ITUPublications

International Telecommunication Union
Standardization Sector

ITU-T Technical Report

(01/2023)

JSTR-OPTR

Optimizing bit rates and transmission resolution by considering display characteristics and available bandwidth

Technical Report ITU-T JSTR-OPTR

Optimizing bit rates and transmission resolution by considering display characteristics and available bandwidth

Summary

This Technical Report provides subjective test results that can be used for optimal video transmission methods in terms of bit rates and resolution, which use the minimum bandwidth while providing equivalent perceptual video quality by considering content characteristics and display size/resolution.

Keywords

Bit rates, content characteristics, display size/resolution, equivalent perceptual video quality, minimum bandwidth, optimal video transmission.

Note

This is an informative ITU-T publication. Mandatory provisions, such as those found in ITU-T Recommendations, are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change Log

This document contains Version 1 of the ITU-T Technical Report JSTR-OPTR on "*Optimizing bit rates and transmission resolution by considering display characteristics and available bandwidth*" approved at the ITU-T Study Group 12 meeting held in Geneva, 18–26 January 2023.

Contacts:

Chulhee Lee Yonsei University Korea (Rep. Of)	Tel.: +82 2 2123 2779 E-mail: chulhee@yonsei.ac.kr
Vincent BARRIAC Orange France	Tel.: +33 296 071 810 E-mail: vincent.barriac@orange.com

Table of Contents

Page

Technical Report JSTR-OPTR

Optimizing bit rates and transmission resolution by considering display characteristics and available bandwidth

1 Scope

This Technical Report provides subjective test results that can be used for optimizing video transmission over networks in terms of bit rates and resolution, which use the minimum bandwidth while providing equivalent perceptual video quality by considering content characteristics and display size/resolution.

This Report can be used for:

1) Determining optimal bit rates and transmission resolution (resizing) by considering content characteristics and display characteristics (resolution and size) when there is no bandwidth limitation (sufficiently large bandwidth is available).

2) Determining optimal transmission resolution (resizing) by considering content characteristics and display characteristics (resolution and size) when there is bandwidth limitation (maximum available bandwidth is given).

3) Designing balanced subject tests by considering content characteristics and display characteristics (resolution and size).

2 References

[ITU-T H.264] Recommendation ITU-T H.264 (2019), *Advanced video coding for generic audiovisual services*.

[ITU-T H.265] Recommendation ITU-T H.265 (2021), *High efficiency video coding*.

[ITU-T J.342] Recommendation ITU-T J.342 (2011), *Objective multimedia video quality measurement of HDTV for digital cable television in the presence of a reduced reference signal.*

[ITU-T P.800.1] Recommendation ITU-T P.800.1 (2016), *Mean opinion score (MOS) terminology*.

[ITU-T P.910] Recommendation ITU-T P.910 (2022), *Subjective video quality assessment methods for multimedia applications*.

[ITU-T P.911] Recommendation ITU-T P.911 (1998), *Subjective audiovisual quality assessment methods for multimedia applications*.

[IDMS1] International Committee for Display Metrology (ICDM) (2012), *Information Display Measurement Standard*.

[IR.94] GSM Association official document IR.94 (2019), *IMS profile for conversational video service,* Version 14.0.

[SSIM] Wang, Z., Bovik, A.C., Sheikh, H.R. Simoncelli, E.P. (2004), *Image quality assessment: from error visibility to structural similarity*, IEEE transactions on image processing, vol. 13, No. 4, April 2004.

[VMAF] GitHub – Netflix/vmaf (2022), *Perceptual video quality assessment based on multi-method fusion*.

3 Definitions

3.1 Terms defined elsewhere

This Technical Report uses the following term defined elsewhere:

3.1.1 mean opinion score (MOS): See [ITU-T P.910].

3.2 Terms defined in this Technical Report

None.

4 Abbreviations and acronyms

This Technical Report use the following abbreviations and acronyms:

ACR	Absolute Category Rating
HD	High Definition
Kbit/s	kilobits per second
Mbit/s	megabits per second
MOS	Mean Opinion Score
PSNR	Peak Signal-to-Noise Ratio
QVGA	Quarter Video Graphics Array
SI	Spatial Information
SRC	Source Reference Channel or Circuit
SSIM	Structural Similarity Index Measure
TI	Temporal Information
VGA	Video Graphics Array
ViLTE	Video over LTE
VMAF	Video Multi-Method Assessment Fusion
UHD	Ultra High Definition

5 Perceptual video quality as functions of display size/resolution and transmission resolution

It was reported that, depending on content characteristics and display size/resolution, the perceptual video quality shows a large variation even at the same bit rates. To effectively utilize the available bandwidth, it is desirable to optimize the bit rate by taking into account the content characteristics and display size/resolution.

