INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC1/SC29/WG11
MPEG2006/N8043-Annex4
March 2006, Montreux, Switzerland

Source Video

Title **Annex4 to Status Report on Wavelet Video Coding Exploration**

1. Introduction

This part of the document N8043 provides results and a description of the subjective quality evaluation for the wavelet video coding activity that was performed at the 76th MPEG Meeting in Montreux.

2. Test scenario

As the main interest of MPEG in this area lies in combined scalability video coding it was decided to perform an evaluation for this special case of scalability. In addition it was agreed to perform an evaluation for SNR scalability only.

Given are the tables from [1] and [2] pointing out the test points for these two test cases.

Sequence	Format	Bit rates (kbit/sec)						
City	QCIF 15Hz	64	80	96	120	144	168	192
	CIF 30Hz	256	320	384	480	576	672	768
	4CIF 60Hz	672	848	1024	1280	1536	1792	2048
Crew, Harbour, Soccer	QCIF 15Hz	96	120	144	180	216	252	288
	CIF 30Hz	384	480	576	720	864	1024	1152
	4CIF 60Hz	1024	1280	1536	1920	2304	2688	3072

Table 1: SNR scalability test points, evaluated were the test points marked green

	QCIF	CIF			4CIF		
	15Hz	7.5 Hz	15Hz	30Hz	15Hz	30Hz	60Hz
City	64; 96; 128	192	224	256; 448; 512	1024	896; 1024	1280; 2048
Crew, Harbour	96; 128; 192	256	320	384; 640; 768	1536	896; 1536, 1280	1780; 3072
Soccer	96; 128; 192	256	320	384; 640; 768	1536	896; 1536, 1280	1780; 3072

Table 2: SNR scalability test points, evaluated were the test points marked green

The system to be evaluated was Vidwav (MPEG wavelet video group). As a reference the current JSVM was tested. A description of the JSVM contribution for this test can be found in [3].
Due to the lack of a PC suitable for continuous playback of 60fps the decoded 60 Hz 4CIF sequences were temporally subsampled to 30 Hz taking the odd frames only with frame 1 being the very first frame of the sequence.

3. Conduction of the test

3.1. Test methodology

The evaluation of coded video in absence of an unimpaired reference, demands for the usage of a particular test method, i.e. the Single Stimulus MultiMedia (SSMM) test method. The Single Stimulus MM test method is basically derived from the Single Stimulus method, as described in ITU-R rec. BT 500-11, and the Single Stimulus with two repetitions, as it was used in the MPEG-4 1995 Competition test. The SSMM has proven to provide valid results in the test for the CfP on Scalable Video Coding [4,5].

The SSMM test method requires the use of progressively scanned display (computer monitors, LCD displays or DLP projectors).
The distance between the screen and the viewer has to be defined according to the dimensions of the images which have to be assessed. For the SVC tests there are three cases:

1. 4CIF=704x576
2. CIF=352x288
3. QCIF=176x144

A DLP projector was used for projection of the sequences, so that the relationship between display height and viewing distance could be preserved. A viewing distance of 3H was used for the 4CIF case, 4H for the CIF case, and 6H for the QCIFcase. In this way the absolute distance to the screen is increased, and the relative impact of viewer movements is decreased.

It is well known that all the test methods are more or less are affected by an effect due to the order of presentation of the material.[1] This effect is particularly strong in the SS category test methods where no reference is present; to reduce this effect SSMM is designed to present twice any condition under test to the subjects. This allows minimizing the contextual effect. The protocol of the SSMM test method is shown in

---

[1] This effect is known as “Contextual effect” and characterises, more or less, quite all the test methods; it is due to the short term memory of the humans that tend to be more relaxed in their judgement when the video material they’ve just seen was of a quality close the one currently under evaluation; in other term when two subsequent conditions have highly different quality the judgement is not as fair as when two subsequent conditions have quite the same quality.

Figure 1; a sequence is presented for 10 seconds and then followed by a message inviting the subjects to vote, that stands for 4 seconds.

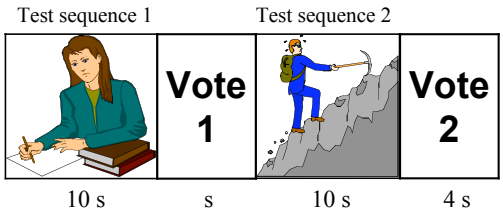


Figure 1 – SSMM test method presentation protocol

The voting scale is based on an 11 grade scale (from 0 to 10); a sample of the scale is given in Figure 2

Figure 2 – SSMM test method voting scale

Subjects were advised to make use of the complete scale for each individual test session.

Tests were performed by 12 people divided into 3 groups with 4 people each. All of the participants were experts in video coding and well familiar with the sequences shown. For both the combined scalability test and SNR test 3 out of this 12 people came from the MPEG Vidwav group. Other test participants were from the multi view video coding group.

A training session of 10 clips preceded every test session, to allow the subjects to adapt to the quality range present in the test.

To allow the subjects to adopt to the illumination during the test 13 seconds of a mid grey image was presented at the beginning of each test. A stabilization phase of 3 sequences preceded every test session, this stabilization phase was repeated once in the middle of the test for ‘renormalization’.

