www.atc-labs.com

A New Low-Delay Codec for Two-Way High-Quality Audio Communication

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

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

Abstract

- High-quality audio bit-rate reduction systems are widely used in many application areas involving audio broadcast, streaming and download services. With the advent of 3G mobile and wireless communication networks, there is a clear opportunity for new multimedia services, notably those relying on two-way high-quality audio communication. In this paper we describe a new source/perceptual audio coder that features low-delay, intrinsic error robustness and high subjective audio quality at competitive compression ratios. The structure of the audio coder is described and an emphasis is given on its innovative approaches to semantic signal segmentation and decomposition, independent coding of sinusoidal and noise components, and bandwidth extension using Accurate Spectral Replacement. A few test results are presented that illustrate the operation and performance of the new coder. Audio demos are available at http://www.atc-labs.com/acc/ .

Outline

- Introduction
- Audio Communication Coder (ACC)
- ACC encoder structure
- ACC decoder structure
- Illustrative coding results
- Conclusion

Introduction

- most proprietary and standardized audio compression technologies are designed for
  - broadcast
  - streaming
  - messaging
  - (Internet) download

- these technologies are not suited to two-way high-quality audio communication since they exhibit:
  - high end-to-end coding delay
  - low intrinsic error robustness ($\therefore$ interframe coding)
  - high complexity