To analyse the relationship between transmission resolution/bit rates and display size/resolution, subjective tests were performed. Ten source video sequences were selected. Table 1 shows the source characteristics and Table 2 shows SI/TI information. Figure 1 shows the test booth.

Table 1 – Source characteristics

	Class	High frequency	Colour	Temporal movement
SRC01	Document	Medium	Low	Medium
SRC02	Sports	Medium	Medium	High
SRC03	Drama	High	Medium	Medium
SRC04	Document	High	Low	Medium
SRC05	Animation	Medium	Low	High
SRC06	Document	High	High	Medium
SRC07	Sports	Low	Medium	High
SRC08	Movie	Medium	High	Medium
SRC09	Advertisement	Medium	High	High
SRC10	Drama	Medium	High	Medium

Table 2 – SI/TI information of the source sequences

KT01	SI	TI	SI_avg	TI_avg
SRC01	121	21	111	20
SRC02	99	39	69	10
SRC03	93	35	63	16
SRC04	127	50	61	10
SRC05	59	56	25	10
SRC06	66	26	58	24
SRC07	58	22	38	13
SRC08	75	48	43	17
SRC09	66	53	45	11
SRC10	84	47	47	12

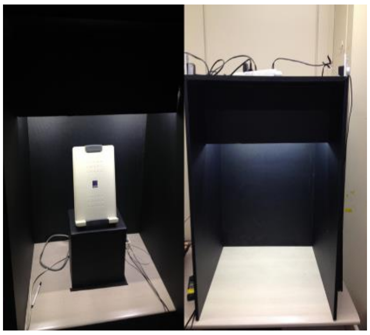


Figure 1 – Test booth

Figure 2 shows the scatter plots of the four displays (4-inch smartphone, 5-inch smartphone, tablet, 40-inch TV). At low bit rates, lower transmission resolutions produced better perceptual quality in all the cases. For small displays, 720p produced essentially the same perceptual quality as 1080p. It appears that only the 40-inch TV monitors showed noticeably improved quality for the 1080p signals compared with the 720p signals. Depending on source characteristics, some variations were observed.

