Speex: A Free Codec For Free Speech

http://www.speex.org/

Presented by: Jean-Marc Valin

27/01/2006

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Overview

- Introduction to Speex
- Speex and CELP
- Speex features
- Using Speex
- Some samples
- Recent developments and roadmap
- Advocacy

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

What is Speex?

- Audio codec specifically designed for speech and VoIP
  - Can also be used for file compression (Ogg)
- Open-source/Free software (BSD-licensed)
- Designed to avoid patents*
- Developed within the Xiph.Org Foundation
- Included in most Linux distributions
- Provides an alternative to closed, expensive proprietary codecs
- Based on old, reliable CELP technology

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

A Brief History of Speech Codecs

- Pre 1875: Voice over Acoustic Waves
- 1875-1972: Analog telephony
- 1972: G.711 (aka μ-law and A-law)
- 1984: First CELP codec (Schroeder & Atal)
- 1990: GSM Full-Rate (13 kbps, poor quality)
- 1995: Standardisation of G.723.1, G.729 (ACELP)
- 1995-200x: Tons of proprietary speech codecs
- February 2002: Speex project started
- October 2002: Speex joined the Xiph.Org Foundation
- March 2003: Version 1.0 released, bit-stream frozen

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Goals and Design Decisions

- VoIP requirements
  - Frame size and algorithmic delay must be small
  - Encoding and decoding must work with limited resources
  - Minimal distortion when packets are lost
  - Support for narrowband and wideband
  - Support for multiple bit-rates (quality)
  - Achieve good compression while avoiding patents
- The above lead to the choice of CELP
  - Proven at both low and high bit-rate
  - Many patents (not all) have expired
  - Minimise inter-frame dependency
    - Without going as far as iLBC

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

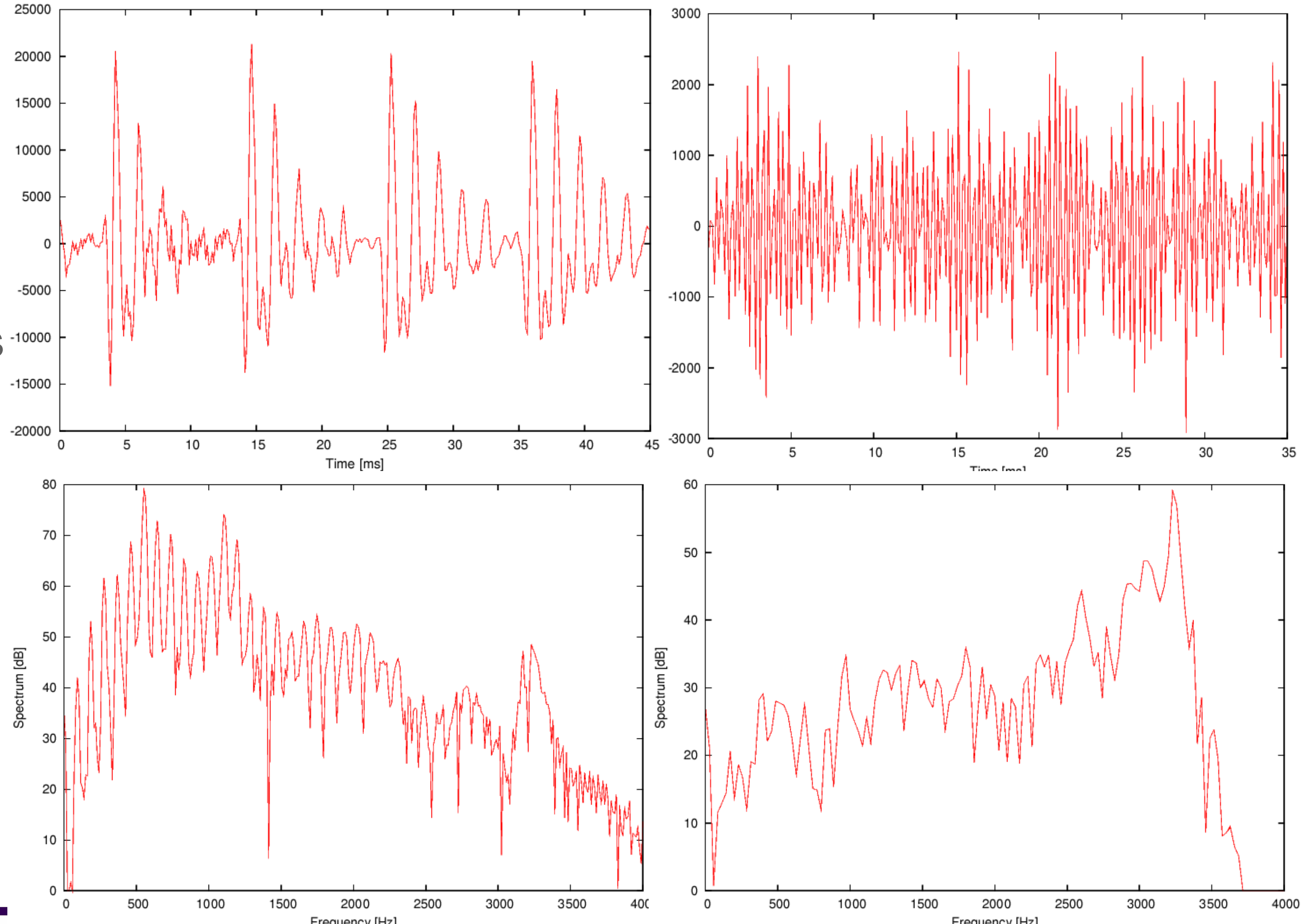
Code-Excited Linear Prediction (CELP)