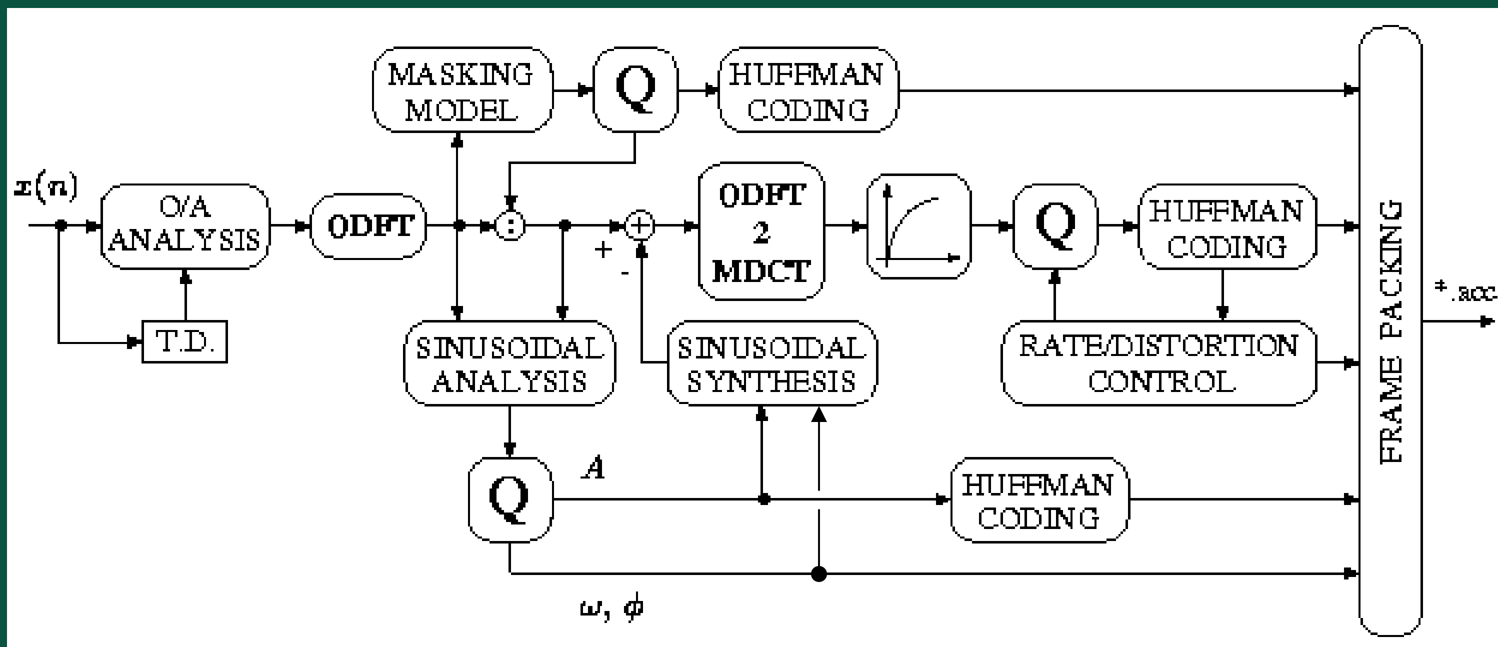
Introduction

- however, real-time audio communication is gaining increased importance in 3G mobile and wireless communication networks, namely due to
  - increased consumer expectation of higher quality and new functionalities in mobile communication
  - pressure of operators who have strongly invested in 3G licenses and expect corresponding return by offering consumers new services

Audio Communication Coder (ACC)

- new source/perceptual audio coder that targets real-time, two-way, high-quality audio communication and that features:
  - competitive compression efficiency
  - low-delay coding (< 50 ms) by minimizing the size of the transform, and look-ahead and bit-stream buffer
  - intra-frame coding only
    - insuring high intrinsic error robustness
    - facilitating error concealment
    - facilitating fast-rewind and fast-forward play modes
  - efficient bandwidth extension using Accurate Spectral Replacement (ASR)

