www.atc-labs.com

Accurate Spectral Replacement

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

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

Abstract

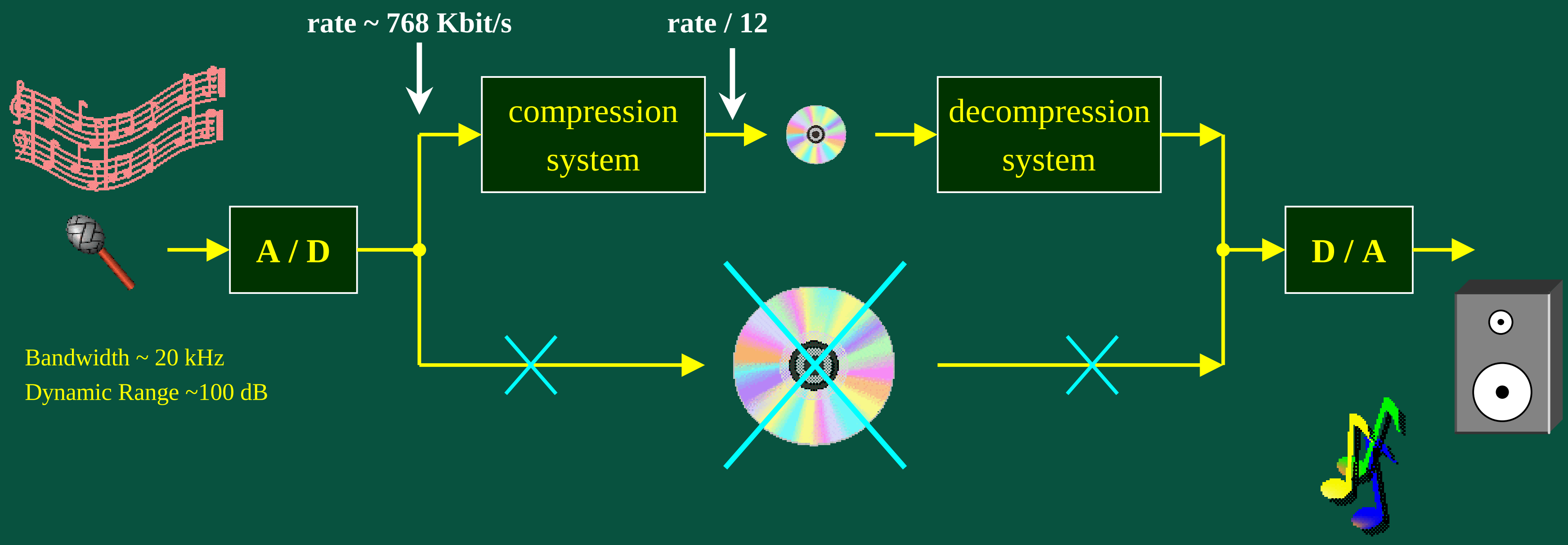
- Recent advances in perceptual audio coding are strongly based on the concept of bandwidth extension. Most techniques implementing bandwidth extension require an analysis/synthesis filter bank in addition to that used by the associated perceptual audio coder, with a clear penalty in system complexity and coding delay. In this paper we present a Accurate Spectral Replacement (ASR) as one of a new class of bandwidth extension techniques applied directly to the high frequency representation of the signal. ASR is based on a suitable decomposition of the MDCT filter bank, and implements synthesis of sinusoidal components with an accuracy much higher than the natural frequency resolution of the MDCT. The ASR technique is described, its performance is assessed with both synthetic and natural audio signals, and its main areas of application are addressed. Audio demos are available at http://www.atc-labs.com/asr/

Summary

- Perceptual coding of high quality audio
  - Perceptual coding paradigm
  - New trends
  - Bandwidth extension
- Accurate Spectral Replacement (ASR)
  - Concept
  - ASR encoder
  - ASR decoder
  - Results
- Conclusion