3.2. Test setup

Tests were performed at the meeting room of the test subgroup. The room setup was as close to a laboratory for visual quality assessment as possible. This especially included that there was no room illumination except the reflection from the screen.
A standard screen was used and sequences were displayed using a DLP projector. The center of the display area was 1,50 meters above the floor.
Distance from the projector to the screen was 3,10 meters, this resulted in the size of the sequences and the viewing distances as given in Table 2.

	Image height	**Viewing Distance**
QCIF	23,0 cm	138 cm
CIF	45,5 cm	182 cm
4CIF	91,0 cm	273 cm

Table 2: Image height and viewing distance

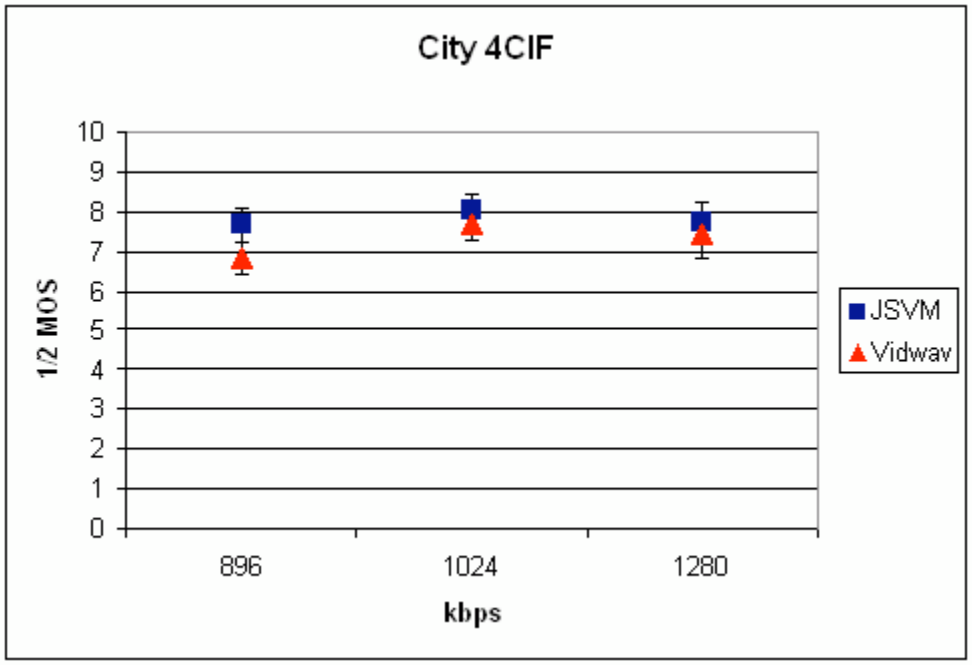
4. Data processing

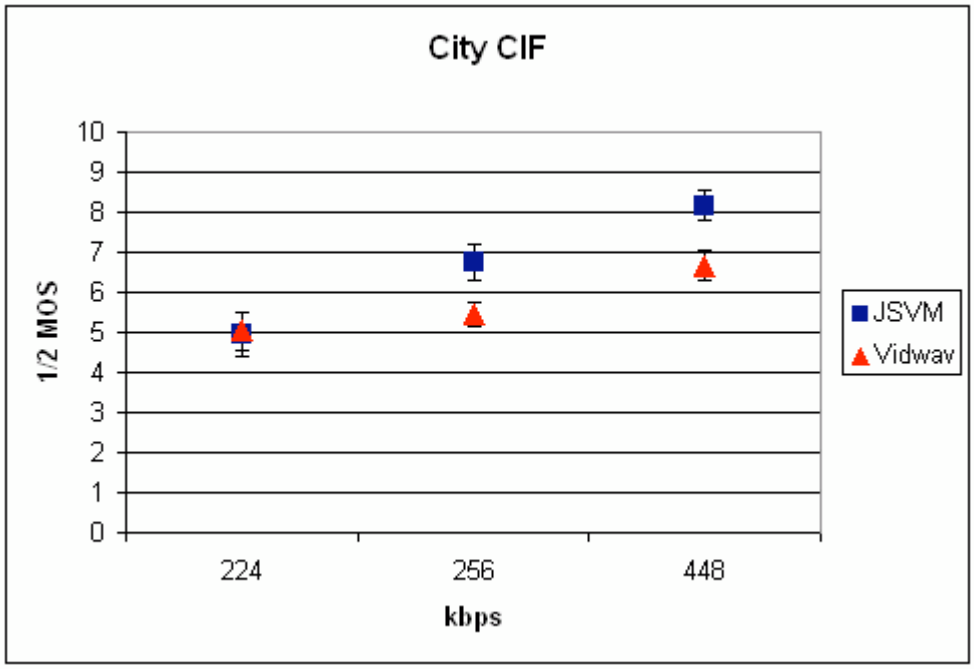
The votes were collected by means of paper scoring sheet. Each condition under test received two votes and the final score was obtained making the mean of the two values. After removal of outliners (below 1 % of all votes) variance was computed on this mean values as well as the 95% CI values.
Due to the wide variety of the votes for the SNR test the highest and the lowest mean vote was removed.