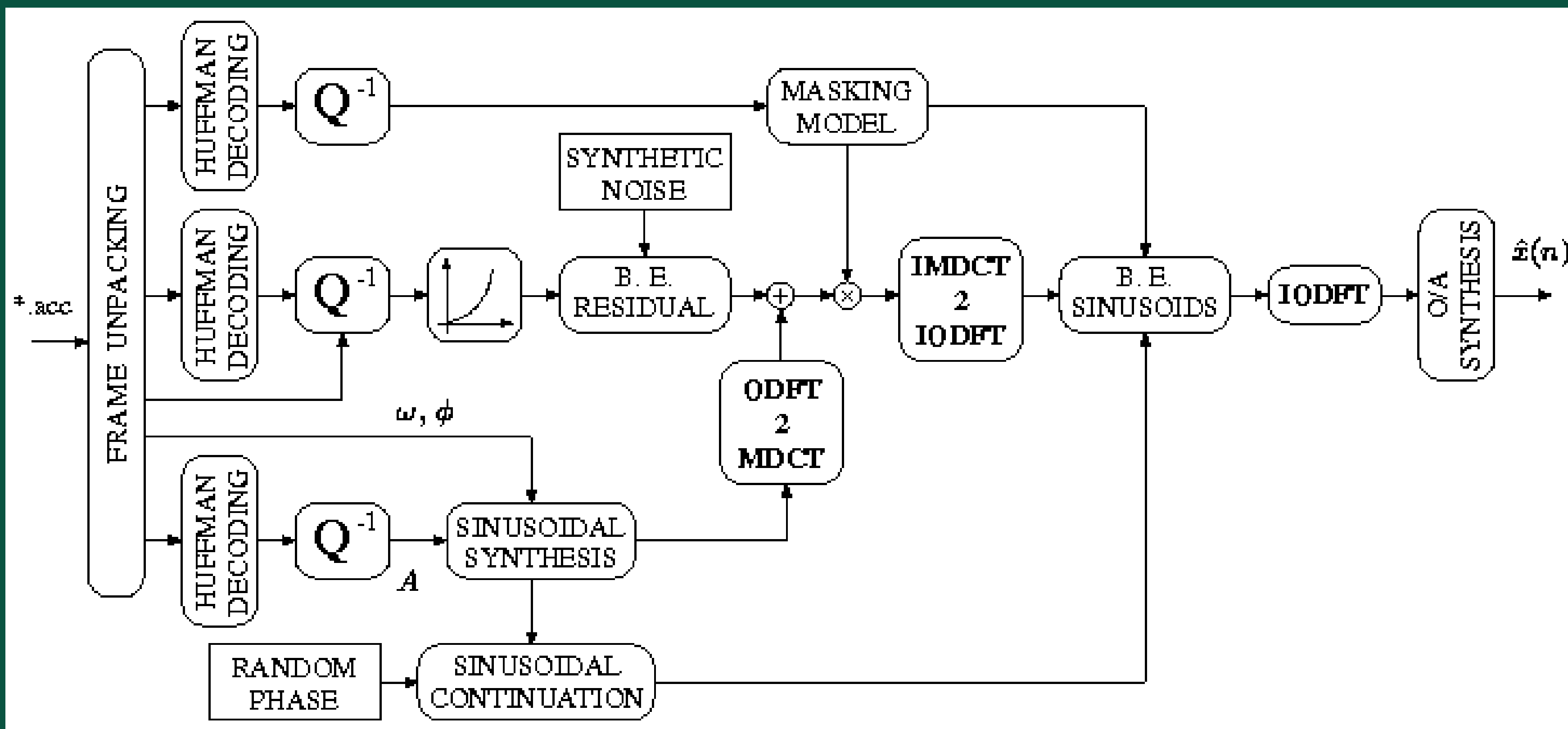
ACC encoder structure



ACC encoder structure

- takes advantage of:
  - source coding tools
    - time-frequency transformation
    - entropy coding
    - parametric coding
    - optimum quantization
  - 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



ACC decoder structure

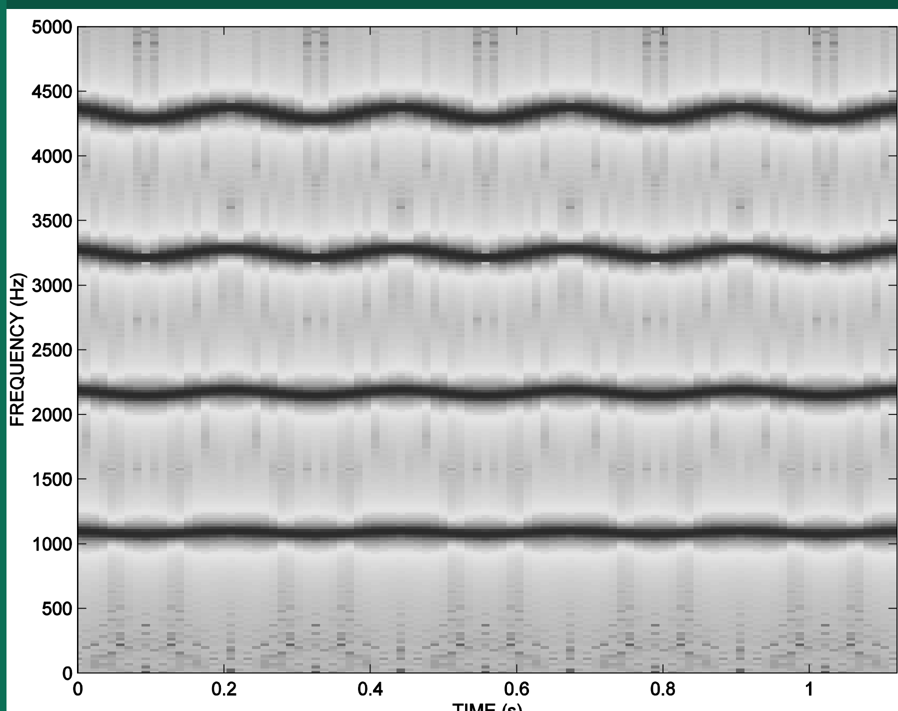
- distinctive features:
  - sinusoids falling within the bandwidth of the coded residual are added back in the MDCT domain
  - bandwidth extension of sinusoids is implemented in the ODFT domain
  - bandwidth extension is separately implemented for sinusoids and noise, according to the Accurate Spectral Replacement (ASR) technique