Perceptual coding of high quality audio

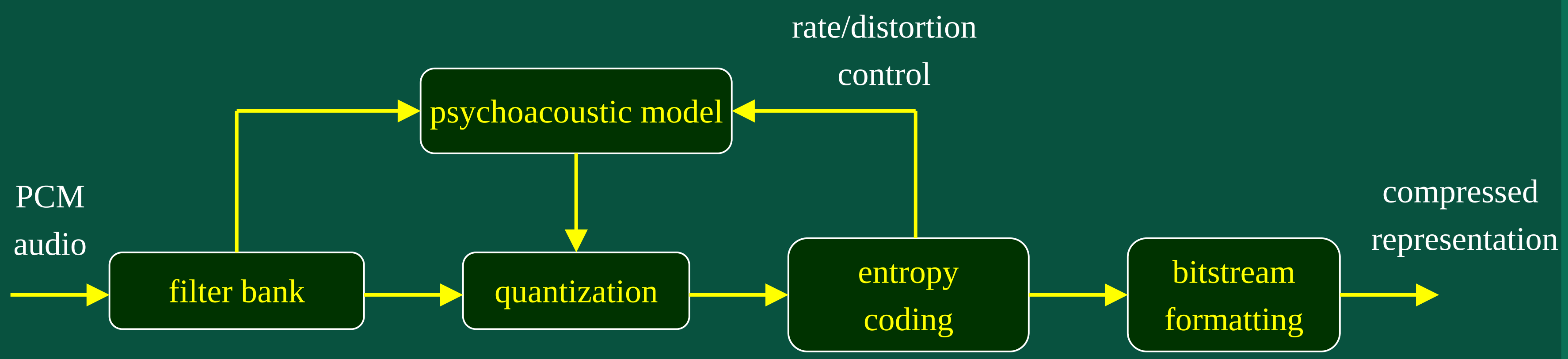
- Perceptual coding paradigm
  - successful approach in the compression of natural audio



  - provides high compression gains
    - e.g., gain in transparent coding of monophonic audio ~ 12:1

Perceptual coding of high quality audio

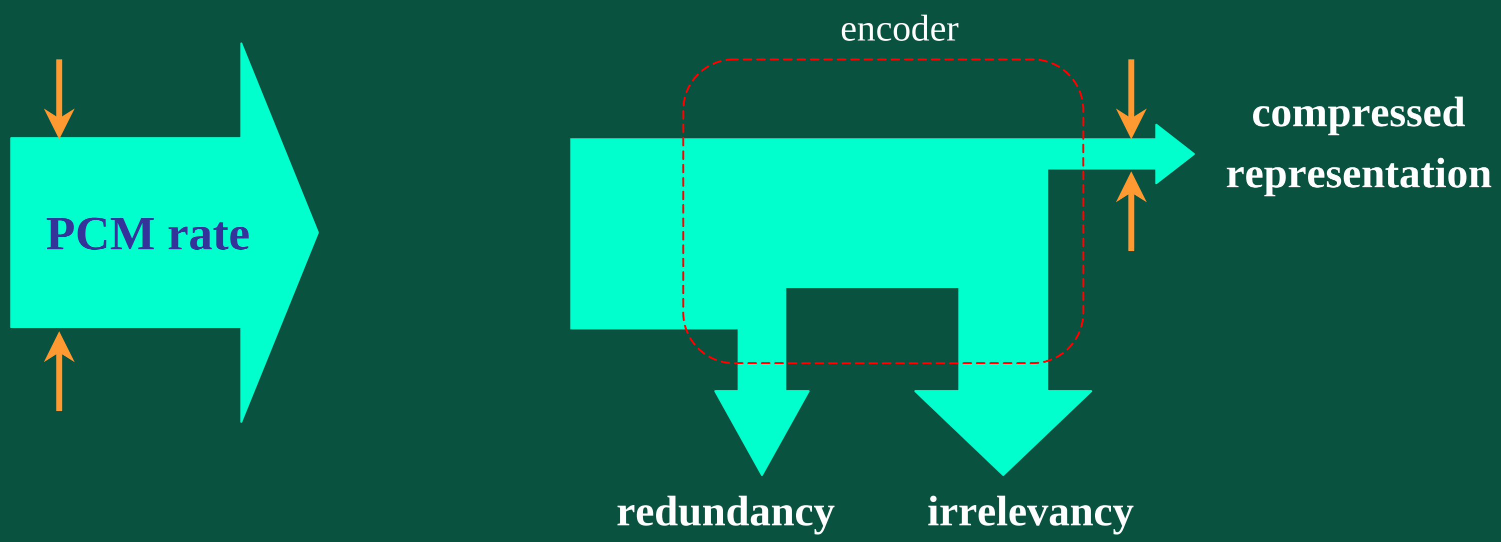
- Perceptual coding paradigm
  - takes advantage of source coding tools and perceptual coding tools



  - several examples exist of proprietary and standardized coding algorithms
    - ATRAC (Sony)
    - PAC (Bell Labs, Lucent)
    - Dolby AC-3
    - MPEG-1/2 Layer 3, MPEG-2 AAC

Perceptual coding of high quality audio

- Perceptual coding paradigm
  - coding gains result from redundancy reduction (source coding) and irrelevancy reduction ("sink" or perceptual coding)