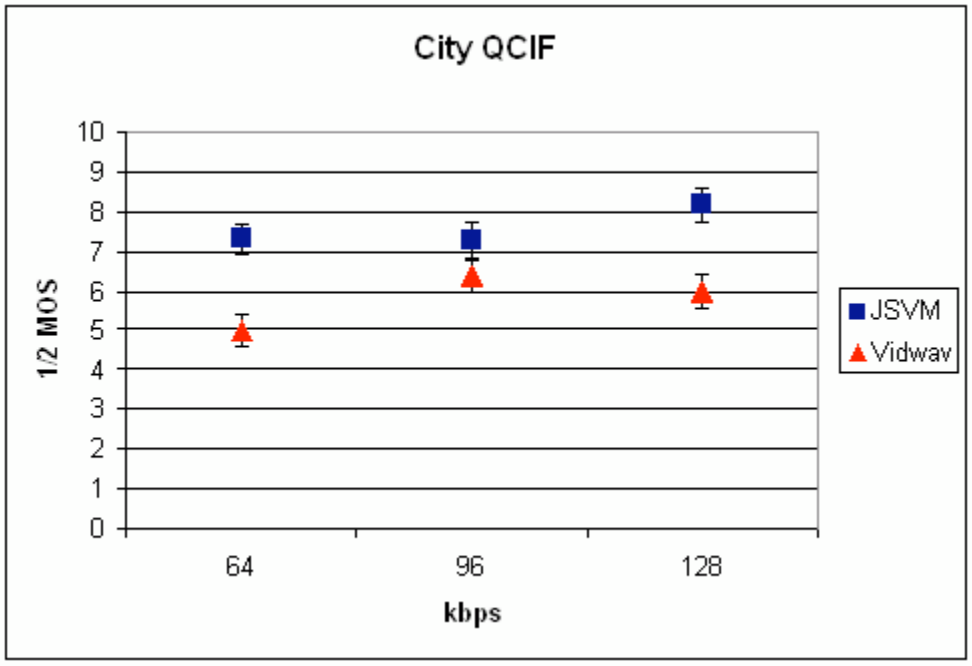
5. Results

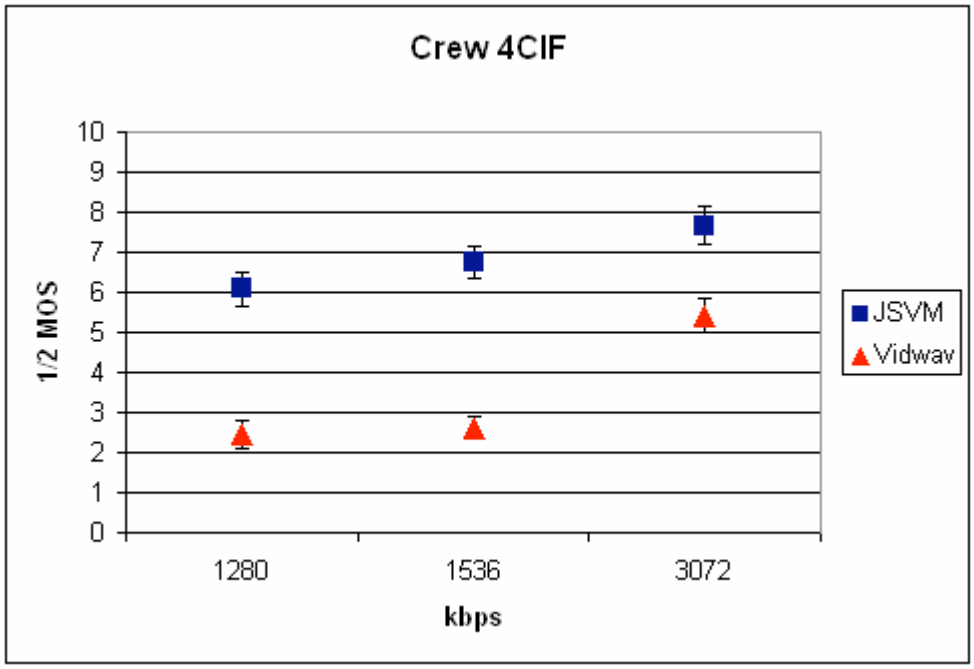
All results are based on a 11 grade scale, where 0 stands for the worst and 10 stands for the best quality. All data points show a 95% CI interval based on the "1 sigma" method. The mean value is located in the middle of each CI-bar.
It has to be noted, that in a Single Stimulus environment it is not possible to mix up results from two different test sessions into one graph or to compare results from to different test sessions. Results are only valid and comparable within one single test session.

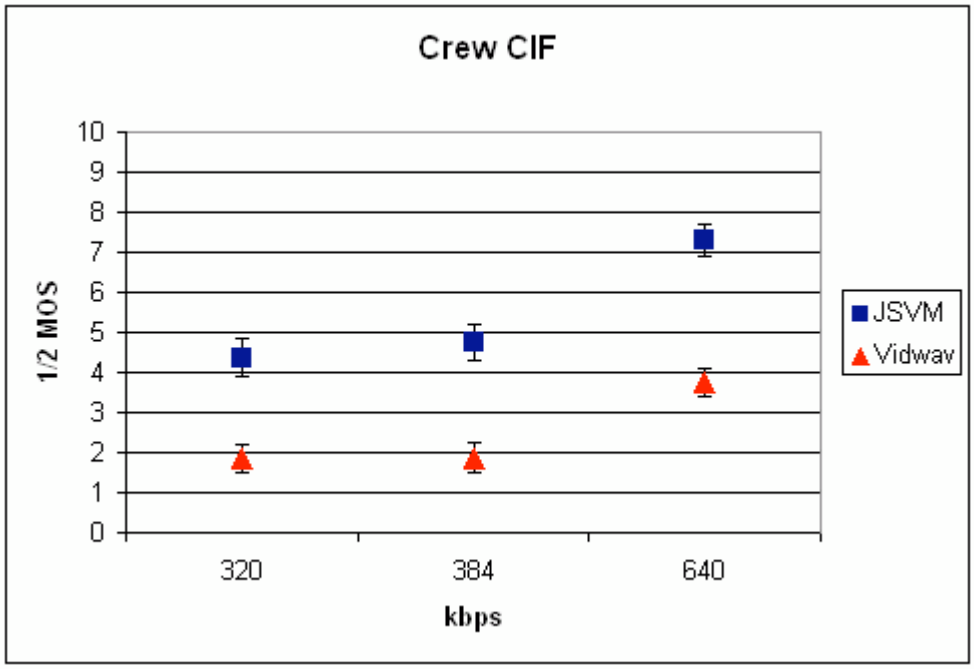
5.1. Results combined scalability:

