Polycom® Siren14™

Information for Prospective Licensees

Jeff Rodman, mailto:jeff.rodman@polycom.com

Polycom, Inc.

Revision 3: 2009-02-20

1.0 Overview

Siren14 is an extension of ITU-T Recommendation G.722.1, Polycom's patented "Siren7 ™" audio coding algorithm ("Siren7"), which was approved by ITU as an international standard in 1998. The Annex C extension which standardized Siren14 was formally approved by ITU on 14 May 2005.

Siren14 features very high audio quality and extremely low computational complexity compared to other super-wideband algorithms, which is critical for applications where other tasks (such as video coding) consume most of the available computing resources, or where low-cost or low-power is important.

The Siren14 algorithm is based on the same transform coding used in G.722.1, a Modulated Lapped Transform (MLT) which makes use of 28 sub-bands, each covering 500 Hz.

Siren14 is a transform codec, not a speech-model based codec, although its design is optimized for speech performance. The codec is also suitable for music, although it is not optimized for this.

Polycom offers a unique licensing program for Siren14 – *royalty-free* licenses to all parties for use of Siren14 (the Annex C mode of G.722.1).
The differences between Siren7 and Siren14 are summarized in Table 1:

Table 1 – Siren7 and Siren14

Algorithm	**Siren7**	**Siren14**
ITU standard	**ITU-T Rec. G.722.1**	**ITU-T Rec. G.722.1 Annex C**
Bitrates	16[1], 24, and 32 kbps	24, 32, and 48 kbps
Sample rate	16 kHz	32 kHz
Audio bandwidth	50 Hz to 7000 Hz	50 Hz to 14,000 Hz
Frame size	20 milliseconds	20 milliseconds
Latency	40 milliseconds algorithmic	40 milliseconds algorithmic
Complexity	< 5.5 WMOPS for encoder + decoder (exceptionally low)	< 11 WMOPS for encoder + decoder (exceptionally low)
Memory	11.0 kBytes RAM[2], 20.4 kBytes ROM[3]	18.0 kBytes RAM[3], 29.5 kBytes ROM[4]
Licensing	***Royalty-free*** licenses offered (see details below)	***Royalty-free*** licenses offered (see details below)

2.0 Complexity

Table 2 below presents the complexity of Siren14 in units of Weighted Million Operations Per Second (WMOPS), an ITU-T measure of computational complexity, similar to MIPS. WMOPS are roughly equivalent to MIPS for fixed-point processors. Polycom supplies both fixed-point and floating-point C code implementations to licensees.

Siren14's low complexity is a major technical advantage compared to other algorithms with similar performance in this bitrate range. It represents a separate class of codec for low-complexity applications.

Table 2 - Siren14 (G.722.1 Annex C) and Siren7 (G.722.1) Complexity - WMOPS

Codec/bitrate	**Encoder**	**Decoder**	**Enc+Dec**
Siren14/24 kbps	4.5	5.3	9.7
Siren14/32 kbps	4.8	5.5	10.3
Siren14/48 kbps	5.1	5.9	10.9
Siren7/24 kbps	2.3	2.7	5.0
Siren7/32 kbps	2.4	2.9	5.3
Siren7/16 kbps	2.2	2.6	4.8

[1] The Siren7 16 kbps rate is not part of ITU-T G.722.1, but is available to Polycom Siren7 licensees.

[2] RAM is used for temporary variables and arrays used by internal procedures of the codec.

[3] ROM is used for constant tables.

[4] The AMR-WB+ and eAAC+ WMOPS values were obtained by measuring the fixed-point implementations on the 3GPP website as of April 2005. In each case the 32 kbps mode is reported; the 24 kbps modes are 2% to 5% less complex. The other values are based on published figures.

For comparison, Table 3 below shows the complexity of Siren14 vs. four well-known super wideband codecs, 3GPP AMR-WB+, 3GPP eAAC+[4], MPEG-4 AAC-LC ("Low Complexity"), and MPEG-4 AAC-LD ("Low Delay"). Siren14 is between 1/4 and 1/6 the complexity of these other codecs.

Table 3 - Siren14 Complexity Compared (encoder+decoder)[5]

	Polycom Siren14 / ITU-T G.722.1 Annex C	**3GPP AMR-WB+**	**3GPP eAAC+**	**MPEG-4 AAC-LC**	**MPEG-4 AAC-LD**
WMOPS	10.9 WMOPS	63.4 WMOPS	41.1 WMOPS		
TI TMS320C64	< 5.6 MHz[6]			34.3 MHz	98.1 MHz
Equator BSP-15	<= 14.3 MHz				
Philips TriMedia (floating-point)	<= 7.6 MHz[7]	62.5 MHz			

3.0 Quality

Polycom has made available a PC executable version of Siren14 so potential licensees may evaluate the quality themselves on their own test material and under their own conditions, and can confirm for themselves the adequacy of the quality for the intended applications. (See section 5 below.) Polycom has been offering a proprietary version of the Siren14 codec as the premier audio coding mode in our video conferencing gear for some time – users overwhelmingly prefer this mode to G.722.1 or G.722.

In March 2005, as a part of the G.722.1 Annex C approval process in the International Telecommunications Union (ITU), audio performance characterization tests were performed on Siren14 by an independent test lab according to a test plan designed by the ITU Speech Quality Experts Group.

The testing was in two phases – the first for speech quality using the MOS method, and the second for music and mixed signals using the MUSHRA method. In each phase the ITU testing used the MPEG-4 AAC-LD codec as the reference codec for comparison. The ITU requirement was that the algorithm to be used for G.722.1 Annex C be proven "not worse than the reference" with a 99% statistical confidence level. In addition, for purposes of comparison ITU also tested Siren14 against the 3GPP codecs AMR-WB+ and eAAC+ at the rates of 24 and 32 kbps (these were not tested at 48 kbps).

