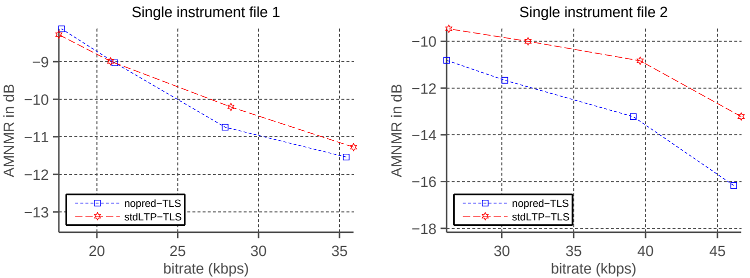
Motivation

- Objective results for reference AAC coder without LTP



Motivation

- Objective results for reference AAC coder with and without LTP



Motivation

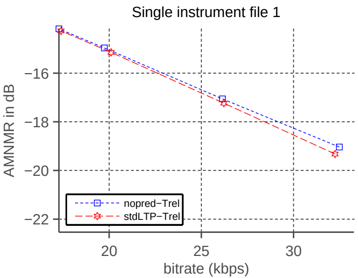
- We know that TLS is sub-optimal for core AAC parameters selection
  - Could this be the reason for the poor RD performance?
- Replace TLS with RD optimal Trellis based core AAC parameters selection [Aggarwal et al. 2006]

Motivation

- We know that TLS is sub-optimal for core AAC parameters selection
  - Could this be the reason for the poor RD performance?
- Replace TLS with RD optimal Trellis based core AAC parameters selection [Aggarwal et al. 2006]

Motivation

- Objective results for Trellis based AAC coder with and without LTP



Motivation

- LTP tool is expected to give improvements for single instrument harmonic files
- Shortcomings attributed to the sub-optimal LTP parameter selection
  - RD optimal approach has to select all encoder parameters with the objective of minimizing perceptual distortion for a given rate
  - Current approach clearly sub-optimal as LTP parameters selected to minimize mean squared prediction error, and independent of core AAC parameters
  - Lag and gain selection ignores eventual prediction switching off in select bands
    - Time domain lag and gain selection effectively considers all transform coefficients
    - Lag and gain thus selected not the best when considering a reduced set of coefficients
  - The heuristically estimated bit savings due to LTP doesn't match actual bit savings reflected after the quantization and coding process

Motivation

- LTP tool is expected to give improvements for single instrument harmonic files
- Shortcomings attributed to the sub-optimal LTP parameter selection
  - RD optimal approach has to select all encoder parameters with the objective of minimizing perceptual distortion for a given rate
  - Current approach clearly sub-optimal as LTP parameters selected to minimize mean squared prediction error, and independent of core AAC parameters
  - Lag and gain selection ignores eventual prediction switching off in select bands
    - Time domain lag and gain selection effectively considers all transform coefficients
    - Lag and gain thus selected not the best when considering a reduced set of coefficients
  - The heuristically estimated bit savings due to LTP doesn't match actual bit savings reflected after the quantization and coding process

Motivation

- LTP tool is expected to give improvements for single instrument harmonic files
- Shortcomings attributed to the sub-optimal LTP parameter selection
  - RD optimal approach has to select all encoder parameters with the objective of minimizing perceptual distortion for a given rate
  - Current approach clearly sub-optimal as LTP parameters selected to minimize mean squared prediction error, and independent of core AAC parameters
  - Lag and gain selection ignores eventual prediction switching off in select bands
    - Time domain lag and gain selection effectively considers all transform coefficients
    - Lag and gain thus selected not the best when considering a reduced set of coefficients
  - The heuristically estimated bit savings due to LTP doesn't match actual bit savings reflected after the quantization and coding process

Motivation

- LTP tool is expected to give improvements for single instrument harmonic files
- Shortcomings attributed to the sub-optimal LTP parameter selection
  - RD optimal approach has to select all encoder parameters with the objective of minimizing perceptual distortion for a given rate
  - Current approach clearly sub-optimal as LTP parameters selected to minimize mean squared prediction error, and independent of core AAC parameters
  - Lag and gain selection ignores eventual prediction switching off in select bands
    - Time domain lag and gain selection effectively considers all transform coefficients
    - Lag and gain thus selected not the best when considering a reduced set of coefficients
  - The heuristically estimated bit savings due to LTP doesn't match actual bit savings reflected after the quantization and coding process