- First presented in 1984 by Schroeder and Atal and is still the most popular speech coding algorithm
- First version was 100x slower than real-time on a Cray!
- Many variants (ACELP, QCELP, RCELP, LD-CELP, ...) and patents on improvements, mostly standard-specific
- My summary: If you select the right noise and filter it carefully, it may end up sounding like speech
- Main ideas are:
  - Use of linear prediction (LPC), excitation-filter model
  - Perceptual weighting of the noise
  - Analysis-by-synthesis (AbS)
  - Vector quantisation (VQ)

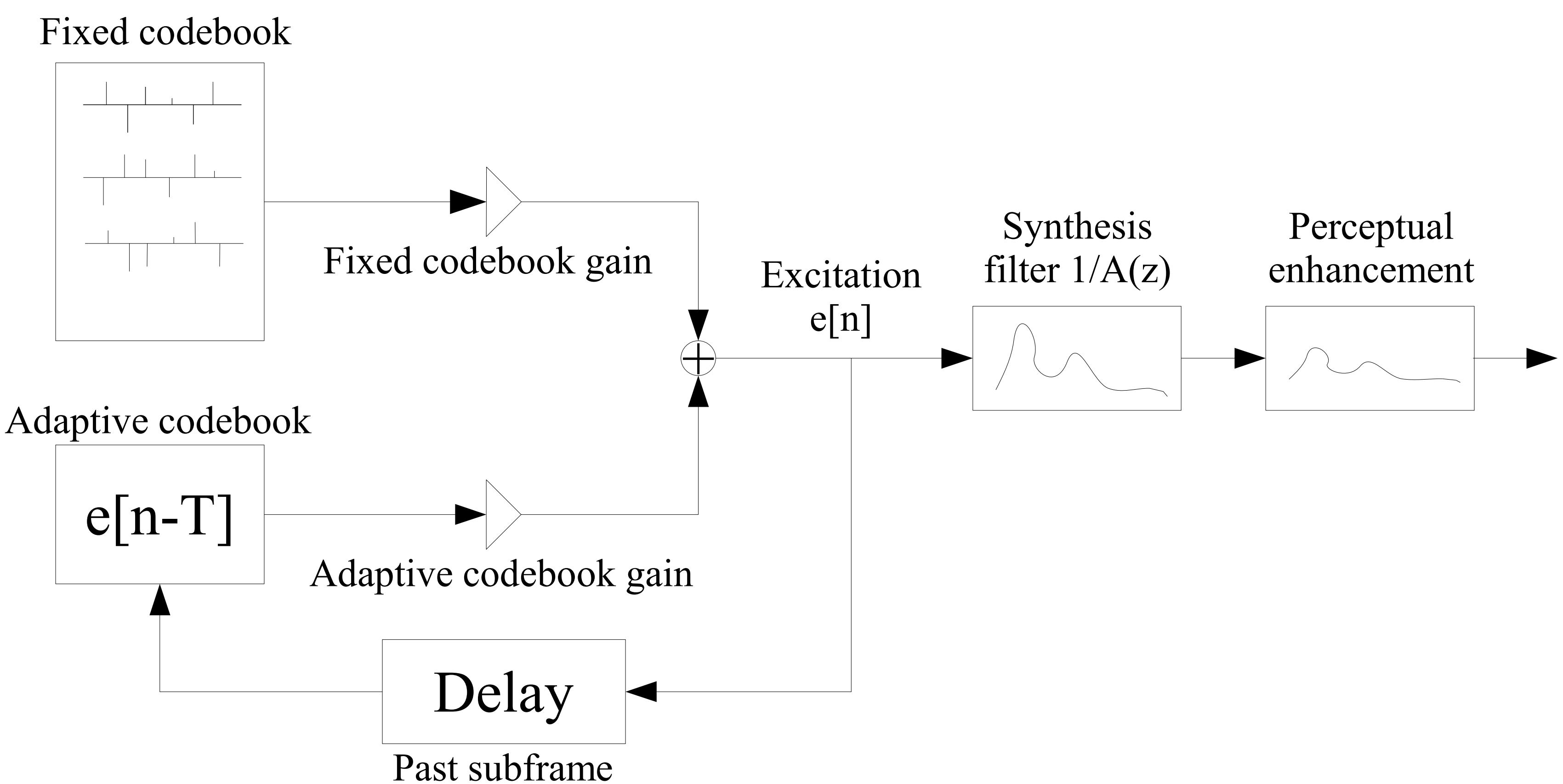
CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Speech Signals