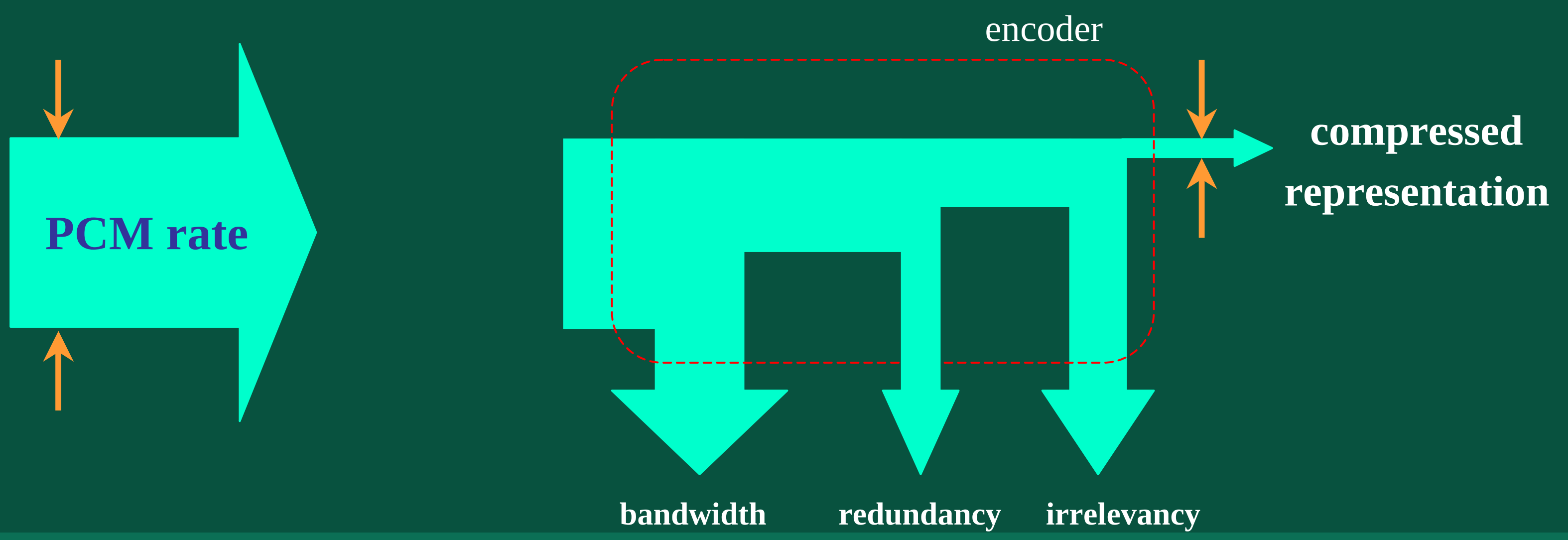


Perceptual coding of high quality audio

- New trends
  - new compression gains
    - are only residual if source and perceptual signal representation techniques focus on the natural audio signal
    - BUT significant gains may result from a shift of the perceptual audio coding paradigm, by combining natural audio coding and synthetic audio such that the sound impression the overall result strongly resembles the original audio signal
  - not new but promising approach: bandwidth extension
    - combines natural audio representation techniques and synthetic audio generation techniques

Perceptual coding of high quality audio

- New trends
  - combination of coding tools
    - natural audio coding
      - redundancy reduction + irrelevancy reduction
    - synthetic audio generation using parametric description
      - bandwidth reduction (encoder) followed by bandwidth extension (decoder)



Perceptual coding of high quality audio

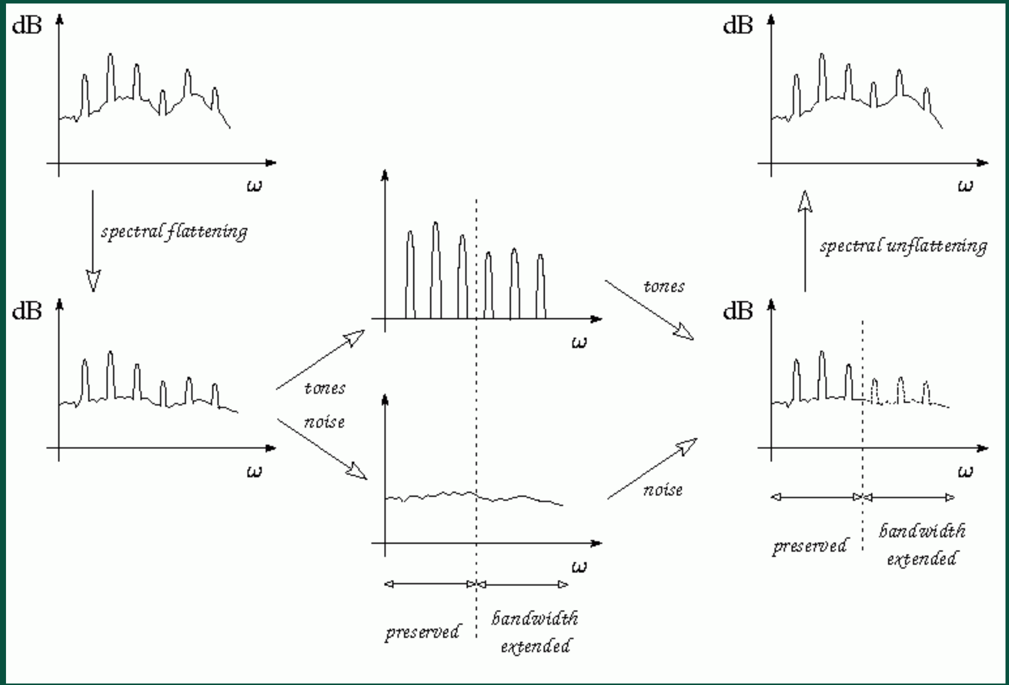
- New trends
  - bandwidth reduction
    - a fraction of the original audio signal is removed at the encoder but a perceptually similar representation is synthesized at the decoder using just a very scarce parametric description
  - redundancy reduction
    - the predictable structure of the original signal is reduced at the encoder and is restored at the decoder without loss
  - irrelevancy reduction
    - Signal components or the original signal which fall below the threshold of masking (are not audible and) are reduced at the encoder (or, equivalently, are replaced by noise), without a subjective loss in quality

