www.atc-labs.com

Audio Communication Coder

Aníbal Ferreira, ajf@atc-labs.com
University of Porto, ATC Labs

Deepen Sinha, sinha@atc-labs.com
ATC Labs

Abstract

- 3G mobile and wireless communication networks elicit new ways of multimedia human interaction and communication, notably two-way high-quality audio communication. This is inline with both the consumer expectation of new audio experiences and functionalities, and with the motivation of Telecom Operators to offer consumers new services and communication modalities. In this paper we describe the design and optimization of a monophonic audio coder (Audio Communication Coder -ACC) that features low-delay coding (< 50 ms) and intrinsic error robustness, while minimizing complexity and achieving competitive coding gains and audio quality at bit rates around 32 kbit/s and higher. ACC source, perceptual and bandwidth extension tools are described and an emphasis is placed on ACC structural and operational features making it suitable for real-time, two-way audio communication. A few performance results are also presented. Audio demos are available at http://www.atc-labs.com/acc/ .

Outline

- Introduction
- ACC encoder
- ACC decoder
- ACC structural features
- ACC operational features
- Performance evaluation
- Conclusion

Introduction

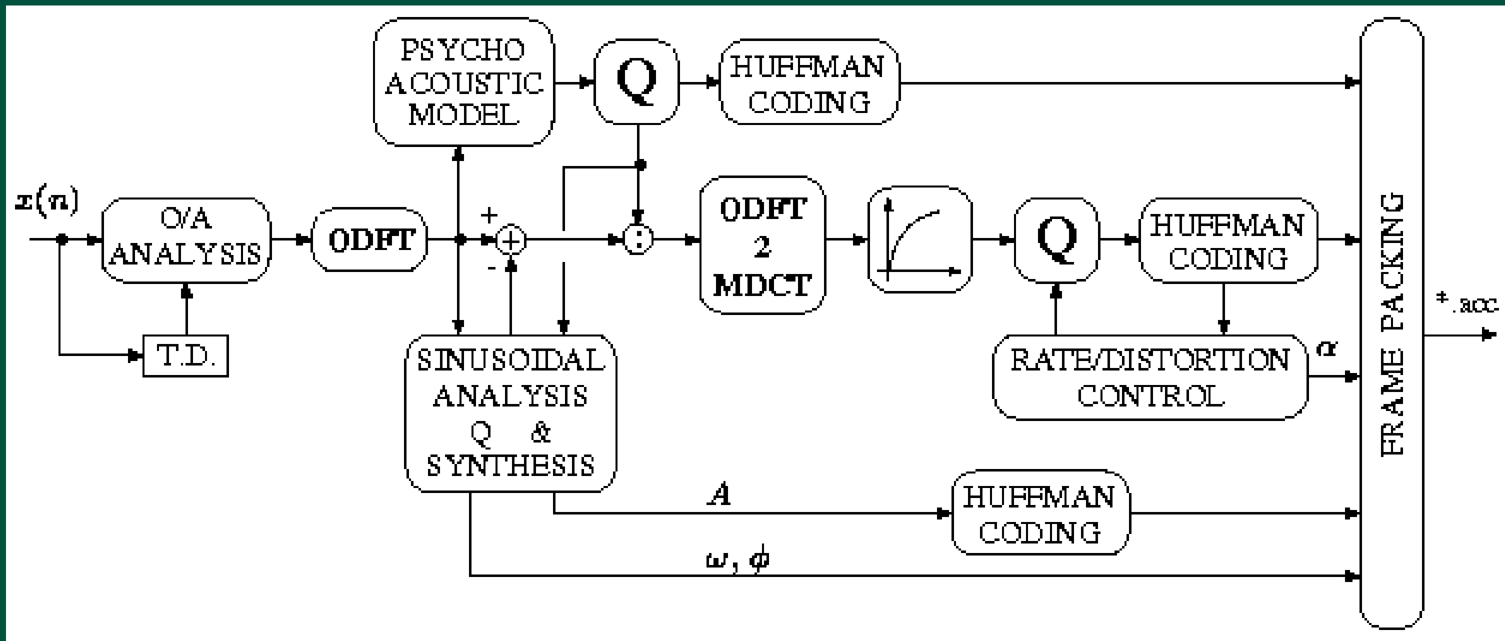
- typical application areas of perceptual audio coders
  - broadcast
  - streaming
  - messaging
  - (Internet) download
- place emphasis on specific features
  - compression efficiency
  - audio quality
- in detriment of other relevant features
  - end-to-end coding delay
  - algorithmic complexity
  - intrinsic error robustness