Motivation

- LTP tool is expected to give improvements for single instrument harmonic files
- Shortcomings attributed to the sub-optimal LTP parameter selection
  - RD optimal approach has to select all encoder parameters with the objective of minimizing perceptual distortion for a given rate
  - Current approach clearly sub-optimal as LTP parameters selected to minimize mean squared prediction error, and independent of core AAC parameters
  - Lag and gain selection ignores eventual prediction switching off in select bands
    - Time domain lag and gain selection effectively considers all transform coefficients
    - Lag and gain thus selected not the best when considering a reduced set of coefficients
  - The heuristically estimated bit savings due to LTP doesn't match actual bit savings reflected after the quantization and coding process

Extension

- Trellis based approach RD optimal for core AAC parameter selection
- Extension
  - All possible LTP parameter combinations formed
  - Each case RD evaluated via Trellis
  - Case which minimizes the distortion for a give rate forms final choice
- Computationally prohibitive as LTP adds significantly more choices of parameters for
  - gain (8)
  - lag (frame length)
  - per band LTP flags (2 power number of bands)

Extension

- Trellis based approach RD optimal for core AAC parameter selection
- Extension
  - All possible LTP parameter combinations formed
  - Each case RD evaluated via Trellis
  - Case which minimizes the distortion for a give rate forms final choice
- Computationally prohibitive as LTP adds significantly more choices of parameters for
  - gain (8)
  - lag (frame length)
  - per band LTP flags (2 power number of bands)

Extension

- Trellis based approach RD optimal for core AAC parameter selection
- Extension
  - All possible LTP parameter combinations formed
  - Each case RD evaluated via Trellis
  - Case which minimizes the distortion for a give rate forms final choice
- Computationally prohibitive as LTP adds significantly more choices of parameters for
  - gain (8)
  - lag (frame length)
  - per band LTP flags (2 power number of bands)

Outline

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Proposed approach

- We achieve computational efficiency by
  - Limiting LTP lag and gain parameter space by careful selection of "prediction survivors"
    - Retains the simplicity of time domain lag and gain calculation
    - Limiting number of LTP parameter combinations also limits the number of full RD evaluations
    - Full RD evaluation enables selection of encoder parameters aligned with the end objective of minimizing perceptual distortion for a given rate
  - Trellis approach, which operates in frequency domain, for selecting the band wise quantization and coding parameters, is extended to select the per band LTP flags as well

Prediction survivors

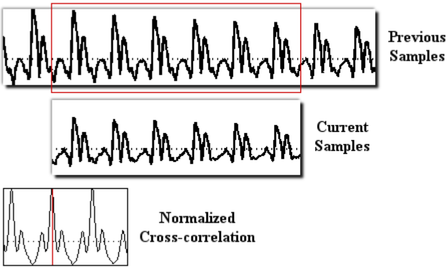
- P lag indices with the highest normalized cross-correlation are retained

$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$

Prediction survivors

- P lag indices with the highest normalized cross-correlation are retained

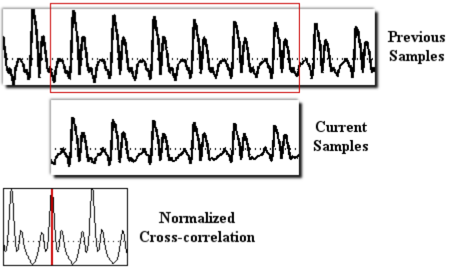
$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$



Prediction survivors

- P lag indices with the highest normalized cross-correlation are retained

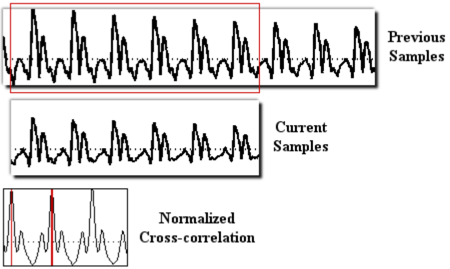
$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$



Prediction survivors

- P lag indices with the highest normalized cross-correlation are retained

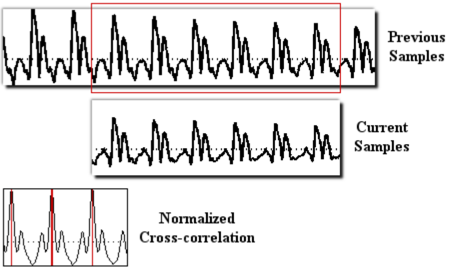
$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$