Perceptual coding of high quality audio

- Bandwidth extension
  - consists in a reconstruction of a spectral region of the speech or audio signal that is missing
  - specific techniques
    - frequency shifting (transposition)
    - non-linear filtering
    - band replication
  - typical implications
    - system complexity and delay are increased because filter banks are used in addition to that of the core audio coder
    - additional signal artifacts arise due to mismatches between natural audio coding and synthetic audio generation

Accurate Spectral Replacement (ASR)

- Concept
  - bandwidth extension is achieved by means of independent and accurate processing of coherent and and incoherent components of the audio signal

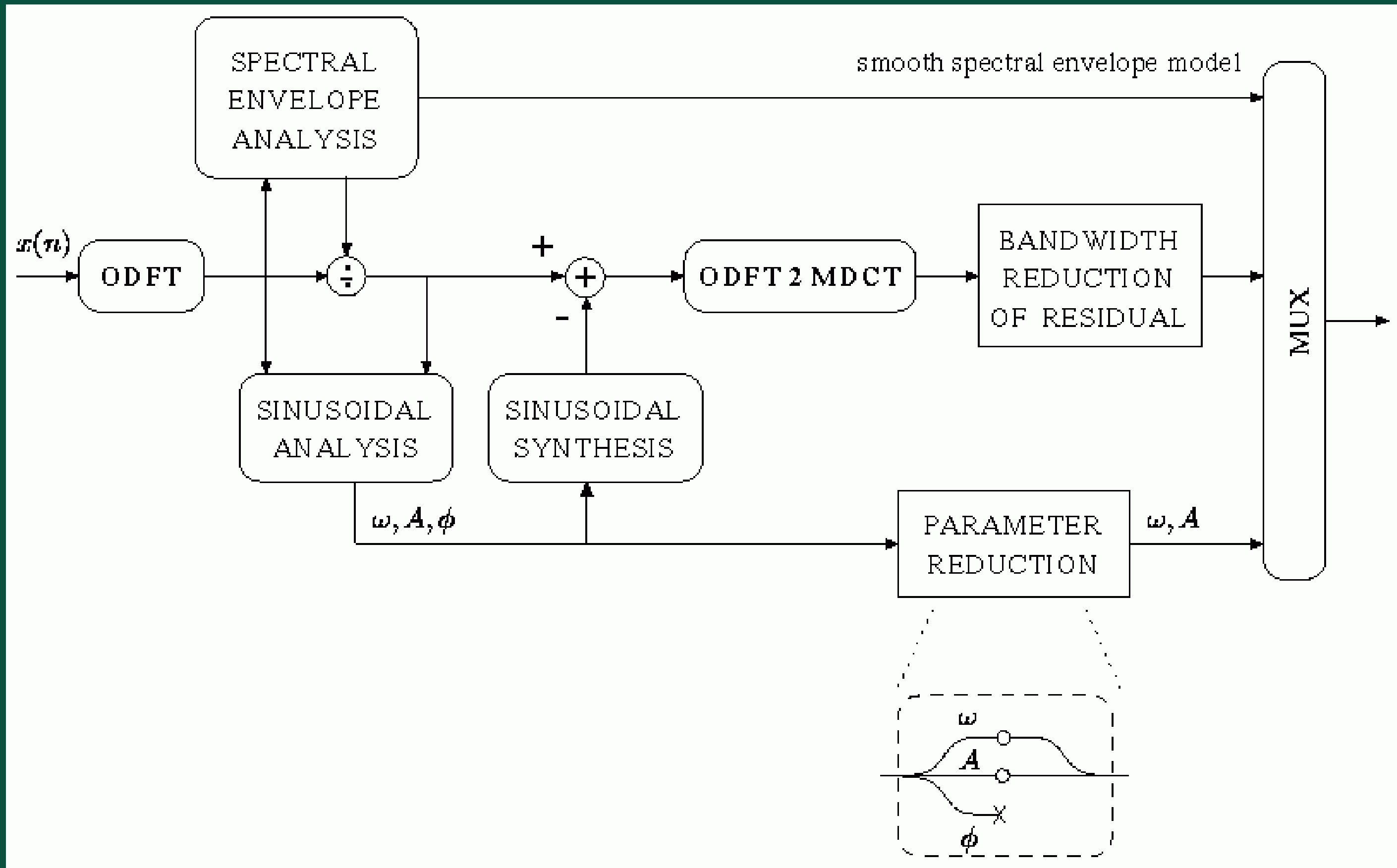


Accurate Spectral Replacement (ASR)

- Concept
  - advantages