- Voiced speech
  - Periodic
  - Regular, filtered impulses
- Unvoiced speech
  - filtered noise

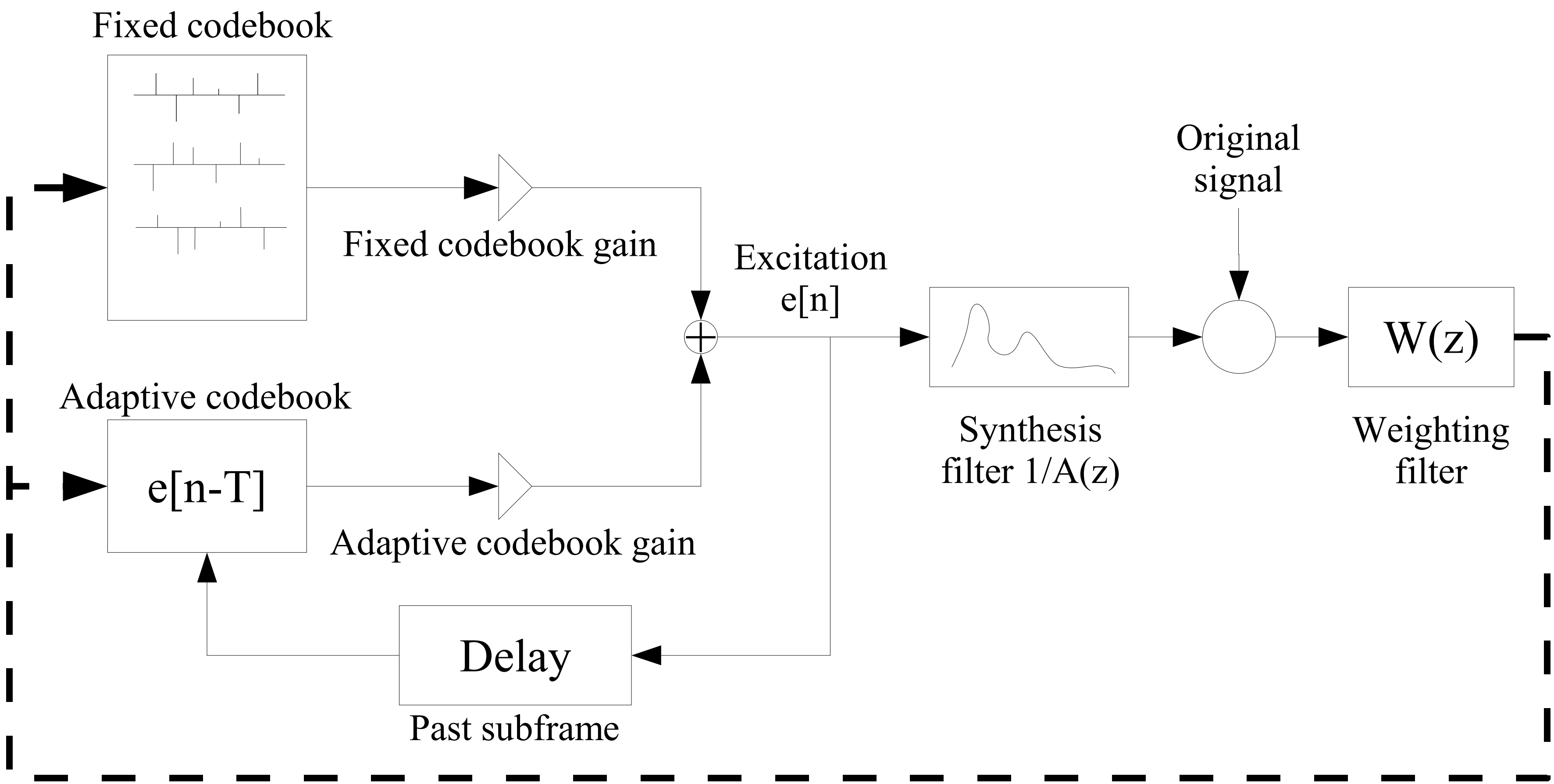


Generic CELP Decoder



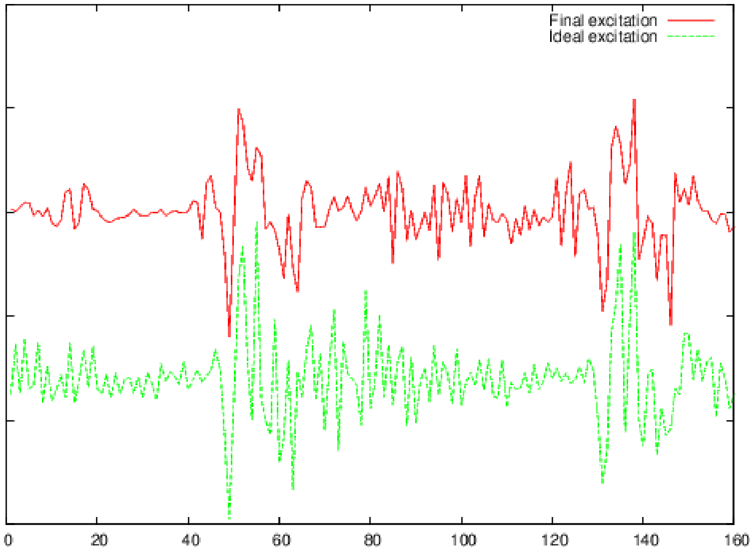
CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