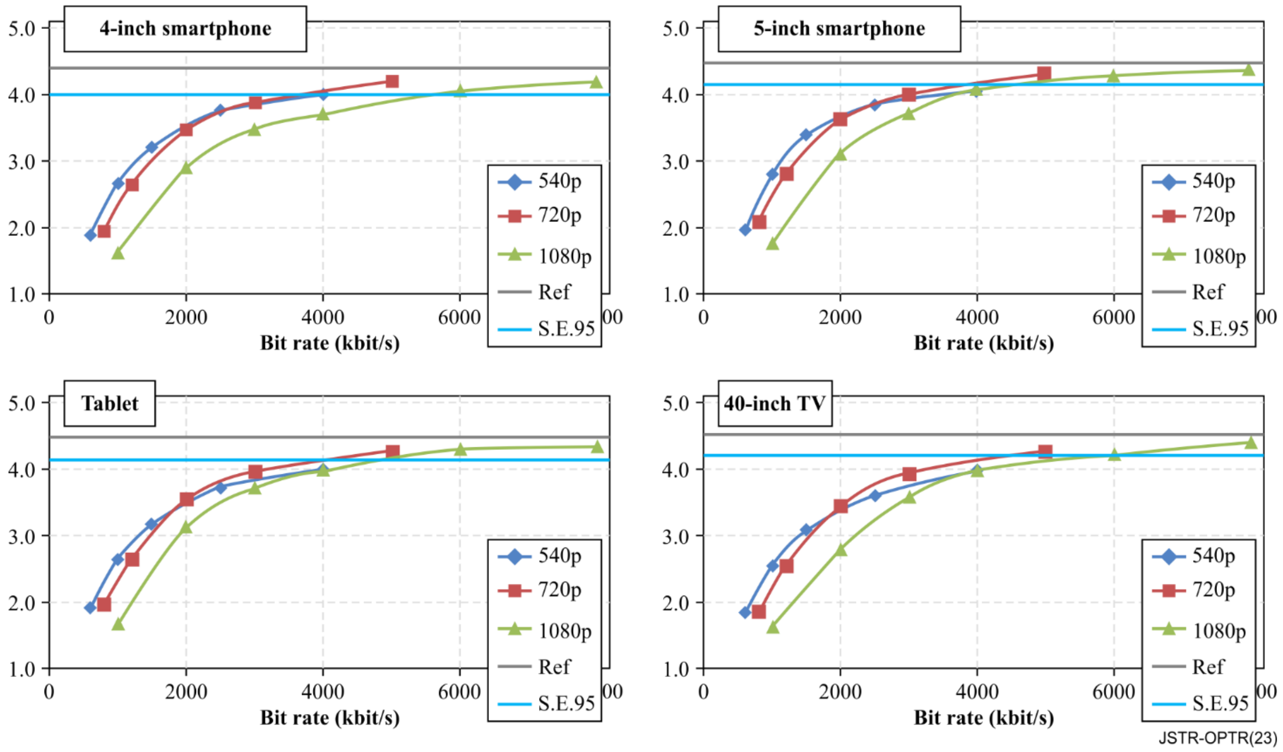


Figure 2 – Subjective scores at various bit rates and transmission resolutions

6 Minimum bit rates for statistically equivalent perceptual quality

A number of subjective tests were performed using various displays (Table 3). The absolute category rating (ACR) assessment method was used in the subjective test and at least 24 subjects participated in each session after screening. A variety of source video sequences (full HD, 1080p) were chosen by considering content characteristics. The source video sequences were reduced to lower resolutions (720p and 540p) and the original and resized video sequences encoded using [ITU-T H.264]. Table 4 shows the bit rate information. The subjective tests were conducted in accordance with [ITU-T P.910]. For smartphones and tablet, a viewing booth was built to control illumination conditions. The booth brightness was set to 500 lux at the desk level (Figure 1).

Table 3 – Displays used in the subjective tests

Description	Size	Resolution
Smartphone	4-inch	1 136 × 640
Smartphone	5-inch	1 920 × 1 080
Tablet	9.7-inch	2 048 × 1 536
TV monitor	42-inch	1 920 × 1 080

Table 4 – Bit rate information

Resolution	Bit rates (Mbit/s)
1080p	1, 2, 3, 4, 6, 8
720p	0.8, 1.2, 2, 3, 5
540p	0.6, 1, 1.5, 2.5, 4

Figure 3 shows the perceptual video quality (MOS) at the various bit rates. It appears that the perceptual video quality of the reduced resolutions was better than that of the original resolution (1080p) at low bit rates. Figure 3 also shows the statistically equivalent perceptual quality levels to the perceptual quality of 1080p at 8 Mbit/s. Tables 5-8 show the minimum bit rates for the four displays, which can provide the perceptual video quality that is statistically equivalent to that of 1080p at 8 Mbit/s.