    - does not imply other filter banks than the MDCT
    - implements direct synthesis of sinusoids in the frequency domain
    - permits independent control of the spectral tilt of sinusoids and the spectral tilt of stationary noise
  - processing steps
    - signal normalization using a smooth spectral envelope model
    - segmentation of the flattened signal into sinusoids and noise
    - synthesis/bandwidth extension of sinusoids with sub-bin accuray
    - bandwith extension of stationary noise with bin accuracy
    - combination of sinusoids and noise
    - signal denormalization using the smooth spectral envelope model

Accurate Spectral Replacement (ASR)

- ASR encoder



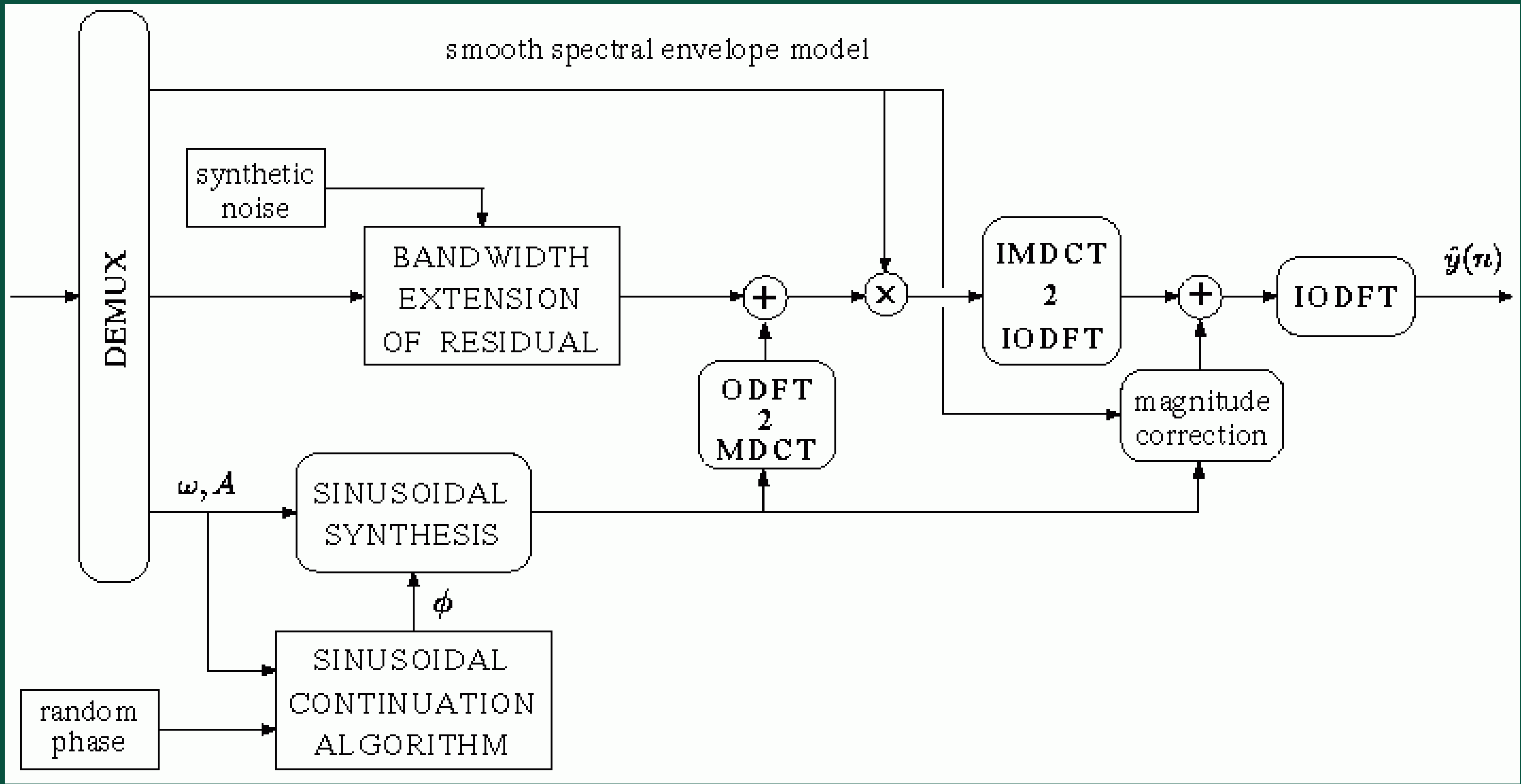
NOTE: this particular encoder implementation discards all phase information and reduces the bandwidth of the stationary noise to about 1/3 of its original bandwidth