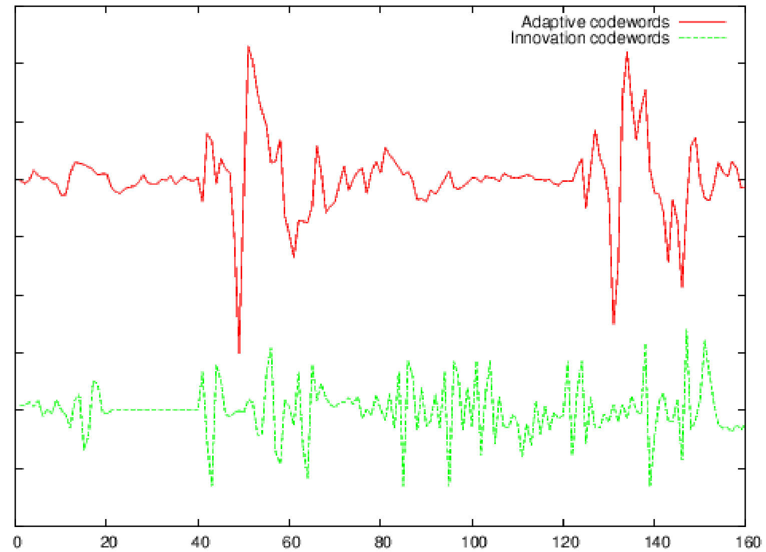
Generic CELP Encoder

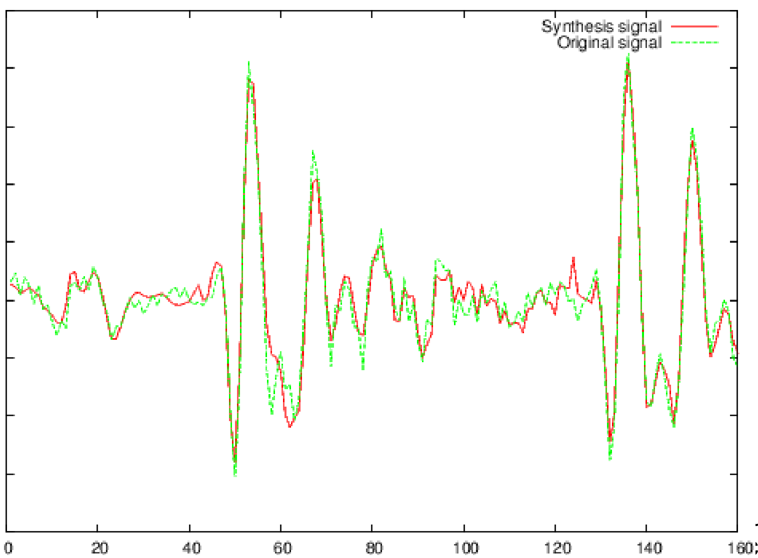


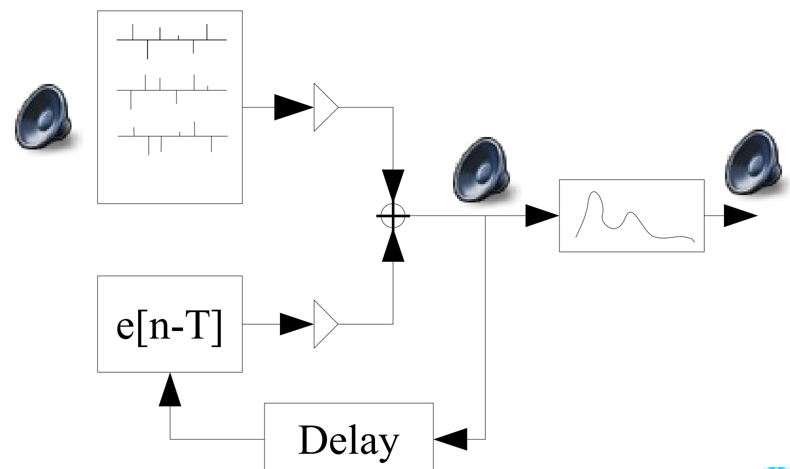
CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Show Me The Signals!









stralian Government through the Advanced Networks Program (ANP) of
ations, Information Technology and the Arts and the CSIRO ICT Centre

Specs

- Bit-rates
  - narrowband: 2.15 – 24.6 kbps
  - wideband: 4 kbps – 42.2 kbps
- Latency
  - narrowband: 30 ms (20 ms frames, 10 ms delay)
  - wideband: 34 ms (20 ms frames, 14 ms delay)
- Features
  - Embedded wideband bit-stream
  - Variable bitrate (VBR)
    - Good for files, bad for VoIP
  - Average bitrate (ABR): VBR with bitrate management
  - Voice activity detection (VAD) and Discontinuous transmission (DTX)

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Implementing Speex Support

- List Requirements
  - How much bandwidth is available?
  - What is the desired quality?
  - What are the latency requirements?
- Choose:
  - Sampling rate
  - Bitrate
  - CBR, VBR, VAD, ...
- Implement using *libspeex*
- Optionally use extra feature (noise suppression, AEC, ...)