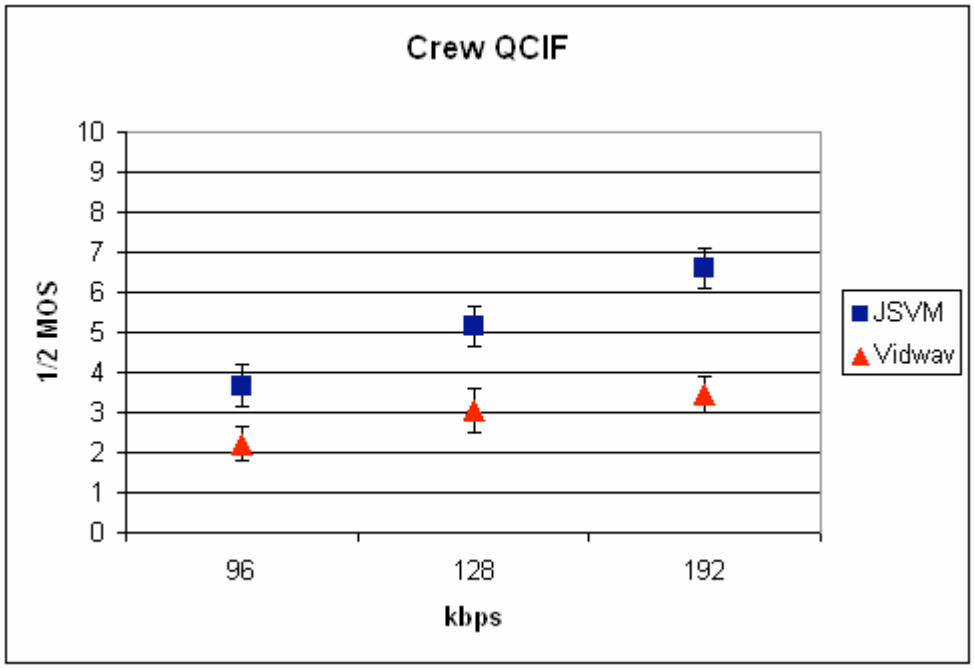


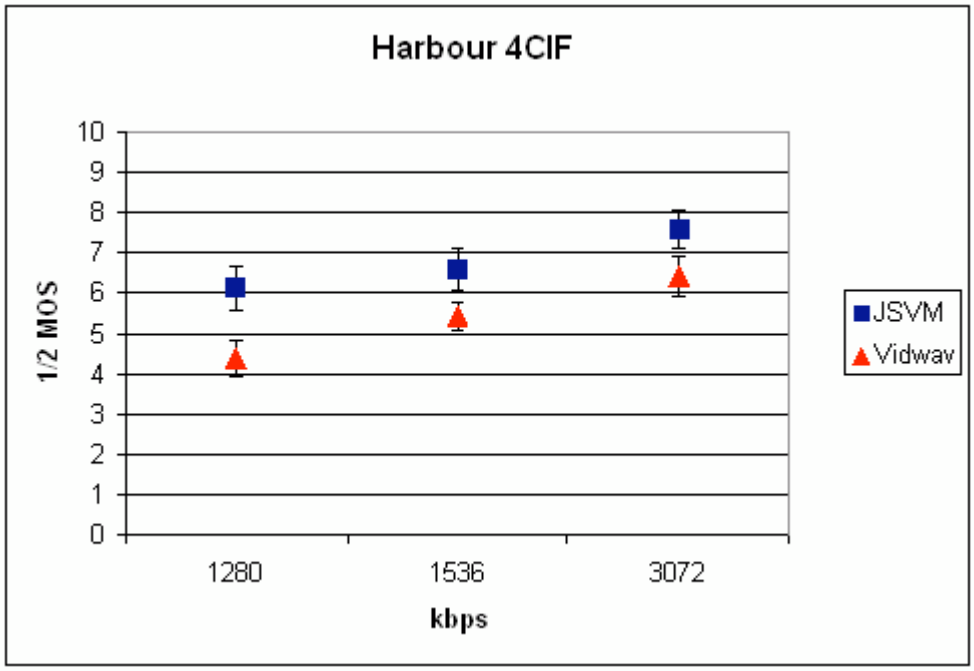


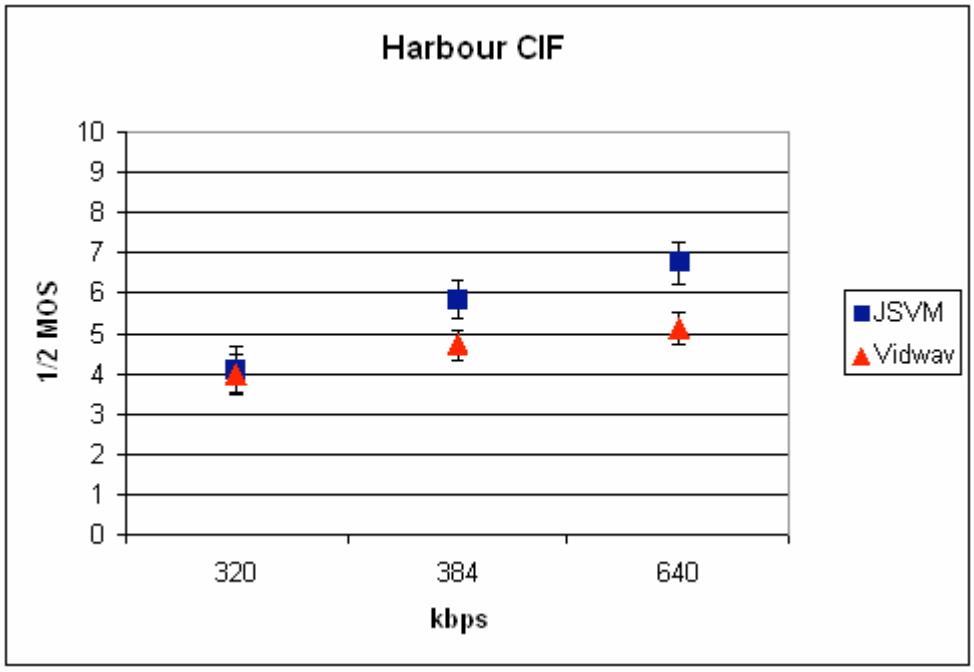

City QCIF
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
