

The Sj32

Copyright

CONTENTS	3
-----------------	----------

Contents

1 Introduction to Speex	6
2 Codec description	8
2.1 Concepts	8
2.2 Codec	10
2.3	10
2.4 Adaptive Jitter Buffer	10
2.5 Acoustic Echo Celler	

1 Introduction to Speex

The Speex project (<http://www.speex.org/>) has been started because there was a need for a speech codec that was open-source⁸⁴ (been)Tessential.078 2 04d (there)condij 561.3894402236(for)Tj 14

This document is divided in the following way. Section 2 describes the different Speex features and defines some terms that will be used in later sections. Section 4 pro

information about the standard con

2 Codec description

This section describes the main features provided by Speex.

One of the

Here are some concepts in speech coding that help better understand the rest of the manual. Emphasis is placed on the Speex features.

First, in


specifying quality, there's no guaranty about the final average bit-rate. Second, for some real-time applications like voice over IP (VoIP), what counts is the maximum bit-rate, which must be low enough for the communication channel.

$$A = \frac{B'}{B} \cdot A_B$$

Average bit-rate solves one of the problems of VBR, as it dynamically adjusts VBR quality in order to meet a specific target bit-rate. Because the quality/bit-rate is adjusted in real-time (open-loop), the global quality will be slightly lower than that obtained by encoding in VBR with exactly the right quality setting to

kHz), the delay is 34 ms. These values don't account for the CPU time it takes to encode or decode the frames.

o c

oc 

This part refers to the preprocessor

- denoising

3 Compiling

Compiling Speex under UNIX or any platform supported by autoconf (e.g. Win32/cygwin) is as easy as typing:

Speex comments

-comment Add the given string as an extra comment. This may be used multiple times.

-author Author of this track.

-title Title for this track.

Raw input options

-rate n Sampling rate for raw input

-stereo

nbBytes is the number of bytes actually written to *byte_ptr* (the encoded size in bytes). Before calling `speex_bits_write`, it is possible to find the number of bytes that need to be written by calling `speex_bits_nbytes(&bits)`, which returns a number of bytes.

When using an unstable release (1.1.x), it

where `audio_frame` is used both as input and output and `echo_residue` is either an array filled by the echo canceller, or NULL if the preprocessor is used without the echo canceller.

In cases where the output audio is not useful for a certain frame, it

5 PROGRAMMING WITH SPEEX (THE LIBSPEEX API)

SPEEX_SET_LOW_MODE*\$PEEX †

SPEEX_GET_ABR* Get average bit-rate (ABR) setting (integer

SPEEX_PREPROCESS_GET_VAD Get VAD status (integer)

SPEEX_PREPROCESS_SET_AGC_LEVEL

SPEEX_PREPROCESS_GET_AGC_LEVEL

SPEEX_PREPROCESS_SET_DEREVERB Turns reverberation removal on(1) or off(2) (integer)

SPEEX_PREPROCESS_GET_DEREVERB Get reverberation removal status (integer)

SPEEX_PREPROCESS_SET_DEREVERB_LEVEL

SPEEX_PREPROCESS_GET_DEREVERB_LEVEL

SPEEX_PREPROCESS_SET_DEREVERB_DECAY

SPEEX_PREPROCESS_GET_DEREVERB_DECAY

c n n n n n n

Sometimes it is desirable to pack more than one frame per packet (or other basic unit of storage). The proper way to do it is to call `speex_encode` N times before writing the stream with `speex_bits_write`. In cases where the number of frames is not determined by an out-of-band mechanism, it is possible

6 FORMATS AND ST

6 FORMATS AND STANDARDS

Field	Type	Size

where $\hat{y}[n]$ is the linear prediction of $x[n]$

8 SPEEX N

8 SPEEX NARROWBAND MODE

9 Speex wideband

A FAQ

o r e o n e c n y o n

Vorbis is a great project but its goals are not the same as Speex. Vorbis is mostly aimed at compressing music and audio in general, while Speex targets speech only. For that

Speex can achieve much better results than Vorbis on speech, typically 2-4 times higher compression at equal quality.

o' n o n e c r k r on o M co c y, e n c e y

First of all, it's not clear whether GSM-FR1
berlin.de/~jutta/toast.html). Also, GSM-FR1

nono

Speex files have the .spx extension. Note, however that the Speex tools (speexenc, speexdec) do not rely on the extension at all, so any extension will work.

nono

Just like Vorbis is not really adapted to speech, Speex is really not adapted for music.
In

'f n n e c o n n o e e n n e n
 f y c on o o

One of the causes could be scaling of the input speech. Speex expects signals to have a $\pm 2^{15}$ (signed short) dynamic range. If the dynamic range of your signals is too small (e.g. ± 1.0), you will suffer important quantization noise. A good target is to have a dynamic range \pm

n e o n co c y

If I could do that I'd be very rich by now :-) Seriously, that would break fundamental laws of information theory.