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Tips

- Start from sample code
- Make sure to send the right input
  - Use the right format, frame size
  - Remove DC offset (if any), possibly high-pass filter
  - Use correct gain (no clipping, enough dynamic range)
- **Listen** to
  - Input speech
  - Decoded speech
  - Result from speexenc/speexdec
- Handle lost packets (at decoder)

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Narrowband

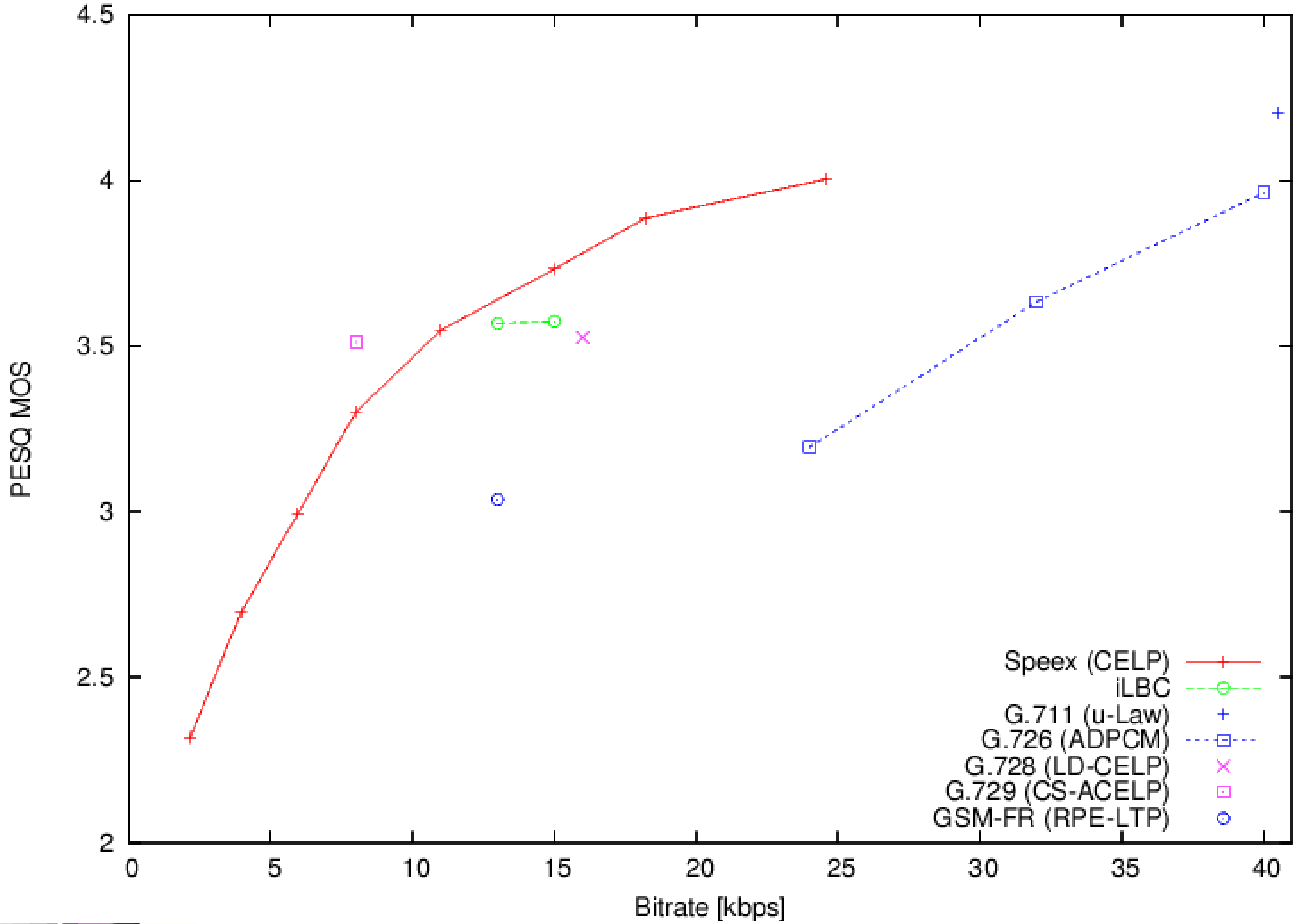
- Sampling rate: 8 kHz (300-3400 Hz effective bandwidth)
- Bit-rates: 2.15 kbps to 24.6 kbps
- Recommended for VoIP: 8 kbps, 11 kbps, 15 kbps
- Samples
  - Original
  - 15 kbps
  - 8 kbps
  - 4 kbps

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Narrowband Evaluation

www.ict.csiro.au

- Results obtained using PESQ (not a real MOS test)