Introduction

- emerging application area in the context of 3G mobile and wireless communication networks
  - real-time two-way high-quality audio communication
- requiring
  - competitive compression efficiency and audio quality
  - low end-to-end coding delay
  - low algorithmic complexity
  - high intrinsic error robustness
- new coder: Audio Communication Coder (ACC)
  - focuses on intraframe coding
  - takes advantage of
    - source/parametric coding techniques
    - perceptual coding techniques
    - bandwidth extension techniques

ACC encoder

- structure

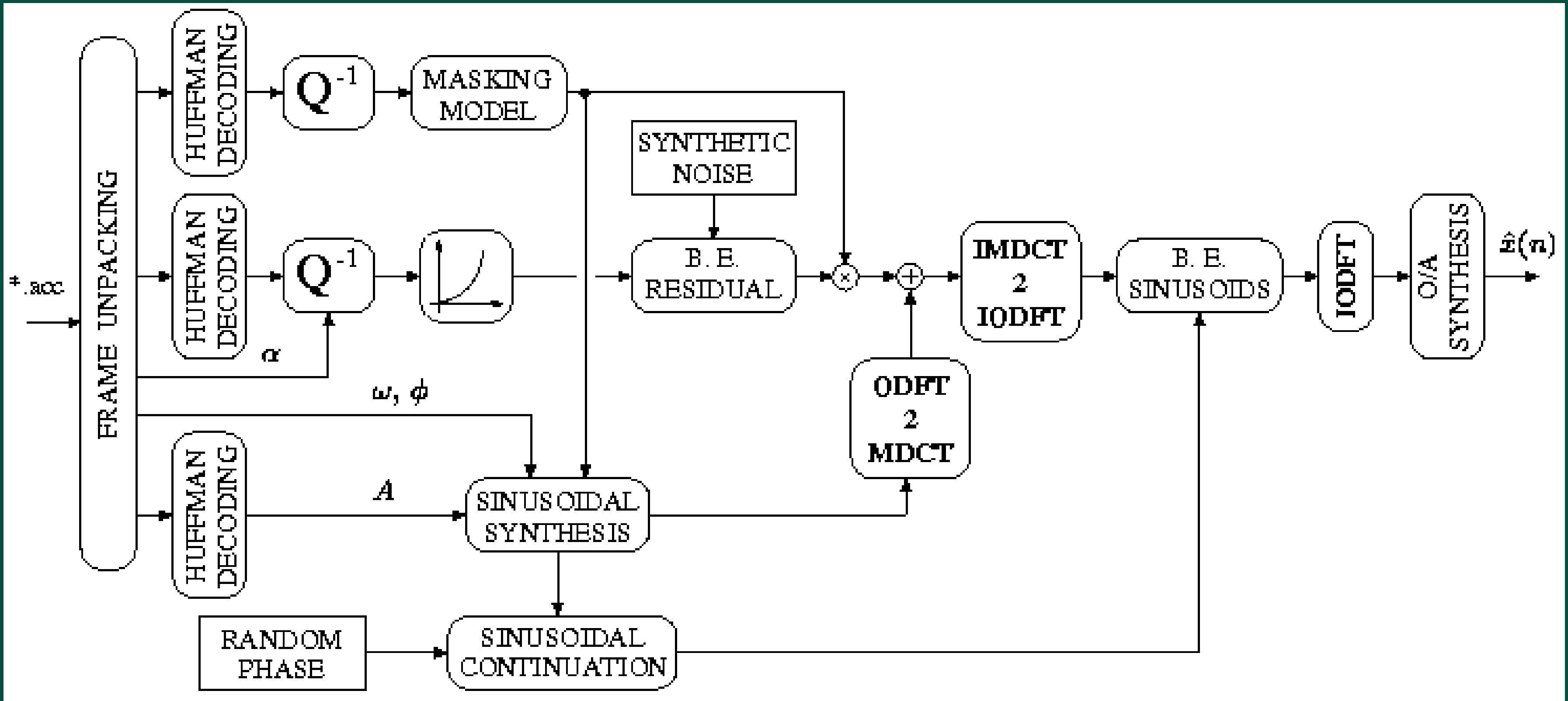


ACC encoder

- combines
  - source coding tools
    - transform coding
    - parametric coding
    - PDF-optimized quantization
    - entropy coding
  - perceptual coding tools
    - psychoacoustic modeling
    - perceptual noise shaping
  - bandwidth extension tools
    - accurate replacement of both sinusoidal and noise components of the original signal that are missing or were bandwidth reduced, by a perceptually similar synthetic sound