Prediction survivors

- P lag indices with the highest normalized cross-correlation are retained

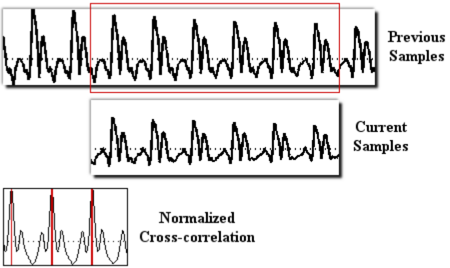
$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$



Prediction survivors

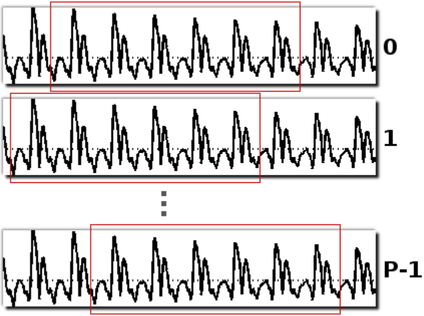
- P lag indices with the highest normalized cross-correlation are retained

$$R[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sqrt{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}}$$



Prediction survivors

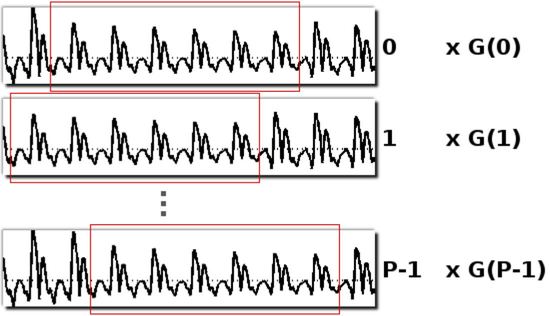
- Forming the P lag survivors



Prediction survivors

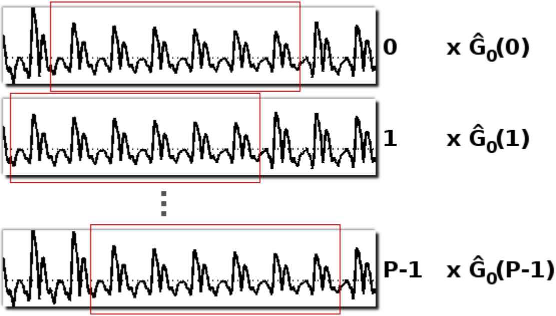
- Gain value for each of these lags found

$$\mathbf{G}[\mathbf{L}] = \frac{\sum_{m=0}^{2K-1} x[m]z[m+2K-\mathbf{L}]}{\sum_{m=0}^{2K-1} z^2[m+2K-\mathbf{L}]}$$