Accurate Spectral Replacement (ASR)

- ASR encoder
  - main features
    - smooth spectral envelope model is obtained by short-pass liftering the ODFT real cepstrum
    - the signal segmentation into sinusoids and noise is implemented in an accurate way in the ODFT/MDCT domain
    - sinusoidal components are parametrically represented and can be transmitted to the decoder in a very flexible way by acting selectively on the magnitude, frequency or phase information
    - stationary noise can be transmitted as
      - full band signal
      - bandwidth reduced signal
      - parametric information

Accurate Spectral Replacement (ASR)

- ASR decoder



Accurate Spectral Replacement (ASR)

- ASR decoder
  - main features
    - the bandwidth extension of the MDCT residual may combine transposition of coded noise and synthesis of artificial noise
    - sinusoids can be synthesized either in the (real) MDCT domain or in the (complex) ODFT domain
    - sinusoids can be synthesized either outside the bandwidth of the transmitted residual or within, in which case the perfect reconstruction property can be met (in the absence of quantization of the residual)
    - a sinusoidal continuation algorithm is needed to generate phase for those sinusoids whose phase is not transmitted or is lost
    - sinusoidal synthesis provides sub-bin accuracy which preserves precise harmonic relation between sinusoids

Accurate Spectral Replacement (ASR)

- Results
  - assumptions
    - sampling frequency: 44100 Hz
    - ODFT/MDCT transform size: 1024
    - bandwidth of preserved residual: 6 kHz
    - sinusoidal parameter reduction: encoder discards all phase info
  - test signal: FM modulated sinusoid
    - Matlab code

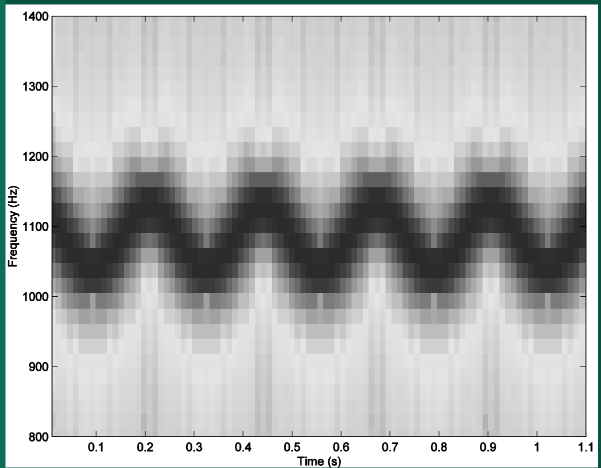
```
samples=512*100;
wave=5000*sin(2*pi*25.1*[0:samples-1]/1024+10.24*sin(2*pi*[0:samples-1]/10240));
```

    - FM carrier: ~ 1081 Hz
    - total maximum frequency deviation: ~ 88 Hz
    - resolution of the analysis/synthesis filter bank: ~ 43 Hz

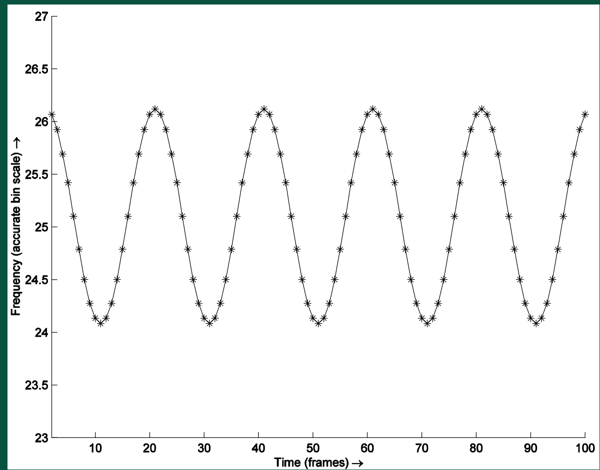
www.atc-labs.com
ajf@atc-labs.com

Accurate Spectral Replacement (ASR)