64
96
128
kbps
JSVM
Vidwav

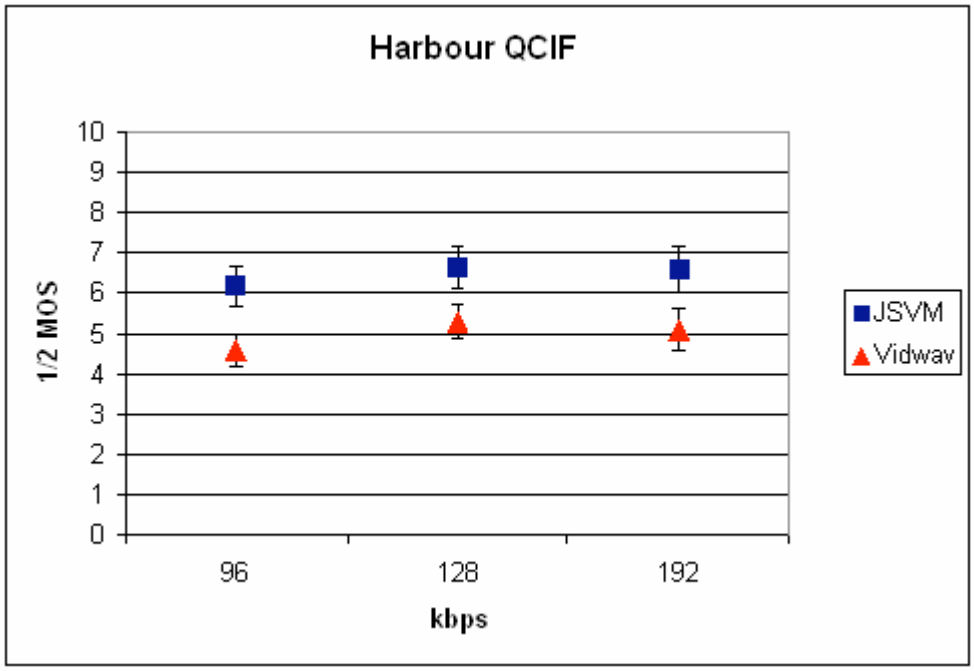

Crew 4CIF
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
1280
1536
3072
kbps
JSVM
Vidwav