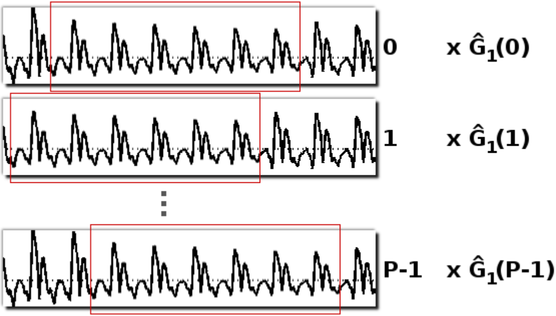
Prediction survivors

- Closest Q quantization levels to each gain value are retained



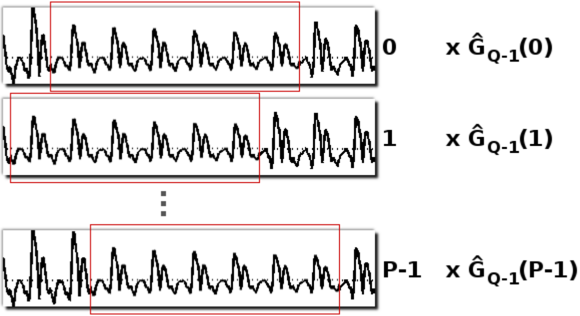
Prediction survivors

- Closest Q quantization levels to each gain value are retained



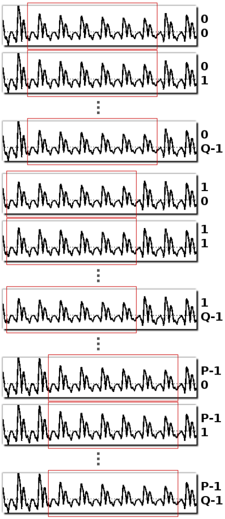
Prediction survivors

- Closest Q quantization levels to each gain value are retained



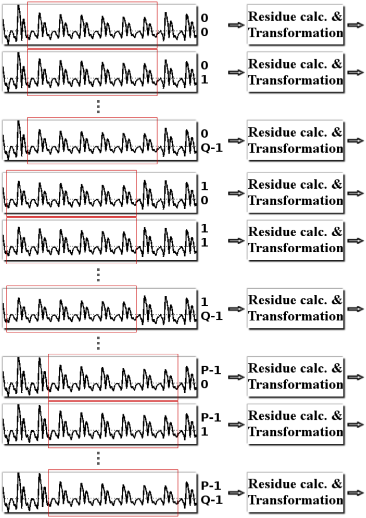
Prediction survivors

- To form the overall PQ survivors



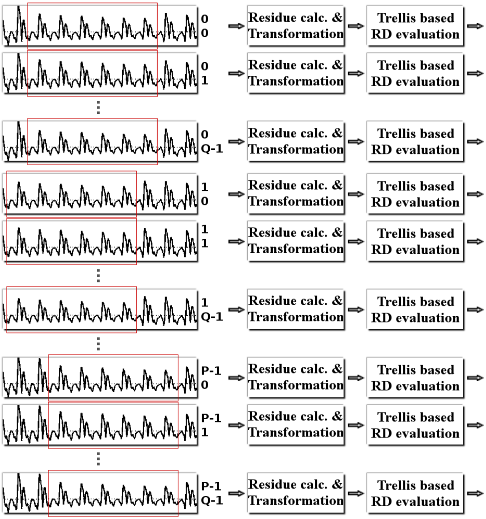
Transformation

- For each survivor, prediction residue is calculated and transformed



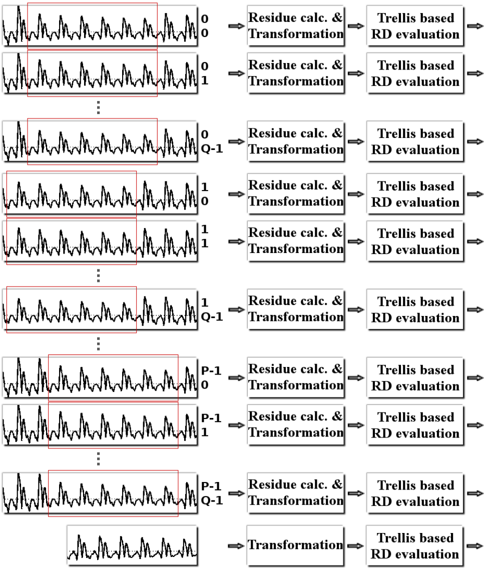
RD evaluation

- Each of these are rate-distortion evaluated via Trellis



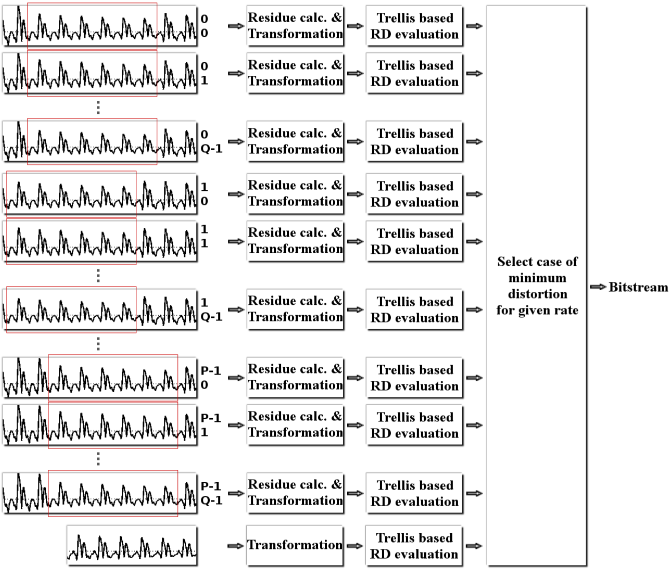
Per frame LTP flag

- To find per frame flag, current frame is transformed and RD evaluated

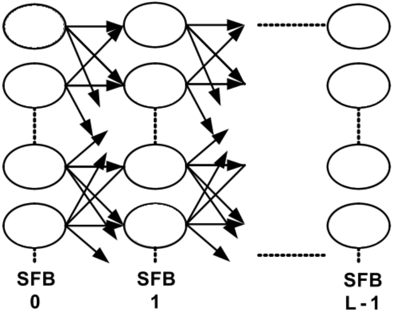


Final selection

- Parameters resulting in minimum distortion for the given rate chosen



Trellis optimization



- Trellis with stages for each band l, states in each stage for every combination of per band LTP flags, quantization and coding parameter values

Trellis optimization

- Each state associated with corresponding distortion and rate costs
- Transition between states associated with costs to differentially encode quantization and coding parameters