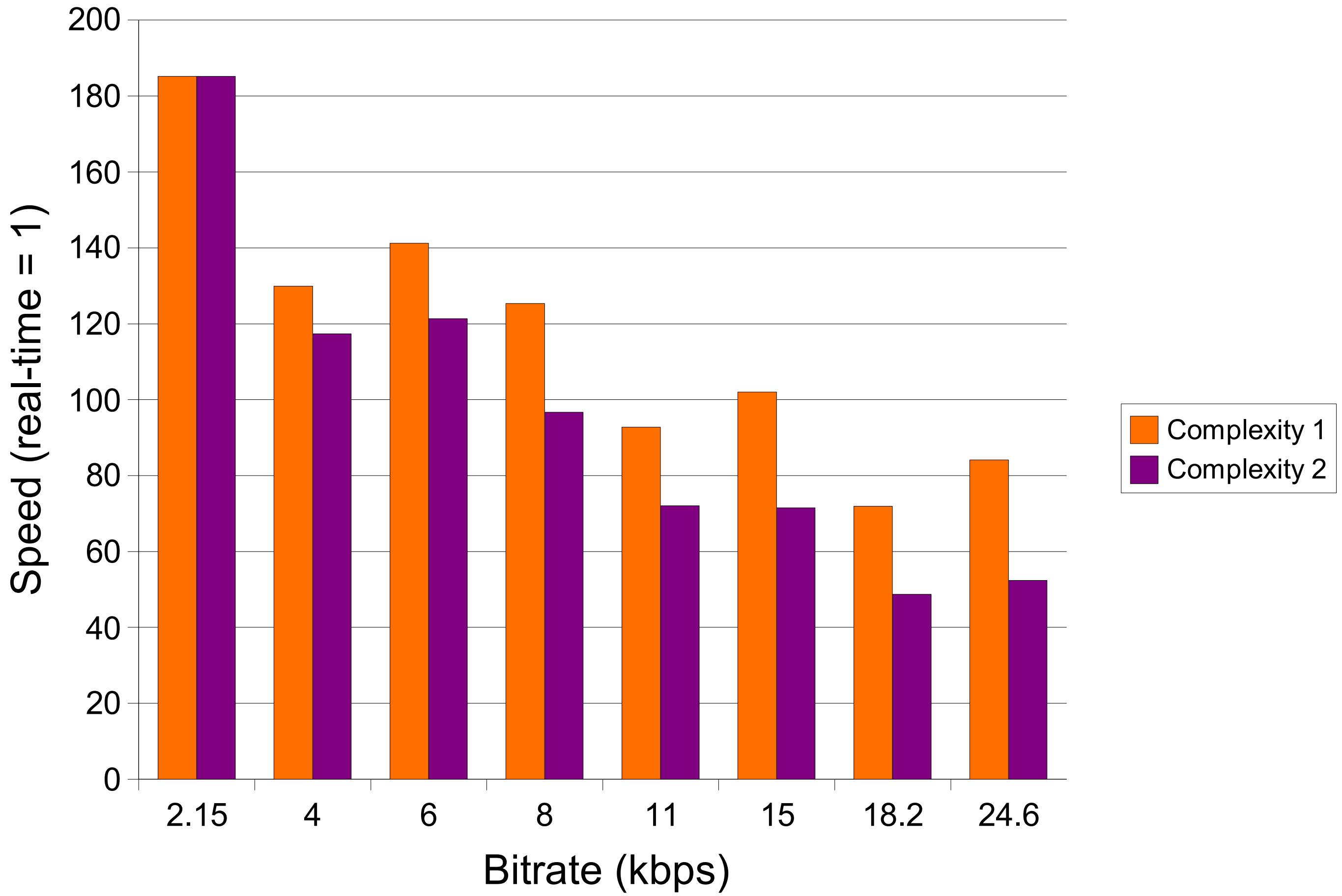


CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

CSIRO

Complexity (Narrowband)