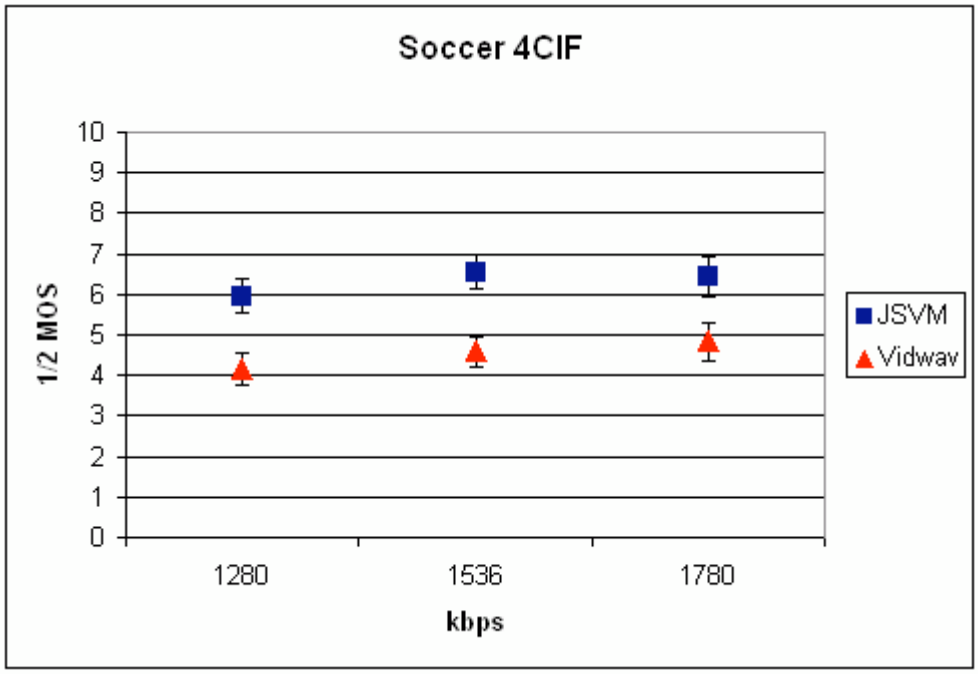

Crew CIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
320
384
640
kbps
JSVM
Vidwav

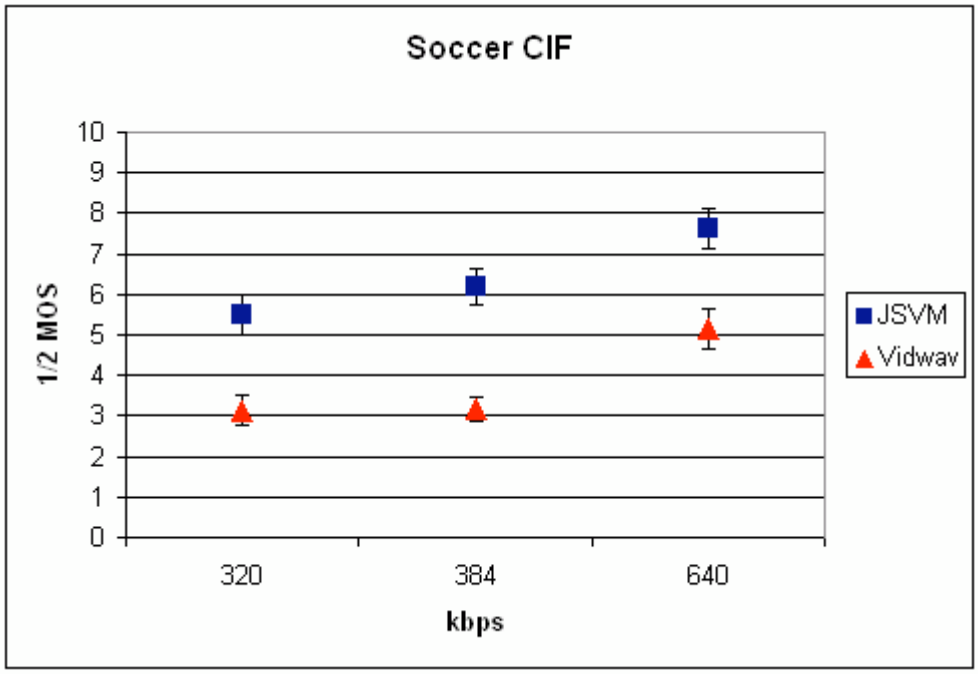

Crew QCIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
96
128
192
kbps
JSVM
Vidwav

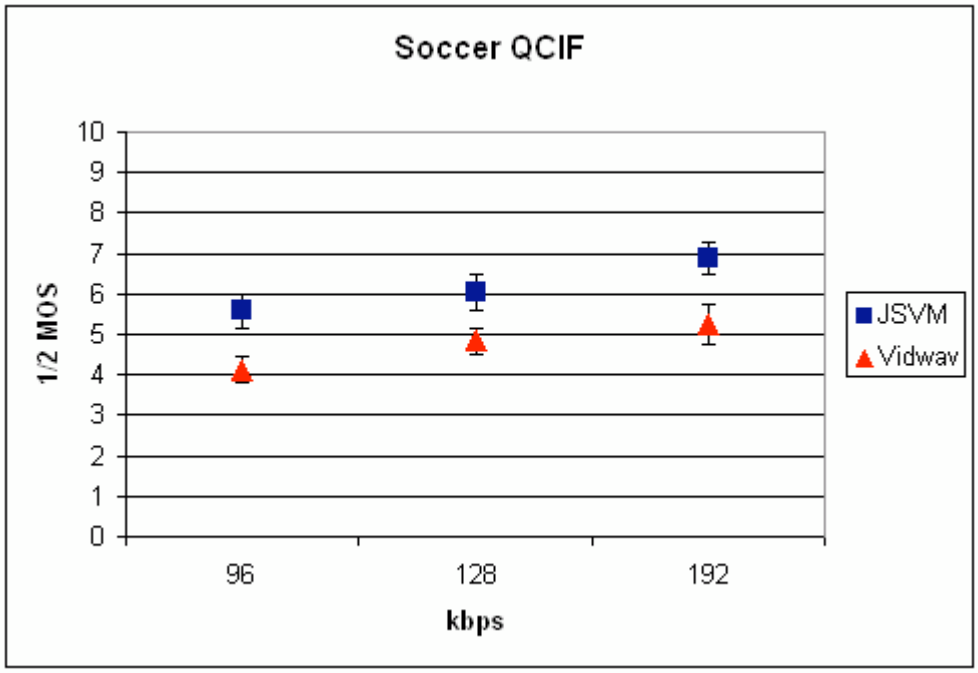

Harbour 4CIF
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
1280
1536
3072
kbps
JSVM
Vidwav


