MUSHRA Listening Tests
Focusing on Stereo Voice Coding

Christian Hoene, Mansoor Hyder
August 28th, 2011

Questions Addressed in These Tests

Can Opus compress

- stereo voice in the Silk and the Hybrid mode?
- two simultaneous voices
- binaural content?

How does Opus perform compared to other stereo voice codecs?

- No open source stereo voice codec available
- thus, used AMR-WB+

Measurement Methodology

Following MUSHRA ITU-R BS.1384-1
Using software “rateit” version 0.1
with modifications and German translation
Analysis and summaries using software “rateit.parse”

Headphones (Sennheiser ABC)
Sound card: PC Dell DEF

Participants were not informed
about the presence of hidden references

Reference Items

1. One Voice Stereo
   8s, stereo voice recording, female German speakers
2. Two Voices Stereo
   9s, two stereo female voices mixed together
3. One Voice Binaural
   13s, one female voice, rendered with HTRF and added room impulse response, moving
4. Two Voice Binaural
   13s, two female voices at different stationary positions, rendered with HTRF and added room impulse response
5. Acappella Song „Mein Fahrrad“ by „Die Prinzen“
   10.5s, mono

Degraded Items 1/2

draft-ietf-codec-opus-07/test_opus

opus.12k SILK, 12kbps, stereo, 60ms
Args.: 0 48000 2 12000 -cbr -framesize 60 -bandwidth NB

opus.16k SILK, 16kbps, stereo, 20ms
Args.: 0 48000 2 16000 -cbr -framesize 20 -bandwidth WB

opus.32k HYBRID, 32kbps, stereo, 20ms
Args.: 0 48000 2 32000 -cbr -framesize 20 -bandwidth FB

opus.64k CELT, 64kbps, stereo, 20ms
Args.: 1 48000 2 64000 -cbr -framesize 20 -bandwidth FB

Degraded Items 2/2

AMR-WB+ using 26304_ANSI-C_source_code_v6_6_0

amrwbp.12k 12kbps, 80ms
Args.: -rate 12
amrwbp.16k 15.2kbps, 80ms
Args.: -rate 16
amrwbp.32k 32kbps, 60ms
Args.: -rate 32

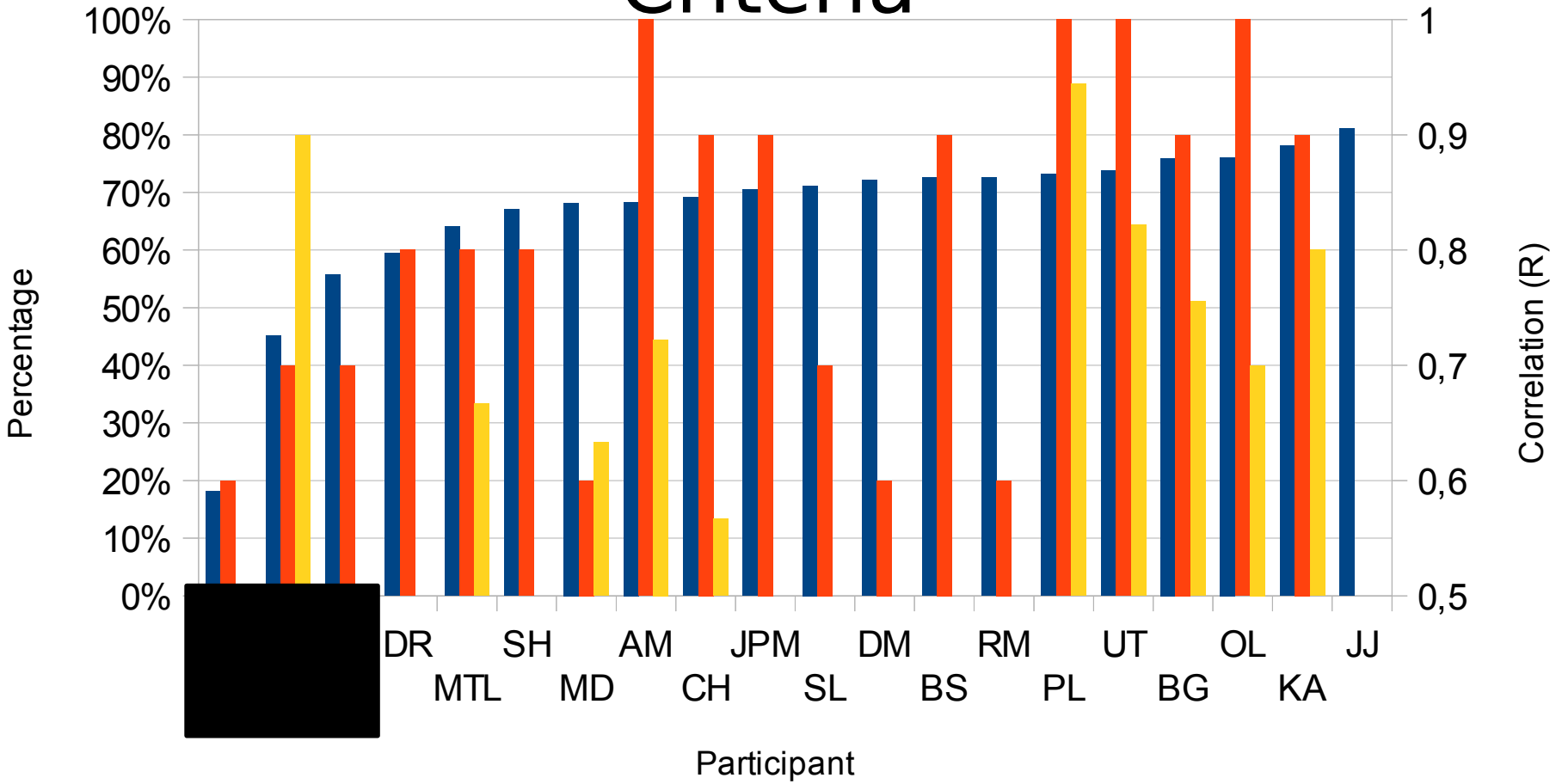
Anchor lowpass 3.5k mono
Args.: sox in.wav -r48000 -c1 out.wav lowpass 3500