- Dynamic programming pursued to find optimal path through trellis
- This path corresponds to optimal set of per band parameters

Trellis optimization

- Each state associated with corresponding distortion and rate costs
- Transition between states associated with costs to differentially encode quantization and coding parameters
- Dynamic programming pursued to find optimal path through trellis
- This path corresponds to optimal set of per band parameters

Trellis optimization

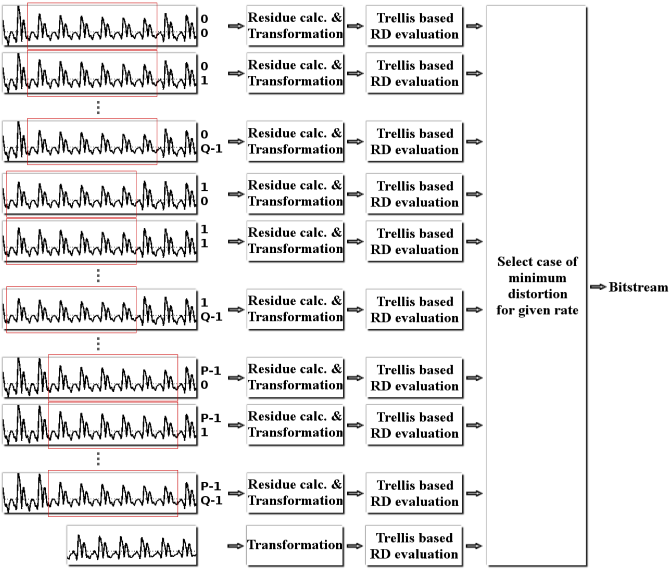
- Each state associated with corresponding distortion and rate costs
- Transition between states associated with costs to differentially encode quantization and coding parameters
- Dynamic programming pursued to find optimal path through trellis
- This path corresponds to optimal set of per band parameters

Trellis optimization

- Each state associated with corresponding distortion and rate costs
- Transition between states associated with costs to differentially encode quantization and coding parameters
- Dynamic programming pursued to find optimal path through trellis
- This path corresponds to optimal set of per band parameters

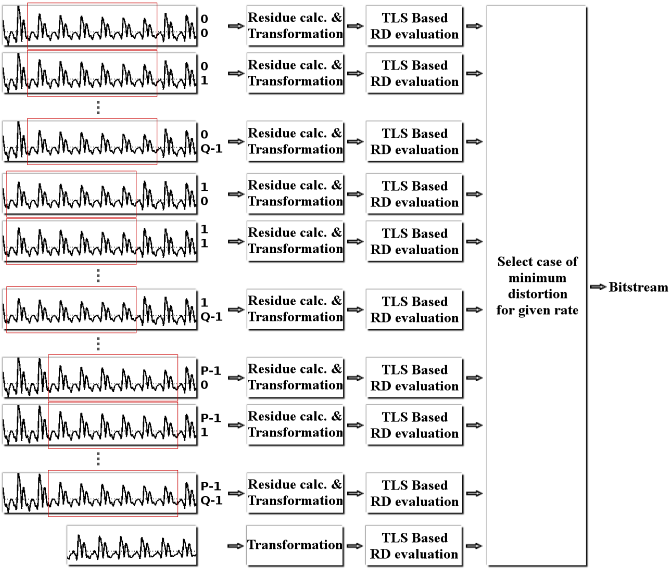
Low complexity variant

- For low complexity Trellis replaced with Two Loop Search



Low complexity variant

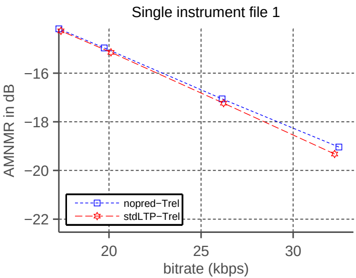
- For low complexity Trellis replaced with Two Loop Search