- Encode+decode, SSE enabled on 2.13 GHz Pentium-M



CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Wideband

- Wideband is the future
  - Only way for VoIP to be better than PSTN
  - Not very expensive considering the 16 kbps overhead (IP+UDP+RTP)
- Speex wideband and narrowband are compatible (embedded)
- Recommended for VoIP: 12.8 kbps to 27.8 kbps
- Samples
  - Original
  - 27.8 kbps
  - 20.6 kbps
  - 12.8 kbps
  - 15 kbps narrowband again!
  - 15% packet loss (zero pad)
  - 15% packet loss (Speex PLC)

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Recent Development

- Speex development is still active
- Preprocessor
  - Noise suppression
  - Automatic gain control (AGC)
  - Improved voice activity detection (VAD)
- Acoustic echo cancellation (AEC)
  - Improved hands-free phones
  - Sound from the speaker is subtracted from the microphone (locally)
- Fixed-point

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Fixed-Point

- Speex is being modified so it can optionally use integer arithmetic only (no FPU required)
  - Assumes a 32-bit accumulator and a 16-bit multiplier (result in 32 bits)
  - Quality is very close to float version
- Parts that are fully implemented in fixed-point
  - CBR narrowband modes from 5.95 kbps to 18.2 kbps
  - Echo canceller
- Partially implemented (fast enough with float emulation)
  - All other narrowband bit-rate, VBR, ...
  - Wideband
- Not implemented in fixed-point
  - Preprocessor

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Embedded World

- ARM Architecture
  - Assembly optimisations for ARMv4
  - Some extra optimisations for ARMv5E
  - Can be used with Linux/gcc
- Analog Devices Blackfin
  - Assembly optimisations
  - Free development kit based on μClinux, gcc and Linphone
  - GPL-licensed STAMP development board (http://blackfin.uclinux.org/)
- Texas Instruments C54x, C55x and C6x
  - Known to work (not tested by me)
  - No Free operating system or development tools
  - C54x not recommended for now

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Roadmap

- In progress
  - Speex over RTP IETF draft
  - Porting to fixed-point
  - Speex using the Vorbis psycho-acoustic model
- Possible improvements (volunteers?)
  - Tuning work (perceptual enhancement, noise shaping)
  - Better VAD and VBR
- In the future
  - High-quality, real-time audio codec

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Conclusion (Why Should I Use Speex?)

- Open source
  - No cost for software, no vendor lock-in
  - The codec is still evolving
  - Compatibility with Free Software (even for a proprietary app)
- One codec to rule them all!
  - Supports narrowband and wideband
  - Wide range of bit-rates (2-44 kbps)
  - Very customisable
- Easy
  - Easy to use library
  - Community support
    - Mailing list: speex-dev@xiph.org
    - IRC: irc.freenode.net #speex

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

Questions?

Unofficial OggPCM3 Header Packet

bits	value	Meaning
1	0	Codec identifier. Please make sure no other format starts with a bit set to zero.
16	0x00	Version Major (increment and have fun breaking other people's applications)
16	0x00	Version Minor (should be compatible, be more creative as to how to break stuff)
32	[uint]	PCM format
32	[uint]	Phase of the moon
32	[uint]	Sampling rate in ROT13 format
32	[uint]	Number of channels (make use of all the bits here)
16	[uint]	Number of flames since creation of the spec (if 16 bits aren't enough, steel from next field)
16	[uint]	Number of developers implementing the spec (will go down as previous field increases)
32	[uint]	Favorite colour (RGBA)
1	[bool]	Evil bit. Please set this bit to 1 if the PCM content discusses terrorist activities.
1	[bool]	Clueless bit. Please set this bit to 1 if you don't know what to set it to.
1	[bool]	Wiretap bit. If you are wiretapping stream content. Please alter this bit in the transmission.
1	[bool]	Steganography bit. If set to 1, undetectable information is encoded in the samples' LSB.
16	[uint]	Annoyance field. Rate annoyance of the content from 0 to 65535.
128	[uint]	Magic number. Guess the right magic number or else the file won't play.
8x12	[char]	CC field. Please leave your credit card number here.

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

CSIRO

Software Support

OpenH323
ANNODEX
Sweep
IAXClient
XINE
GnomeMeeting
MPLAYER
Asterisk
The Open Source PBX
Videolan.org
Linphone
gstreamer

CeNTIE2

CeNTIE is supported by the Australian Government through the Advanced Networks Program (ANP) of the Department of Communications, Information Technology and the Arts and the CSIRO ICT Centre

CSIRO