Participants

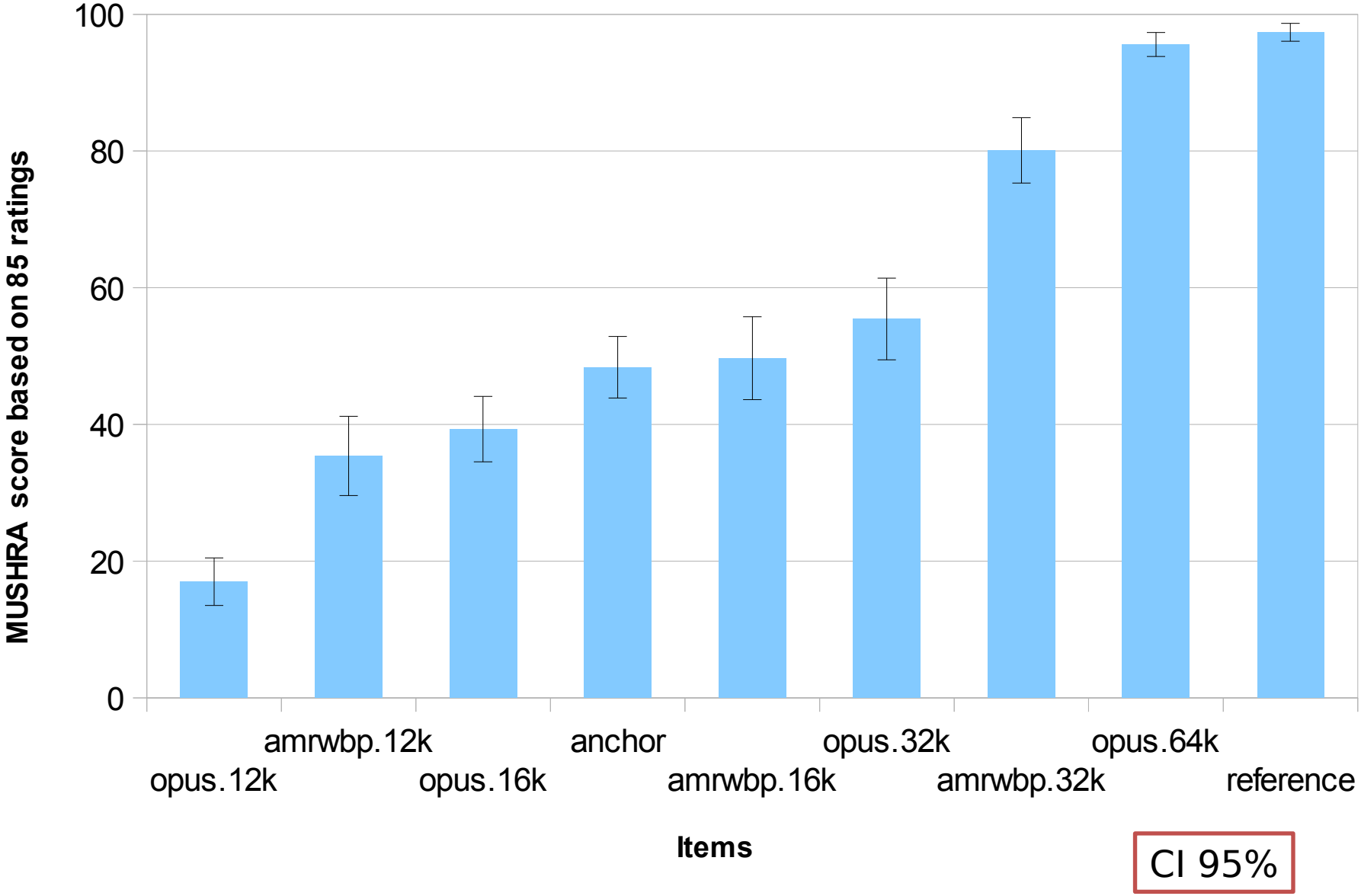
20 German native speakers
Age: between 20 and 59,
Avg. Age: 30.55
9 male, 11 female
All have academic backgrounds

Quality of the individual ratings was verified with Correlation (R) between individual ratings and averaged ratings
If $R \geq 0.8$, individual ratings are good enough.
The results of three participants were removed.