ACC decoder

- structure



www.atc-labs.com
[ajf, sinha]@atc-labs.com

ACC decoder

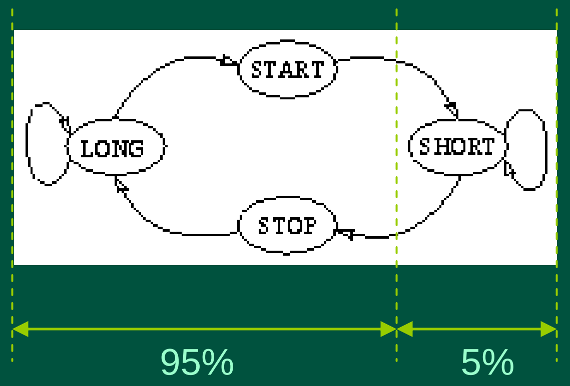
- major processing steps
  - Huffman decoding
  - inverse quantization
  - signal reconstruction
  - bandwidth extension (using FSSM and ASR)
    - sinusoids are bandwidth extended in the ODFT domain
    - noise is bandwidth extended in the MDCT domain

ACC structural features

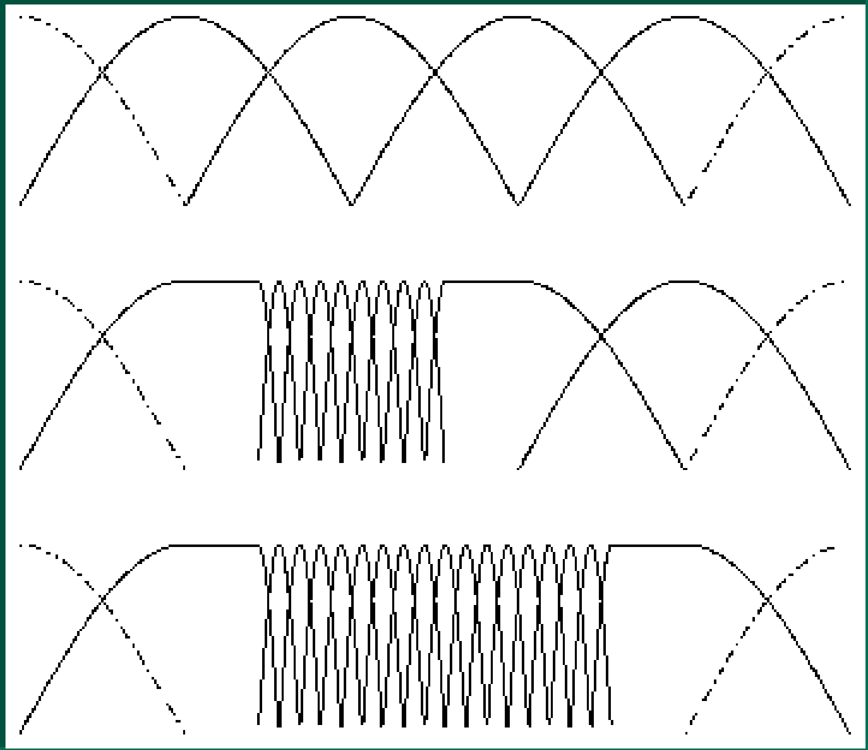
- parametric coding of sinusoids
  - as individual sinusoids or harmonic structures
  - spectral subtraction of the modeled sinusoids leads to an effective spectral flattening
- window switching
  - typical audio material
    - quasi-stationary regions
      - about 95% of the time
      - Human Auditory System (HAS) focuses on spectral details
      - coding using high frequency resolution
    - non-stationary regions
      - about 5% of the time
      - HAS focuses on temporal details
      - coding using high time resolution

ACC structural features

- window switching in ACC
  - state diagram ⇓



  - typical examples ⇒
    (frequency/time switching factor: 8)