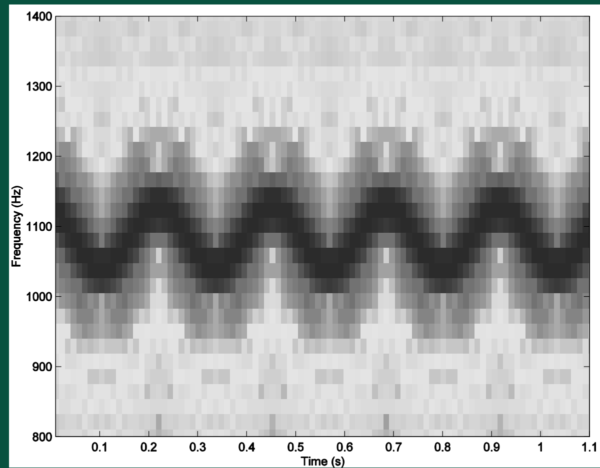
- Results
  - test signal: FM modulated sinusoid



input spectrogram



instantaneous frequency estimation

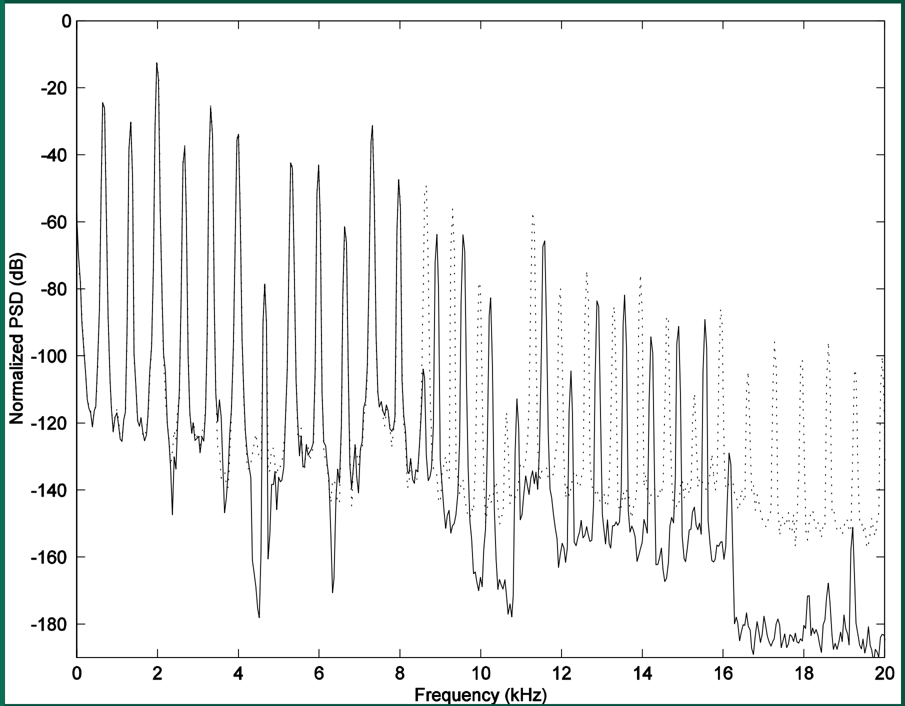


output spectrogram

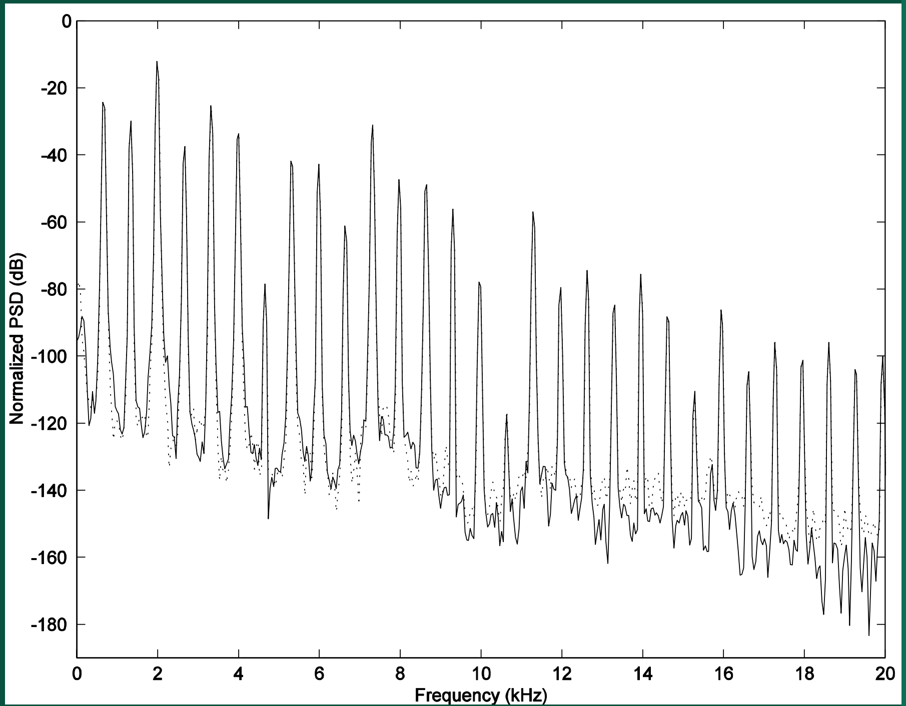
- comments:
  - signal integrity is preserved
  - no signal artifacts are heard