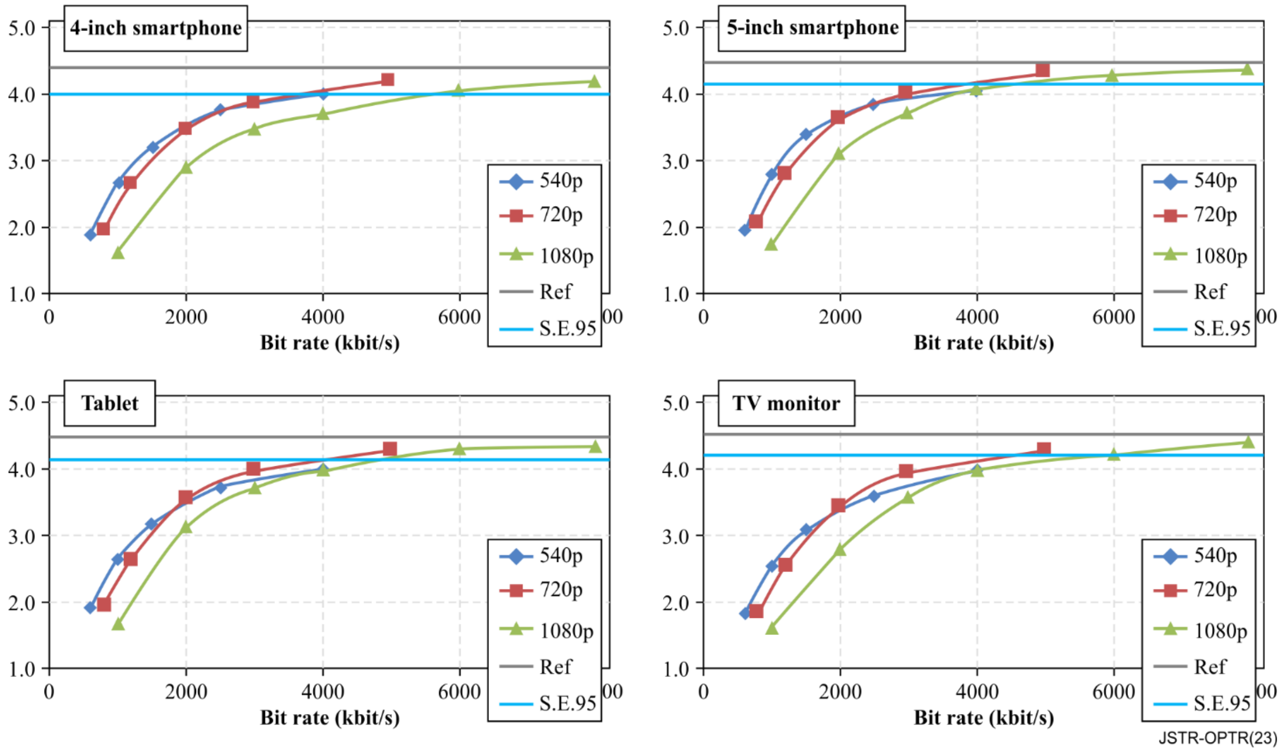


Figure 3 – MOS as a function of bit rates of the four displays

Table 5 – Minimum bit rates (Mbit/s) for statistically equivalent perceptual quality (4-inch smartphone)

	540p	720p	1080p	Min
SRC01	4	5	6	4
SRC02	4	5	6	4
SRC03	2.5	3	6	2.5
SRC04	4	5	6	4
SRC05	1.5	2	2	1.5
SRC06	–	5	8	5
SRC07	4	2	3	2
SRC08	4	–	6	4
SRC09	–	3	8	3
SRC10	2.5	2	3	2
Average				3.2

Table 6 – Minimum bit rates (Mbit/s) for statistically equivalent perceptual quality (5-inch smartphone)

	540p	720p	1080p	Min
SRC01	–	5	6	5
SRC02	–	5	8	5
SRC03	–	5	4	4
SRC04	2.5	5	8	2.5
SRC05	4	3	3	3
SRC06	4	2	3	2
SRC07	–	5	6	5
SRC08	2.5	1.2	2	1.2
SRC09	–	5	6	5
SRC10	–	5	6	5
Average				3.5

Table 7 – Minimum bit rates (Mbit/s) for statistically equivalent perceptual quality (tablet)

	540p	720p	1080p	Min
SRC01	–	–	6	6
SRC02	–	–	6	6
SRC03	4	3	6	3
SRC04	4	3	3	3
SRC05	0.6	1.2	2	0.6
SRC06	4	5	6	4
SRC07	2.5	3	2	2
SRC08	–	–	6	6
SRC09	4	5	4	4
SRC10	4	5	3	3
Average				3.8

Table 8 – Minimum bit rates (Mbit/s) for statistically equivalent perceptual quality (TV monitor)

	540p	720p	1080p	Min
SRC01	4	3	4	3
SRC02	–	–	6	6
SRC03	4	3	4	3
SRC04	–	5	6	5
SRC05	–	2	3	2
SRC06	–	5	8	5
SRC07	–	5	4	4
SRC08	4	5	6	4
SRC09	–	–	8	8
SRC10	–	5	4	4
Average				4.4

7 Optimizing bit rates and transmission resolution for mobile displays based on perceptual video quality

Subjective tests were performed using two mobile displays (Table 9). The ACR assessment method was used and at least 24 subjects participated in each session after screening. A variety of source video sequences (full HD, 1080p) were chosen by considering content characteristics. The source video sequences were reduced to a lower resolution (720p) and the original and resized video sequences were encoded using [ITU-T H.264]. Table 10 shows the bit rate information. The content consisted of five categories: sports, films, animation, documentaries and drama.

The subjective tests were conducted in accordance with [ITU-T P.910]. For smartphones and tablets, a viewing booth was built to control illumination conditions. The booth brightness was set to 500 lux at the desk level (Figure 1).