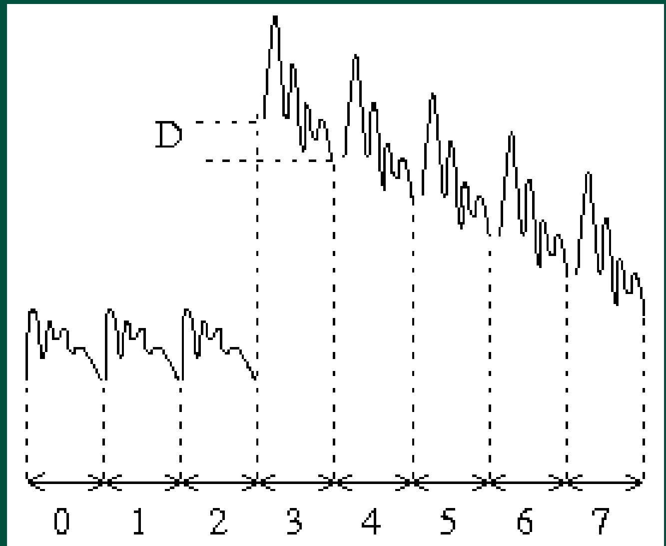


ACC structural features

- window switching issues
  - high time resolution $\Rightarrow$ low coding gain
  - concern: minimize side information due to short blocks
  - opportunity: shape of Threshold of Masking (THR) is quite comparable for those short blocks before the ‘attack’, and also for those short blocks over and after the ‘attack’

ACC structural features

- window switching issues
  - ACC modeling of the Masking Threshold of short blocks



In this example, the audio 'attack' takes place over the 4th short block

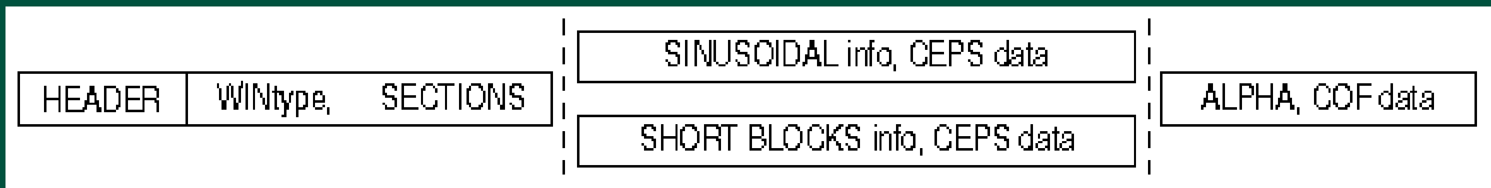
  - ACCurate localization of audio 'attacks' in two steps:
    - time-domain transient detector
    - frequency domain analysis of THR dynamics

ACC structural features

- PDF-optimized quantization
  - residual after spectral subtraction of sinusoids and spectral normalization by the Threshold of Masking, exhibits a well defined PDF
    - ⇒ PDF-optimized quantization through companding and uniform quantization
- bandwidth extension (BWE)
  - allows replacement, in a perceptually suitable way, of a spectral region of the original signal that has not been coded or has been lost
    - ⇒ BWE combining techniques of FSSM and ASR

ACC structural features

- Huffman coding
  - applied to
    - ⇒ quantized MDCT coefficients (COF1 ↔ COF7)
    - ⇒ quantized cepstral coefficients of the THR model (CEPS)
    - ⇒ quantized differential magnitudes of sinusoids (SINM)
- bit stream structure
  - four main fields



ACC structural features

- bit stream structure
  - header (fixed bit size)
    - synchronization pattern
    - number of audio channels
    - length of LONG windows
    - window switching option
    - coding bit rate
    - sampling frequency
  - sectioning scheme (variable bit size)
    - window type
    - Huffman tables assignment
  - sinusoidal/short block info and CEPS data (variable bit size)
  - under/over coding factor and MDCT data (variable bit size)

ACC operational features

- no inter frame coding
  – avoids error propagation across several frames
  – eases re-synchronization after a failure
  – facilitates error concealment
  – improves time resolution when accessing bit streams
  – facilitates fast-rewind and fast-forward play modes
- minimized coding delay
  – by using a residual or no bit stream buffer
- simple rate-distortion control
  – a single parameter controls rate and distortion simultaneously
  – convergence is fast in CBR coding