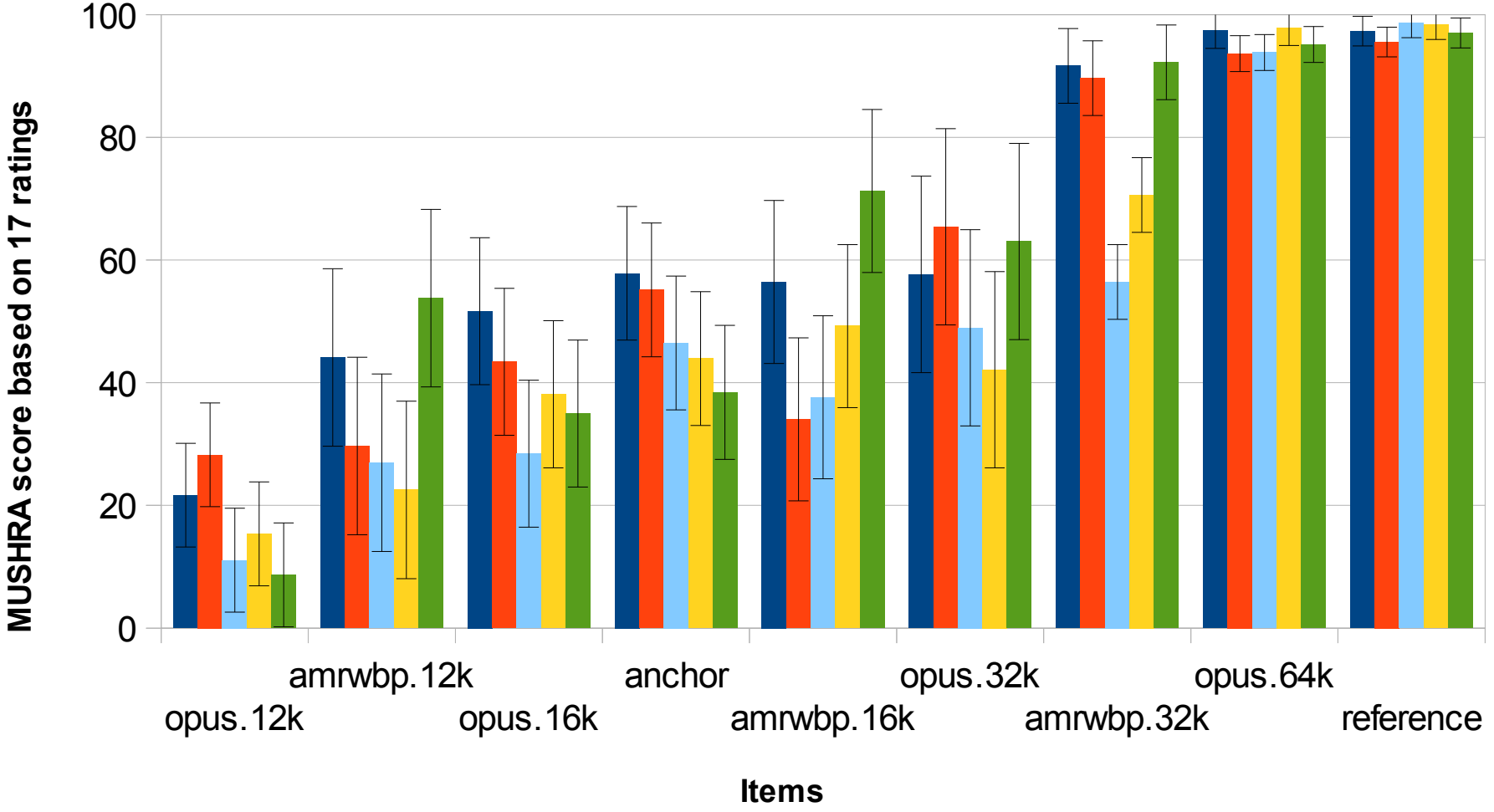
Participants: Different Quality Criteria



Results: Codecs



Codec and Item



Summary

For stereo voice in wideband quality, Opus needs 16kbps in Silk mode

Two (or more) voices are compress ok at 32kbps in the hybrid mode

Binaural contents is only well compressed at 64kbps with CELT

AMR-WB+ (at 80ms framesize) is better than Silk/Hybrid (at 20ms)
AMR-WB+ cannot compress binaural content well.