When comparing codecs, the ITU tests can produce results that the codec being tested is *better than*, *not worse*, or *worse than* the reference codec. "*Not worse*" means the codec performance was too close to distinguish a statistically meaningful difference.

---

5 All Siren14 complexity values are for the most-complex mode (48 kbps). The 32 kbps rate is ~ 5% less complex, and the 24 kbps rate is ~ 11% less complex (see Table 2).

6 Approximated by doubling G.722.1.

7 Estimated based on an actual value of 6.8 MHz for the 24 kbps rate.

3.1 ITU Test Results – Mean Opinion Score (MOS) for Speech Quality

Phase 1 focused on speech quality in video and tele-conferencing applications, and used the Mean Opinion Score (MOS) method[8][8],.[9]

The testing conditions included "clean" (noise-free) speech, and speech with various kinds of background noise including "interfering talkers" (noise which might be experienced by someone in a telecommunication conference in an open area), and "office noise" (typical of the background noise encountered in modern office environments).

In all Phase 1 testing at 24 and 32 kbps, Siren14 was *better than* MPEG-4 AAC-LD, and *not worse* than 3GPP eAAC+. At 32 kbps, Siren14 was *not worse* than 3GPP AMR-WB+.
In the testing at 48 kbps, Siren14 was *not worse* than MPEG-4 AAC-LD operating at either 48 or 64 kbps.

3.2 ITU Test Results – MUSHRA testing on Music and Mixed Signals

Phase 2 of the ITU Characterization Test focused on music and mixed signals using the "Multi Stimuli with Hidden Reference and Anchor points" (MUSHRA) method.[1][10]
The results were very similar – Siren14 was shown to be *better than* MPEG-4 AAC-LD at all bitrates (and 48 kbps Siren14 was also *better than* AAC-LD operating at 64 kbps).
However we note that in practical terms, all the codecs tested by ITU provide very good audio quality; in most cases users will not be able to tell the difference between them. Siren14's more important advantages are its low computational complexity and royalty-free licensing terms.

4.0 Demonstration files

A ZIP file, "**Siren14 Demos.zip"** is available from Polycom's Web site which demonstrates the quality of Siren14 audio (note: this file is over 26 MBytes in size).

The original files "VTX" and "PLCM" originate from Polycom, and these may be used freely. The other files are originally from the EBU via http://sound.media.mit.edu/mpeg4/audio/sqam and may be used "only for the testing and evaluation of sound systems". See the referenced web page for details.

All files with names ending in "_original.wav" are original files, and have not been compressed. The files with names ending in "_24.wav", "_32.wav", and "_48.wav", have been coded and decoded using Siren14 at 24, 32, and 48 kbit/s respectively.

---

[8] Note that MOS score results are affected by the other references in the test; as a result MOS scores can be meaningfully compared only within the same test. MOS is not an absolute measure of quality.

[9] The full Phase 1 results are reported in ITU-T WP3/16 document AC-05-29, "*Characterisation test results of the 14khz low-complexity audio coding algorithm at 24, 32, and 48 kbps extension to ITU-T G.722.1: phase 1*", Strasbourg FR, April 2005.

[10] The full Phase 2 results are reported in ITU-T WP3/16 document AC-05-30, "*Characterisation test results of the 14khz low-complexity audio coding algorithm at 24, 32, and 48 kbps extension to ITU-T G.722.1: phase 2*", Strasbourg FR, April 2005.

5.0 Demonstration software

A PC-executable version of the Siren14 codec is freely available on Polycom's Web site as "**Siren14PCExecutable.zip**". This will encode and decode audio in a batch mode for the purpose of demonstrating the quality of Siren14.

6.0 Source code

Licensees will receive C source code for both a fixed-point and a floating-point implementation of Siren14. The fixed-point software is a part of the formal specification of ITU-T Recommendation G.722.1 Annex C. The floating-point software has been proposed by Polycom to ITU as Annex D/G.722.1, but at this time is unofficial.

The fixed-point and floating-point versions provide identical audio quality, but do not produce bitexact identical output for the same audio input samples. The fixed-point and floating-point implementations are fully interoperable with each other.

Licensees may modify the software to optimize it for their platforms, integrate with other software in a product, etc., provided that they retain conformance with ITU-T Rec. G.722.1 Annex C and interoperability with the supplied reference software.

Note that the Siren14 software source code also implements Siren7 (G.722.1 baseline). Siren7/Siren14/G.719 royalty-free license does include the right to use Siren7.

7.0 Licensing

Our belief is that all patents on **SIREN7/SIREN14/G.719** are owned by Polycom Inc. The algorithm is identical to that of G.722.1, which has been in the market since 1998, and as of this writing no other party has claimed otherwise, even informally.
Polycom offers *royalty-free* licenses to implement, use, and sell the **SIREN7/SIREN14/G.719** to all parties on a nondiscriminatory basis. Included with this license are royalty-free implementations, in both fixedpoint and floating-point C code. This license is with the "Siren7" baseline G.722.1 wideband mode.

The royalty-free license requires that:

a) Credit is given to Polycom Inc. for the algorithm in all specifications, marketing literature, etc. which mentions the audio codec. This credit will be in the form of a description of the codec as "Polycom® Siren14™" (licensees may also use other names, such as the name of the standard). (See section Exhibit B of the license.)
b) Licensee agrees to a provision that the license terminates if the Licensee makes any patent assertion against Polycom. (See section 2.9 of the license.)

The license contains other important provisions as well; please read it carefully.