Performance evaluation

- performance ≠ compression efficiency

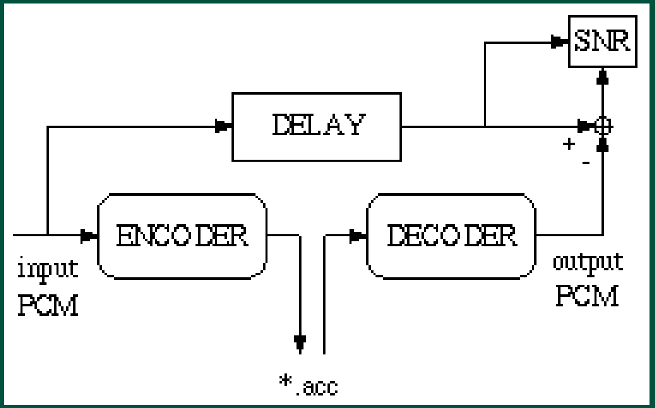
$$\text{performance} = f\,(\text{bitrate, audio quality, delay, complexity, robustness})$$

$$\text{performance} = \underbrace{\underbrace{f_1\,(\text{bitrate, audio quality})}_{\text{compression efficiency}} \times f_2\,(\text{delay, complexity, robustness})}_{\text{performance}}$$

- ACC algorithmic delay
  - ⇒ 21 ms @ 48 kHz sampling frequency
- ACC practical delay
  - ⇒ < 50 ms @ 48, 44.1 or 32 kHz sampling frequency

Performance evaluation

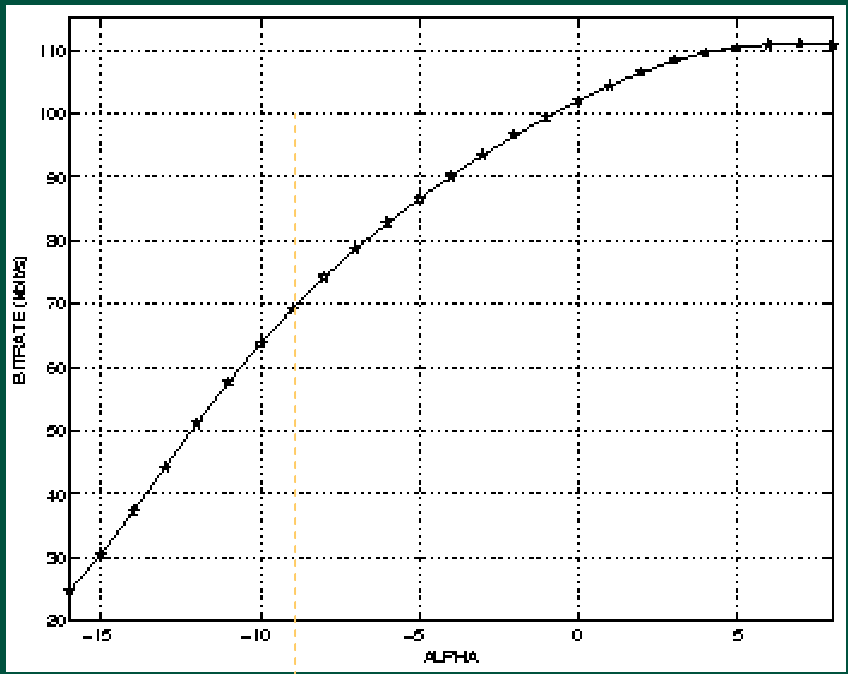
- compression efficiency
  - quality evaluation under VBR / CBR coding
    - 32 kHz sampling frequency
    - no BWE
    - informal subjective evaluation
    - objective evaluation (Signal to Coding Noise Ratio (SNR))