Illustrative coding results

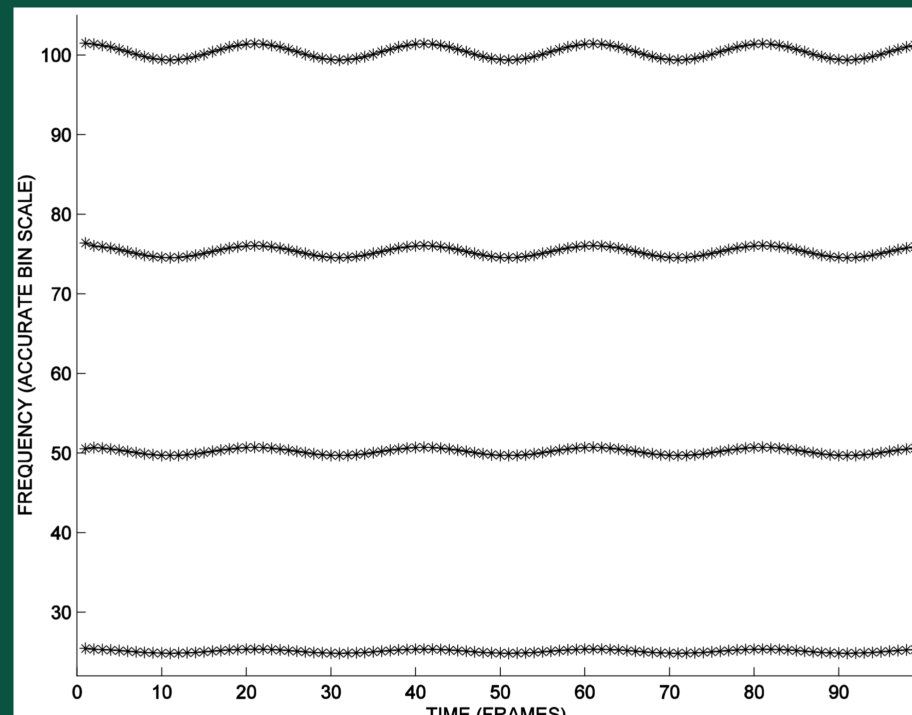
- input: harmonic complex with vibrato

  F_0=25.1 bins, Δf_{MAX}=0.512 bins, partials: F_0, $2F_0$, $3F_0$, $4F_0$

Spectrogram, Fs=44100 Hz, N=1024



instantaneous frequency estimation

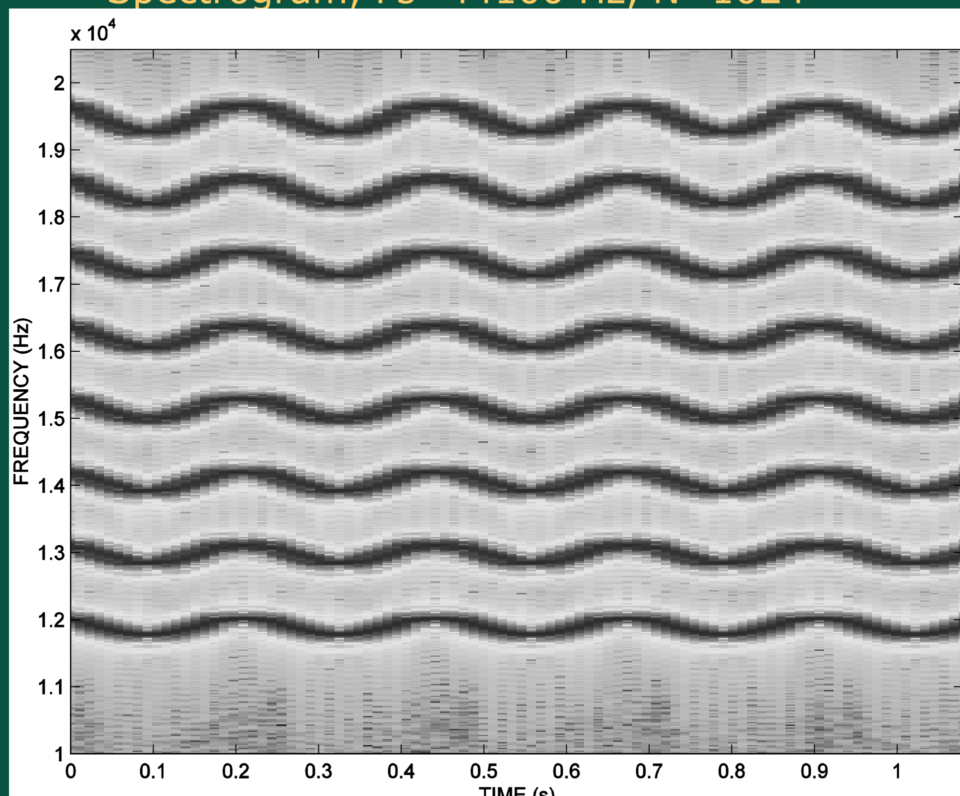


Illustrative coding results

- input: harmonic complex with vibrato

  bandwidth extended partials: $11F_0, \ldots, 18F_0$

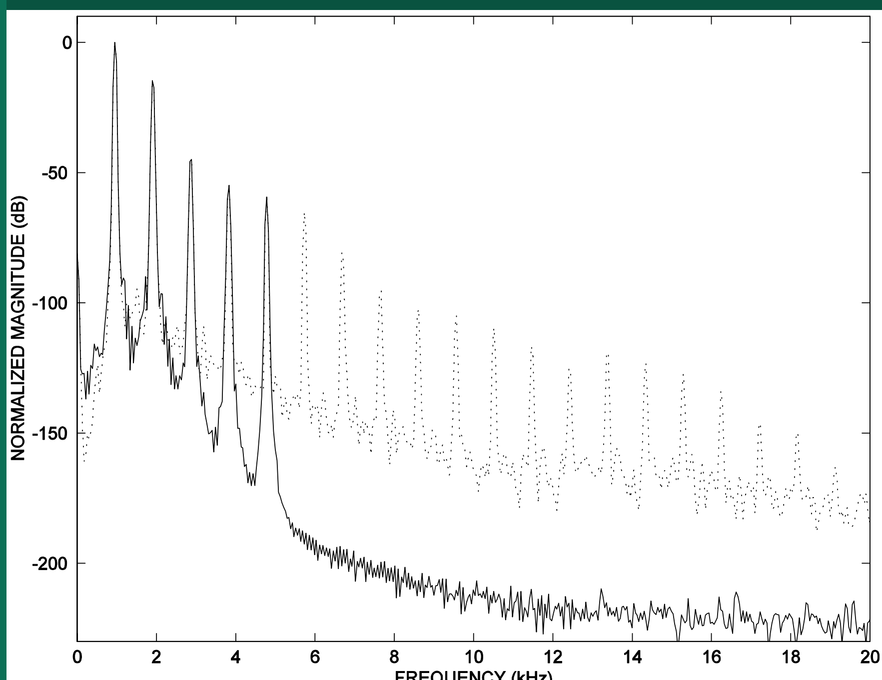
Spectrogram, Fs=44100 Hz, N=1024



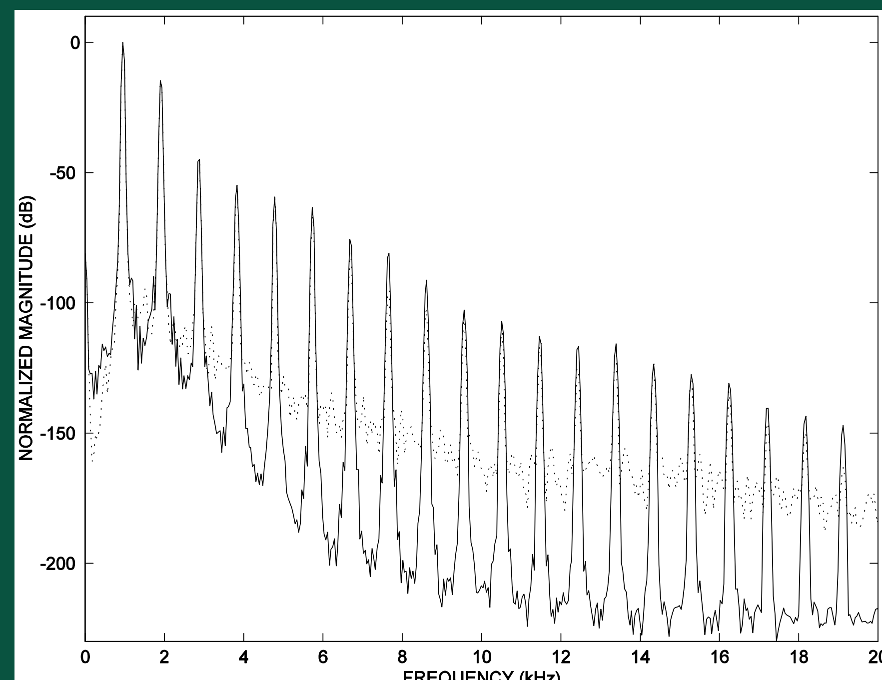