 , yo n.M c ,on , n y o ,
oo n o o ,n o

Currently


```
inFile = argv[1];
fin = fopen(inFile, "r");

/*Initialization of the structure that holds the bits*/
speex_bits_init(&bits);
while (1)
{
    /*Read a 16
```

```
    return 0;  
}
```

B  **c c**

sampledec reads a Speex stream from **stdin**, decodes it and

```

fout = fopen(outFile, "w");

/*Initialization of the structure that holds the bits*/
speex_bits_init(&bits);
while (1)
{
    /*Read the size encoded by sampleenc, this part will likely be
       different in your application*/
    fread(&nbBytes, sizeof(int), 1, stdin);
    fprintf (stderr, "nbBytes: %d\n", nbBytes);
    if (feof(stdin))
        break;

    /*Read the "packet" encoded by sampleenc*/
    fread(cbits, 1, nbBytes, stdin);
    /*Copy the data into the bit-stream struct*/
    speex_bits_read_from(&bits, nbBytes);

    /*Decode the data*/
    speex_decode(&bits, state,
                output);

    /*Copy from float to short (16 bits) for output*/
    for (i=0;i<FRAME_SIZE;i++)
        out[i]=output[i];

    /*Write the decoded audio to file*/
    fwrite(out, sizeof(short), FRAME_SIZE, fout);
}

/*Destroy the decoder state*/
speex_encoder_destroy(speex_encoder);
destroy the bit-stream

```

C IETF RTP Profile

AVT Working Group

Internet-Draft

Expires: October 3, 2005

G. Herlein

S. Morlat

J. Jean-Marc

R. Hardiman

P. Kerr

April 04, 2005

draft-herlein-speex-rtp-profile-02
RTP Payload Format for the Speex Codec

Status of this Memo

This document is an Internet-Draft and is subject to all provisions of section 3 of RFC 3667. By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she become aware

The list of Internet-Draft Shadow Directories can be

Herlein, et al. Expires October 3, 2005 [Page 2]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

1. Conventions used in this document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [1].

2. Overview of the Speex Codec

Speex is based on the CELP [10] encoding technique with support for either narrowband (nominal 8kHz), wideband (nominal 16kHz) or ultra-wideband (nominal 32kHz), and (non-optimal) rates up to 48 kHz sampling also available. The main characteristics can be summarized as follows:

- o Free software/open-source
- o Integration of narrowband and wideband in the same bit-stream
- o range of bit-rates available
- o Dynamic bit-rate switching (and) variable

+++++
| one or more frames of

specification is two [2].

C IETF RTP PROFILE

50

Padding (P): 1 bit

If the padding bit


```

|V=2|P|X|  CC  |M|      PT      |      sequence number      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|                                timestamp                        |
+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+
|
|      synchronization source (SSRC) identifier                |
+===+===+===+===+===+===+=== 118.697 0 Td (+-+)=+

```

Herlein, et al. Expires October 3, 2005 [Page 6]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

```

|      synchronization source (SSRC) identifier      |
+=====+
|      contributing source (CSRC) identifiers          |
|      ...                                             |
+---+---+---+---+---+---+---+---+---+---+---+---+---+
+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      ..speex data..                                |
+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      ..speex data..      |      ..speex data..      |
+---+---+---+---+---+---+---+---+---+---+---+---+---+
|      ..speex data..                                |
+---+---+---+---+---+---+---+---+---+---+---+---+---+

```

8. MIME registration of Speex

Full definition of the MIME [3] type for Speex will be part of the

Og68t.112.7Vorbis((CSRC916)Tj 23.995 0 Td (MIME)Tj 29.7895 0 Td (tyME)Tj 29.9217 0 Td (defi

Security Considerations:

See Section 6 of RFC 3047.

Interoperability considerations: none

Publhet646

Examples:

```
m=audio 8008 RTP/AVP 97
a=rtpmap:97 speex/8000
a=fmtp:97 mode=4
```

This examples illustrate an offerer that wishes to receive a Speex stream at 8000Hz, but only using speex mode 3.

The offerer may suggest to the remote decoder to activate its perceptual enhancement filter like this:

```
m=audio 8088 RTP/AVP 97
a=rtpmap:97 speex/8000
a=fmtp:97 penh=1
```

Several Speex specific parameters can be given in a single a=fmtp line provided that they are separated by a semi-colon:

Herlein, et al. Expires October 3, 2005 [Page 9]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

```
a=fmtp:97 mode=any;penh=1
```

The offerer may indicate that it wishes to send variable bit rate frames with comfort noise:

```
m=audio 8088 RTP/AVP 97
a=rtpmap:97 speex/8000
a=fmtp:97 vbr=on;cng=on
```

c

Herlein, et al. Expires October 3, 2005 [Page 10]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

Speex can find a mutual codec.

12. RTP Payload Types

discard packets from undesired sources, but the processing cost of the authentication itself may be too high.

14. Acknowledgments

The authors would like to thank Equivalence Pty Ltd of Australia for their assistance in attempting to standardize the use of

Terminal Equipment", ITU-T Recommendation H.245.

[7] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video

Herlein, et al. Expires October 3, 2005 [Page 12]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

Conferenad247.7061 .75

France

EMail: simon.morlat@linphone.org

Jean-Marc Valin
Department of Electrical and Computer Engineering
University of Sherbrooke
2500 blvd Universite
Sherbrooke, Quebec J1K 2R1
Canada

EMail: jean-marc.valin@hermes.usherb.ca

Roger Hardiman
49 Nettleton Road
Cheltenham, Gloucestershire GL51 6NR
England

EMail: roger@freebsd.org

Herlein, et al. Expires October 3, 2005 [Page 13]

Internet-Draft draft-herlein-speex-rtp-profile-02 April 2005

Phil Kerr
England

EMail: phil@plus24.com

Herlein, et al.

Expires October 3, 2005

[Page 14]

D Speex License

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of the Xiph.org Foundation nor the names of its contributors

may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS (AND CONTRIBUTORS) "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE FOUNDATION OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

E GNU Free Documentation License

Version 1.1, March 2000

Copyright (C) 2000 Free Software Foundation, Inc. 59

is in

E GNU FREE DOCUMENTATION LICENSE

E GNU

- J. Preserve the network location, if any, given in the Document for public access to a T

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.



You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of the title (the)T 10.16860Tj 3d (the)0.294 0 Td (in)Tj 10.2708 0 Td parentheses,

to the present version, but may differ in detail to address new problems or concerns.

See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. Document specifies that a numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.