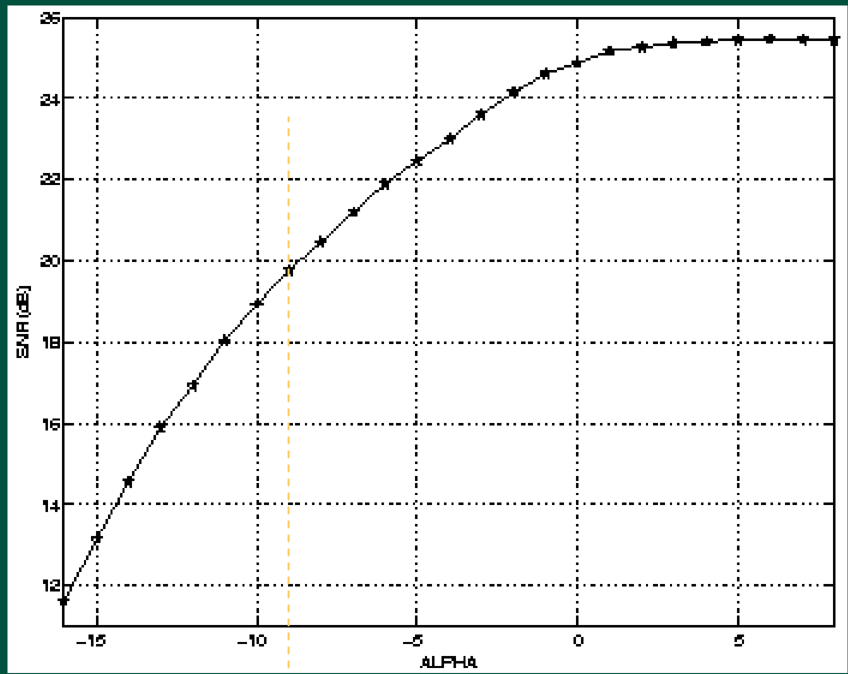


  - comparison of VBR and CBR coding

Performance evaluation

- VBR coding



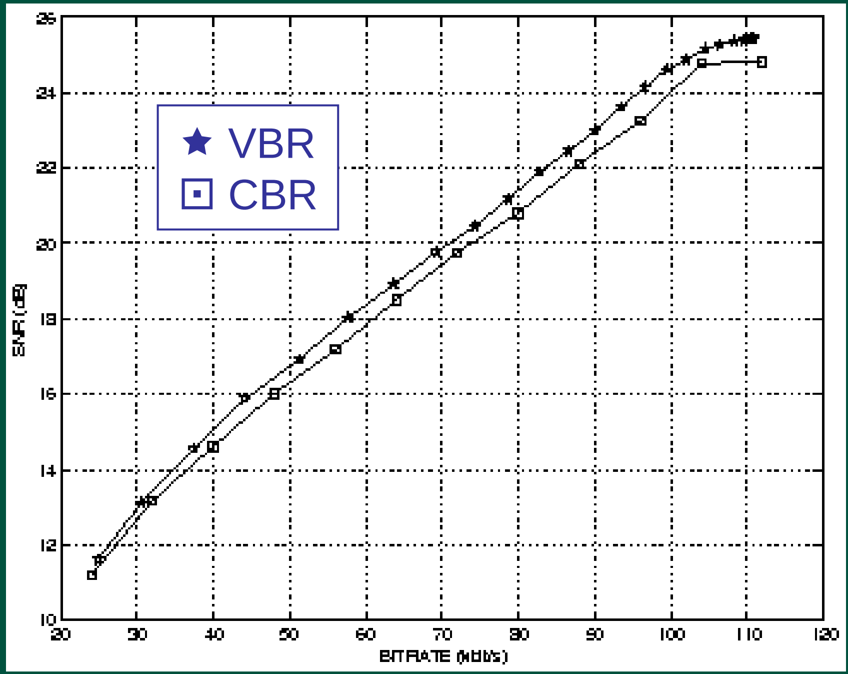


high-quality coding | transparent coding

high-quality coding | transparent coding

Performance evaluation

- VBR versus CBR coding



  - VBR
    - results of previous slide
  - CBR
    - bit rates allowed in bit stream
    - residual bit stream buffer is used
  - main conclusions
    - transparent coding is achieved for most audio material at bit rates between 60 kbit/s and 80 kbit/s (mono)
    - audio quality is not strongly dependent on a large bit stream buffer
    - subjective quality is similar under CBR or VBR coding, for comparable average bit rates

Conclusion

– the structure of Audio Communication Coder (ACC), a source/perceptual coder, has been described

– emphasis has been placed on structural and operational features making ACC suitable for real-time, high-quality audio communication

– ACC features low-delay (< 50 ms) while providing competitive compression efficiency under VBR or CBR

– ACC target application areas include mobile audio communication, B-channel audio communication, and audio conferencing

– ACC audio demos @ http://www.atc-labs.com/acc

www.atc-labs.com
[ajf, sinha]@atc-labs.com