Harbour CIF
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
320
384
640
kbps
JSVM
Vidwav

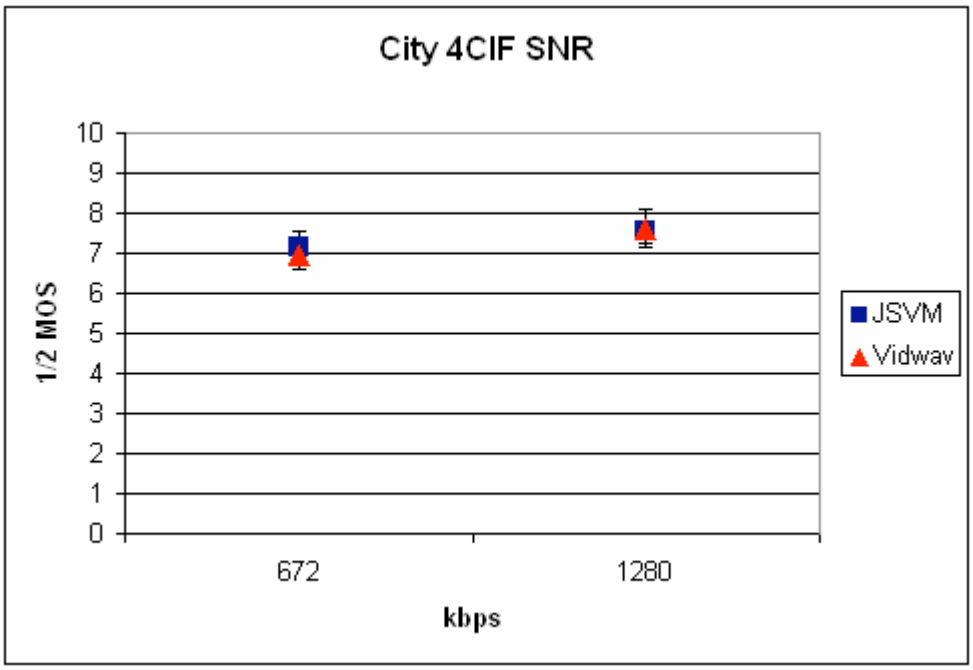

Harbour QCIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
96
128
192
kbps
JSVM
Vidwav

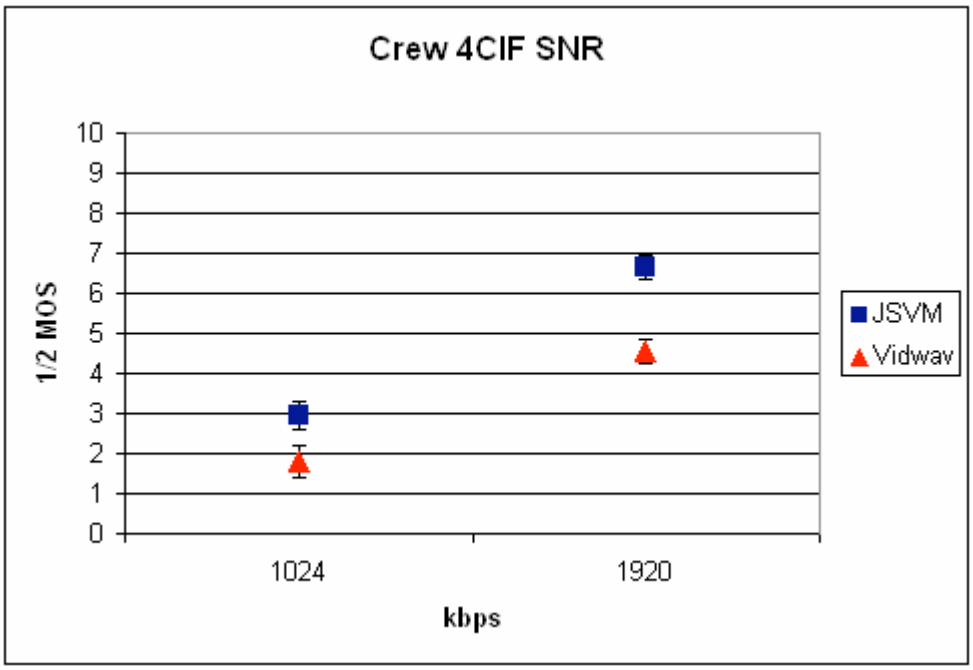

Soccer 4CIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
1280
1536
1780
kbps
JSVM
Vidwav