Accurate Spectral Replacement (ASR)

- Results
  - test signal: natural music signal (*pitchpipe*)



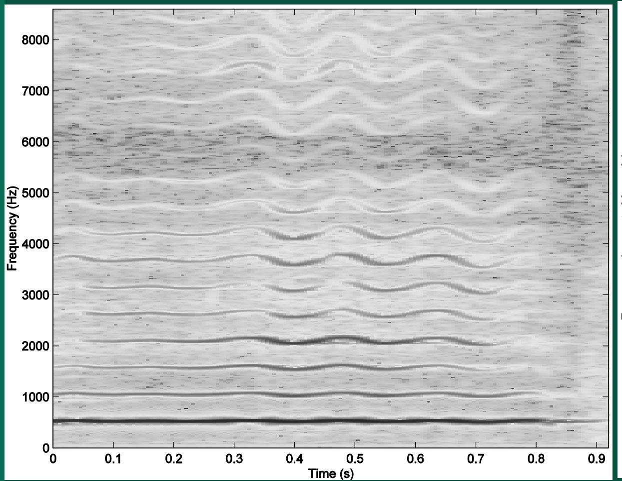
PSD of original (dotted line), and bandwidth extension using MP3+SBR encoding at 64 kbit/s (solid line)



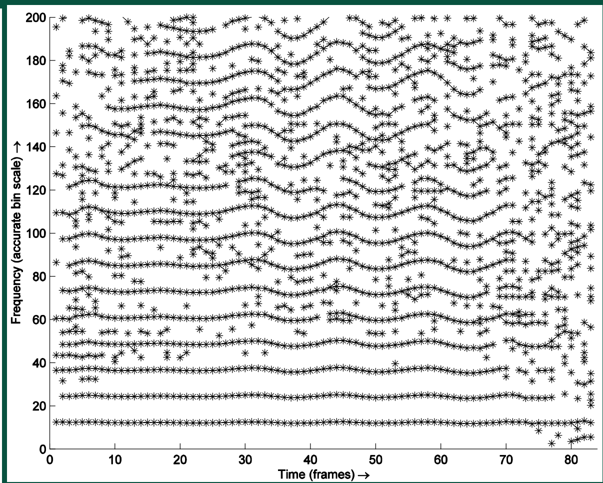
PSD of original (dotted line), and bandwidth extension using ASR processing (solid line)

Accurate Spectral Replacement (ASR)

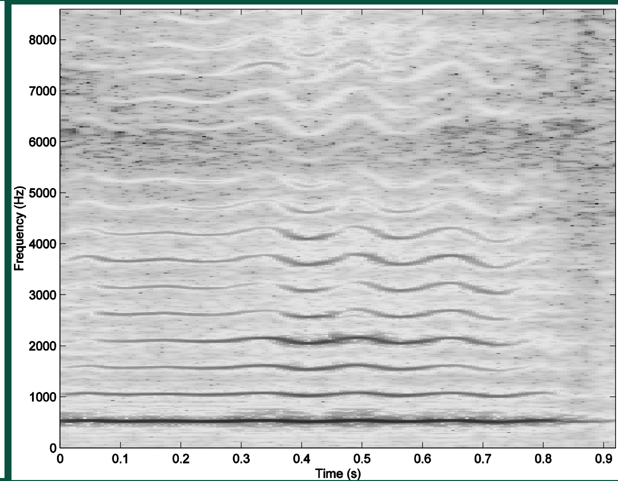
- Results
  - test signal: singing voice



input spectrogram



instantaneous frequency estimation and sinusoidal continuation



output spectrogram

- comments:
  - vibrato effect is preserved
  - subjective quality is very good even if phases are synthetic

Conclusion

– ASR, a new bandwidth extension technique has been described and its performance illustrated
– ASR operates directly in the ODFT/MDCT domain which represents a structural advantage relative to other bandwidth extension techniques
– ASR is very flexible since accurate bandwidth extension/spectral replacement is achieved independently for sinusoids and stationary noise
– the foreseen main application areas include low-delay, low-complexity, and low-bit rate high quality audio communication, special effects in audio (including mutichannel)