Illustrative coding results

- input: natural music signal (trumpet)
  - sampling frequency: 44100 Hz, 24 kbit/s constant bit rate coding, preserved bandwidth: 5 kHz

PSD of output without B.E. (dotted=original)



PSD of output with B.E. of sinusoids only

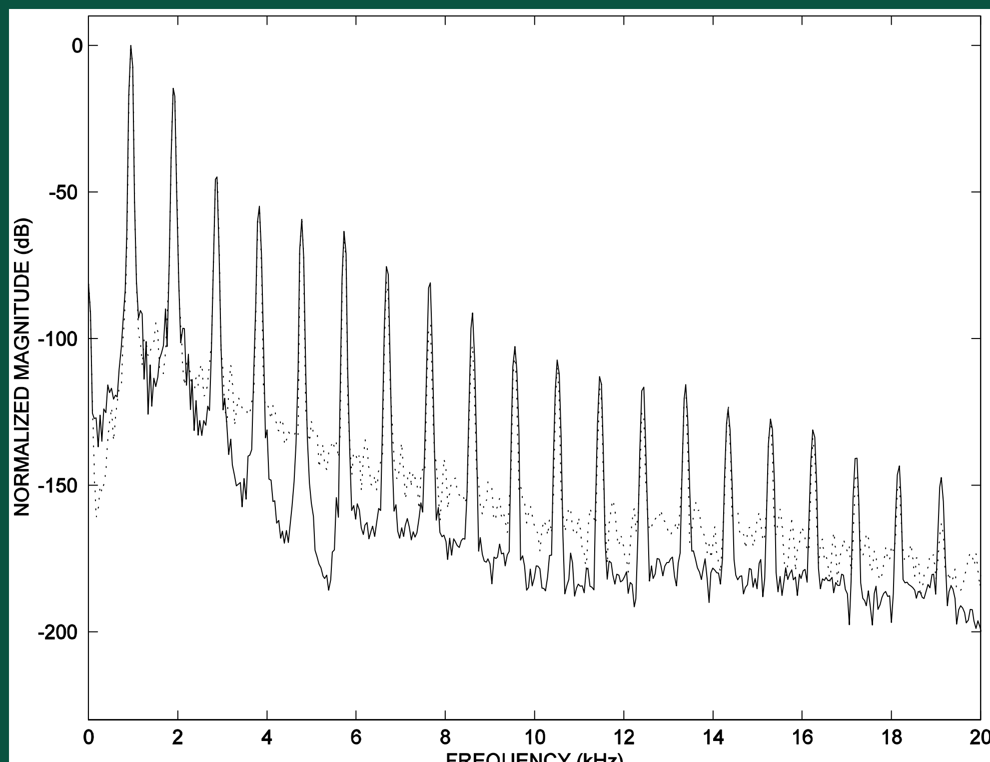


www.atc-labs.com
[ajf, sinha]@atc-labs.com

Illustrative coding results

- input: natural music signal (trumpet)
  – bandwidth extension: 5 kHz $\leftrightarrow$ 20 kHz

PSD of output with B.E. of sinusoids + noise (dotted=original)



Conclusion

- the structure of a new low-delay (< 50 ms) source/perceptual audio coder has been described
- the new coder Audio Communication Coder (ACC) has been designed for real-time, two-way high-quality audio communication
- operation of ACC has been illustrated with both synthetic and natural audio signals
- application scenarios include real-time 3G mobile and wireless communication of high-quality audio
- AAC is amenable to new functionalities including semantic classification, access, filtering and retrieval of audio on the compressed (*i.e.*, bit stream) domain