Table 9 – Displays used in the subjective tests

Description	Size	Resolution
Smartphone	5-inch	1 920 × 1 080
Tablet	9.7-inch	2 048 × 1 536

Table 10 – Bit rate information

Resolution	Bit rates (Mbit/s)
1080p	6, 8
720p	0.5, 0.7, 1, 2, 3

Tables 11 and 12 show the perceptual video quality at various bit rates in terms of the MOS values for the smartphone and tablet for each of the five categories. It can be seen that, depending on the content and display type, low bit rates at lower resolutions can provide equivalent perceptual video quality levels to 1080p at 8 Mbit/s.

Table 11 – Perceptual video quality at various bit rates (smartphone)

	Clip	Bit rates (Mbit/s)						
		0.5	0.7	1	2	3	6	8
Sports	1	1.1	1.2	2.2	3.5	3.9	4.3	4.2
	2	1.1	1.5	2.1	3.6	4.3	4.1	4.3
	4		1.4	2.8				4.2
	5		1.9	2.7	3.6	3.7		4
	6		2.2	2.7	4.1	3.9		4.3
	8		2.8	2.3				4.9
	9		1.6	2.3	3.7	4.3	3.8	4.3
Movies	4		3.3	4				4.1
	5		3.4	3.3	4.3	4.2	4.3	4.2
	6		3.3	4.4				4.2
	7		3.6	3.4	4.2	4.3		4.3

Table 11 – Perceptual video quality at various bit rates (smartphone)

	Clip	Bit rates (Mbit/s)						
		0.5	0.7	1	2	3	6	8
Animations	1	1.2	1.7	2.7	3.6	4.3		4.1
	2	2.7	3.4	3.6				4.5
	3	2.4	2.9	3.5				4.2
	4		2.8	3.5				4.2
	5		1.5	3.5	4.3	4.2		4.6
	6		3.5	3.6				4.5
	7		1.1	3.6	4.3	4.3	4.4	4.4
	8		1.6	2.6	3.4	4.1		4.6
	9		1.5	2.2	3.6	4.1	3.8	4.2
	10		1.8	2.3	3.6	4.1		4.1
	11			3.5	4	3.9		4.3
Documentary	1	2.4	2.8	3.7				4.3
	4		3.7	4				4.1
	5		3.2	3.3	3.5	4.2	4.1	4.1
	6		1.5	2.8	3.7	4.1	4.2	4.2
	7		1.7	2.9	3.7	3.6	3.6	4.2
	8		1.2	3.4	3.7	3.9		4.1
	9		2.8	2.8	3.4	3.8		4
Drama	1	2.1	2.4	2.9				4.5
	2	2.2	3.1	3.4	4.3	4.6	4.8	4.6
	3	2.3	2.8	4.3				4.6
	4		2.8	4.2	4.4	4.4	4.6	4.4
	5		2.2	4.3	4.8	4.6		4.5
	6		3.2	3.8				4.8
	7		2.5	3.8				4.5
	8		2.4	3.2	3.7	3.9	4.4	4.2
	9		3.2	3.7	3.8	3.8	4.1	4.2
	10		1.4	3.9	4.3	4.3		4.2

Table 12 – Perceptual video quality at various bit rates (Tablet)

	Clip	Bit rates (Mbit/s)						
		0.5	0.7	1	2	3	6	8
	1	1	1.3	2.9	4	4	4	4.5
	2	1.4	1.9	2.4	3.7	4	4	4.4
	3	1.9	2.5	2.2	3.5	3.8	3.7	4

Table 12 – Perceptual video quality at various bit rates (Tablet)