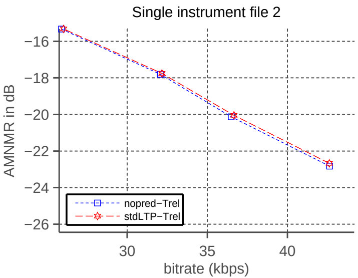


Outline

Objective evaluation results

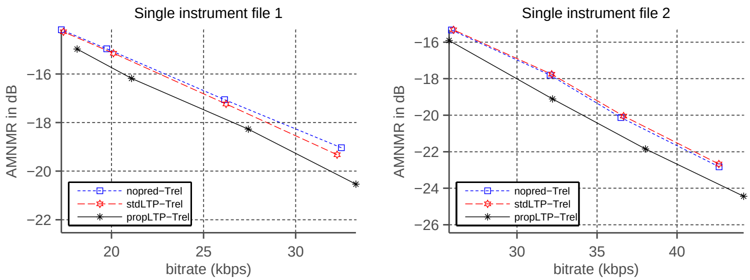
- For Trellis based AAC coder with and without LTP





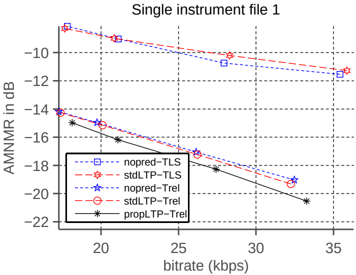
Objective evaluation results

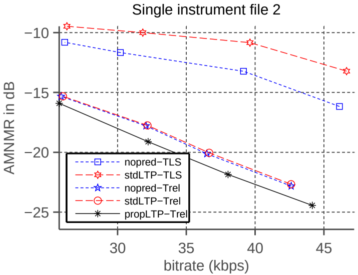
- Along with results for the proposed coder (with $P = 20$ and $Q = 6$)



Objective evaluation results

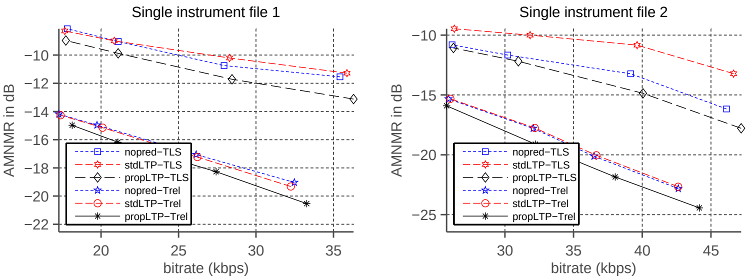
- Trellis based coders compared to the TLS based coders





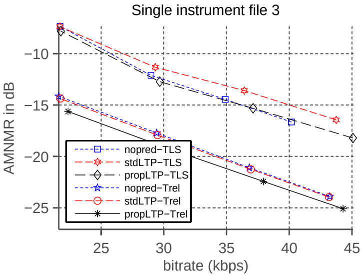
Objective evaluation results

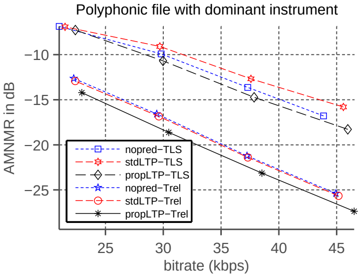
- Along with results for proposed low complexity coder