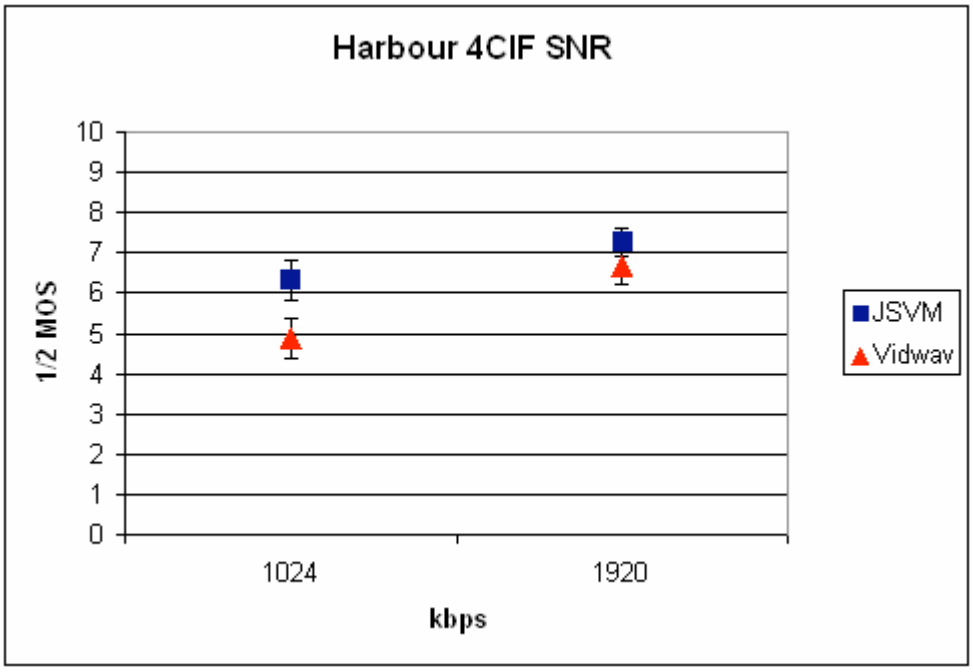

Soccer CIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
320
384
640
kbps
JSVM
Vidwav

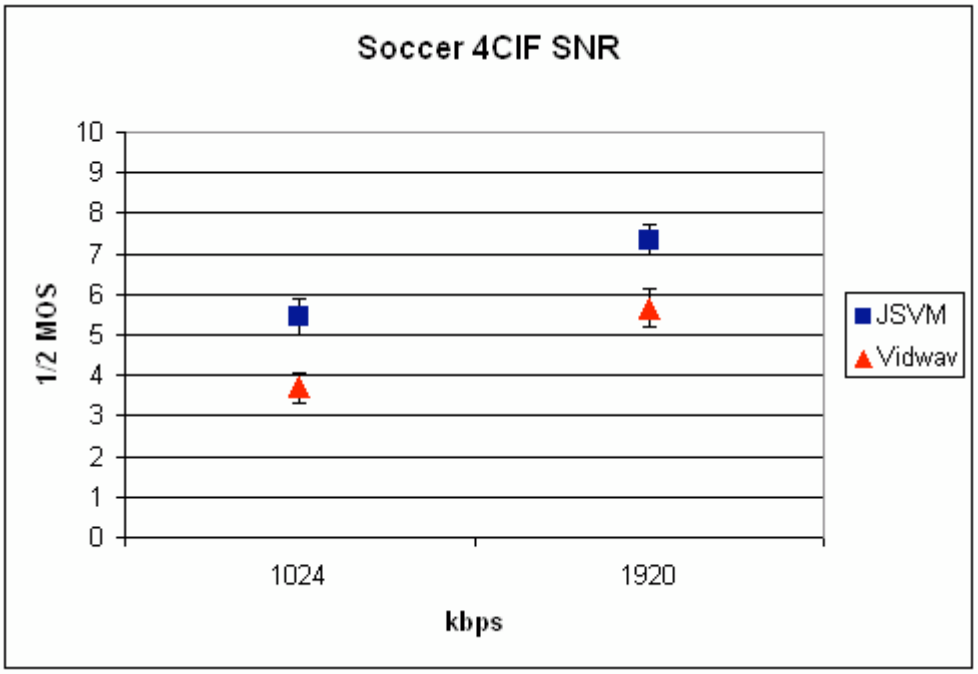

Soccer QCIF
10
9
8
7
6
5
4
3
2
1
0
1/2 MOS
96
128
192
kbps
JSVM
Vidwav

5.2. Results SNR scalability:






Harbour 4CIF SNR
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
1024
1920
kbps
JSVM
Vidwav


Soccer 4CIF SNR
1/2 MOS
10
9
8
7
6
5
4
3
2
1
0
1024
1920
kbps
JSVM
Vidwav

6. References

[1] ISO/IEC JTC1/SC29/WG11, "Description of Testing in Wavelet Video Coding," ISO/IEC JTC1/SC29/WG11, N7823, 75th MPEG Meeting, Bangkok, January, 2004.

[2] Report of the Subjective Quality Evaluation for SVC CE1, ISO/IEC JTC1/ SC29/ WG11, N6736, 70th MPEG Meeting, Palma de Mallorca, Spain, Oct. 2004.

[3] Mathias Wien, "JSVM bitstreams for VIDWAV visual evaluation," ISO/IEC JTC1/SC29/WG11, M13264, , 76th MPEG Meeting, Montreux, Switzerland, April 2006.

[4] ISO/IEC JTC1/SC29/WG11, Call For Proposals on Scalable Video Coding Technology N6193, Waikoloa, December 2003.

[5] ISO/IEC JTC1/SC29/WG11, Subjective test results for the CfP on Scalable Video Coding Technology M10737, Munich, March 2004.

[2] Scoring sheets will be rigorously anonymous, indicating just the seat position occupied by each viewer, the date and the time.