	Clip	Bit rates (Mbit/s)						
		0.5	**0.7**	**1**	**2**	**3**	**6**	**8**
Sports	4		1.3	3.1				4.7
	5		2.1	2.9	3.6	4		4.2
	6		2.6	3.2	4.2	4.3		4.3
	8		3	3.7				4.7
	9		1.5	2.4	3.8	4.3	4.1	4.3
Movies	1	1.4	1.5	2.5	3.7	3.6	3.7	4
	3	2.4	3.4	4.1				4.3
	4		3.3	3.9				4.2
	5		3.6	3.7	4.4	4.3	4.3	4.1
	6		3.5	4.5				4.6
	7		3.8	3.7	4.3	4.3		4.6
	10		1.4	3.6	3.7	3.7		4.1
Animations	1	1.1	1.6	3.1	3.9	4.4		4.5
	2	2.9	3.4	4.1				4.5
	3	2.7	3.3	3.7				4.3
	4		2.7	3.9				4.3
	5		1.8	3.8	4..2	4.3		4.5
	6		3.7	3.4				4.7
	7		1.1	4	4.4	4.3	4.4	4.5
	8		1.6	2.8	4	4		4.5
	9		1.7	2.1	3.7	3.6	4	4.1
	10		2	3	3.9	4.3		4.1
	11			3.5	4.3	3.8		4.3
Documentary	1	2.4	3.4	4				4.7
	4		4	4				4.3
	5		3.7	3.3	4	4.2	4.3	4.3
	6		1.5	2.6	4	3.8	4	4.2
	7		1.9	3	3.9	3.7	3.8	4.1
	8		1.5	3.4	3.6	4.1		4
	9		3	2.8	3.4	4.1		4
Drama	1	2.1	2.2	3				4.6
	2	2.1	3	3.6	4.2	4.5	4.7	4.6
	3	2.6	3	4.1				4.7
	4		3.1	4.3	4.2	4.4	4.5	4.3
	5		2.1	4.3	4.5	4.6		4.8
	6		3.2	4.1				4.7
	7		2.4	4.1				4.6

Table 12 – Perceptual video quality at various bit rates (Tablet)

	Clip	Bit rates (Mbit/s)						
		0.5	0.7	1	2	3	6	8
	8		2	3.4	3.7	3.9	4	4
	9		3.3	3.9	4.1	3.8	4	4.1
	10		1.6	4	4.4	4.3		4.5

8 Influence of video resolution and frame rate on perceived quality for videotelephony applications

8.1 Joint impact of resolution and frame rate

In 2014 and 2015, Orange conducted several ITU-T P.911 subjective tests on perceived audiovisual quality of content to be used in Video over LTE (ViLTE) contexts. At that time, the corresponding standard [IR.94] was not specifying high levels of quality for the video part of the service, in particular in terms of image resolution and frame rate, and some decisions had to be taken in order to properly design the service and its support in networks and end devices.

Several devices (smartphones and tablets) have been tested with a large variety of:

– display resolutions (from 176 × 144 to 1280 × 720 pixels);

– frame rates (from 10 to 25 frames/second).

The value ranges correspond to the state of the art at that time; this may have evolved since.

The results of these tests showed clearly that both factors play an important role in the perception of quality. This is illustrated by the example of Figure 4.

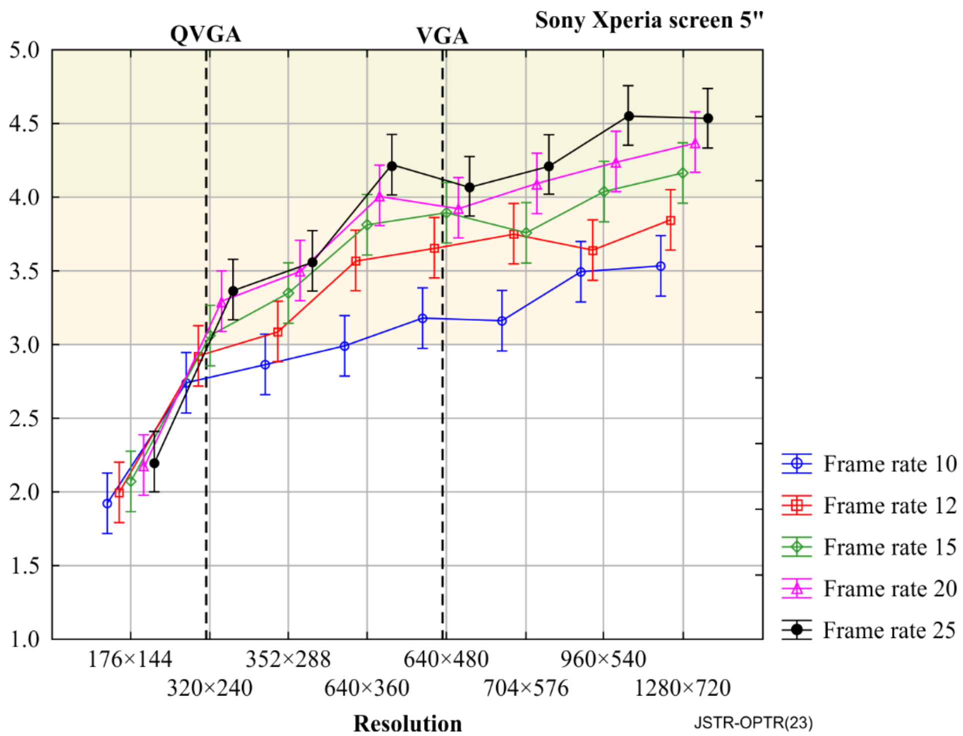


Figure 4 – Example of test result showing the joint impact of video resolution and frame rate