Objective evaluation results

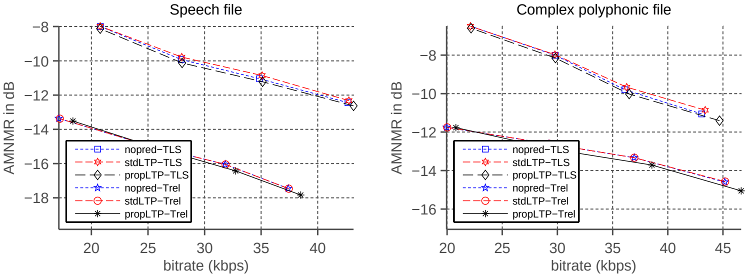
- For other files





Objective evaluation results

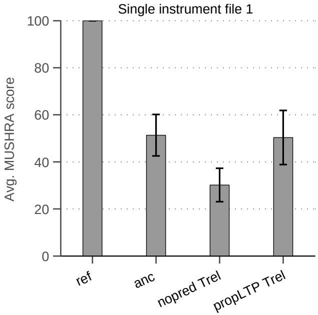
- For other files

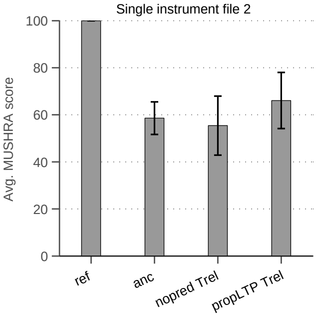


Subjective evaluation

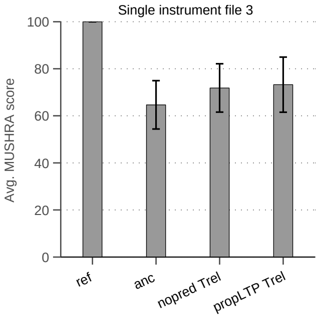
- MUSHRA listening tests for coders operating at 32 kbps
- 12 listeners score on a scale of 0 (bad) to 100 (excellent)
- Plots show average MUSHRA scores and 95% confidence interval

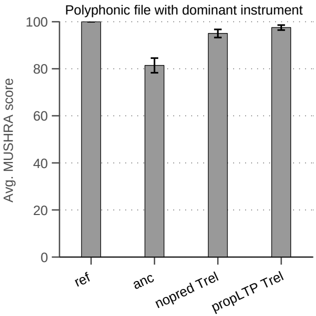
Subjective evaluation results



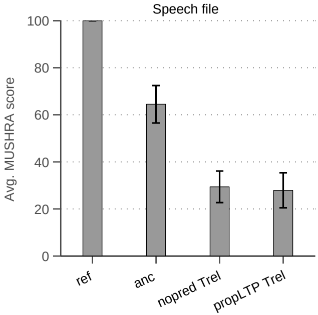


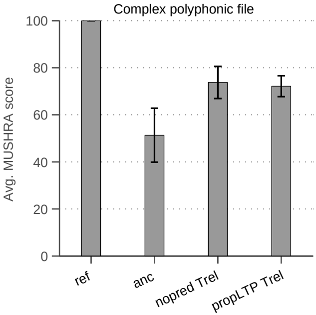
Subjective evaluation results





Subjective evaluation results





Summary

- Current approach for LTP parameter selection is sub-optimal
- Joint selection of LTP and core AAC parameters which optimize perceptual distortion-rate performance proposed
- Low complexity two-loop search based variant also proposed
- Subjective and objective evaluations show substantial improvements
- We conclude that when rightly optimized LTP can be a potent tool

Summary

- Current approach for LTP parameter selection is sub-optimal
- Joint selection of LTP and core AAC parameters which optimize perceptual distortion-rate performance proposed
- Low complexity two-loop search based variant also proposed
- Subjective and objective evaluations show substantial improvements
- We conclude that when rightly optimized LTP can be a potent tool

Summary

- Current approach for LTP parameter selection is sub-optimal
- Joint selection of LTP and core AAC parameters which optimize perceptual distortion-rate performance proposed
- Low complexity two-loop search based variant also proposed
- Subjective and objective evaluations show substantial improvements
- We conclude that when rightly optimized LTP can be a potent tool

Summary

- Current approach for LTP parameter selection is sub-optimal
- Joint selection of LTP and core AAC parameters which optimize perceptual distortion-rate performance proposed
- Low complexity two-loop search based variant also proposed
- Subjective and objective evaluations show substantial improvements
- We conclude that when rightly optimized LTP can be a potent tool

Summary

- Current approach for LTP parameter selection is sub-optimal
- Joint selection of LTP and core AAC parameters which optimize perceptual distortion-rate performance proposed
- Low complexity two-loop search based variant also proposed
- Subjective and objective evaluations show substantial improvements
- We conclude that when rightly optimized LTP can be a potent tool

Thank you for your attention

Questions?