8.1.1 Test 1: ITU-T H.264 only, operational lessons learned

The main conclusions of these tests conducted on smartphones are the following:

– QVGA (320 × 240) at 15 fps can only provide a medium quality user experience (MOS around 3).

– A good quality (MOS above 4) requires a minimum resolution of 640 × 360 at 15 fps.

– VGA (640 × 480) is the widely supported resolution achieving this quality level.

– There is no real impact of screen sizes (3.5", 6" tested).

– Resolutions above VGA (up to 720p) are not needed for smartphones with 5"/6" screens.

– Influence of video bit rate needs also to be considered (with ITU-T H.264 "baseline", see Figure 5).

  – 384 kbit/s is the minimum bit rate to ensure a reasonably good quality experience (around 3.5 MOS). Below 384 kbit/s both the quality and acceptability drop;

  – A very good video quality (above 4 MOS) requires a bit rate of at least 768 kbit/s.

The ViLTE operating range for good quality with the ITU-T H.264 baseline therefore requests at least the application of ITU-T H.264 with Level 3.1. This is what is now reflected in [IR.94].

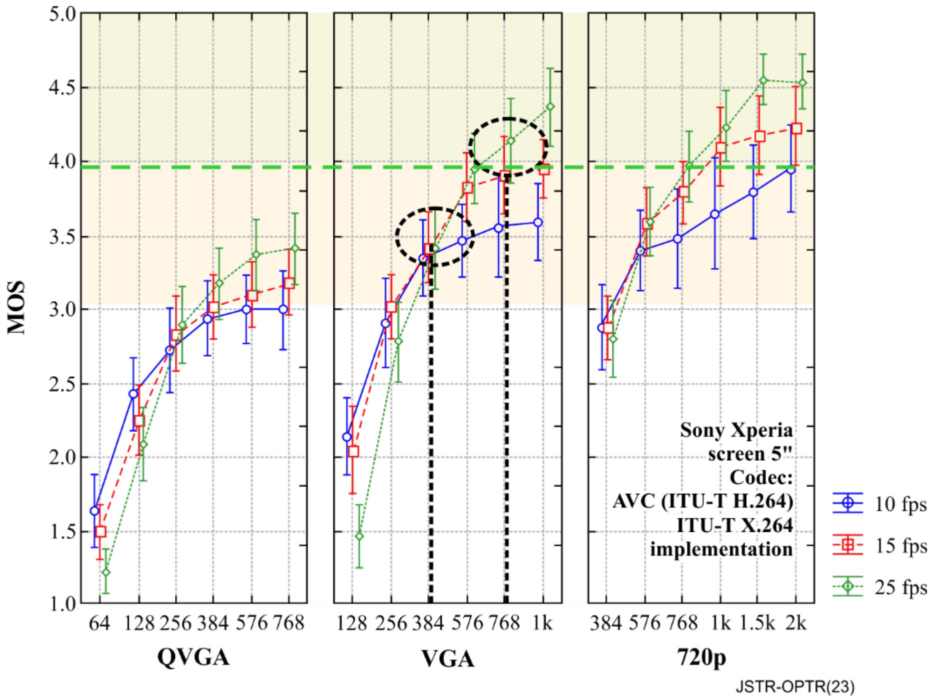


Figure 5 – Example of test result showing the impact video coding bit rate on perceived quality

8.1.2 Test 2: ITU-T H.264 vs ITU-T H.265, operational lessons learned

In opposition to the *a priori* consideration on the interest of [ITU-T H.265] compared with [ITU-T H.264], it has been found during this test (on HM reference software) that this new generation of codec is not only meant for high video qualities (4K, UHD…), but rather that its benefit is at low bit rates: [64-128 kbit/s]. At 64 kbit/s acceptability is still > 80% and MOS > 3 (see Figure 6), which makes low bit rates possible as fallback solutions. At such a low bit rate, it is impossible for [ITU-T H.264] to encode VGA properly.

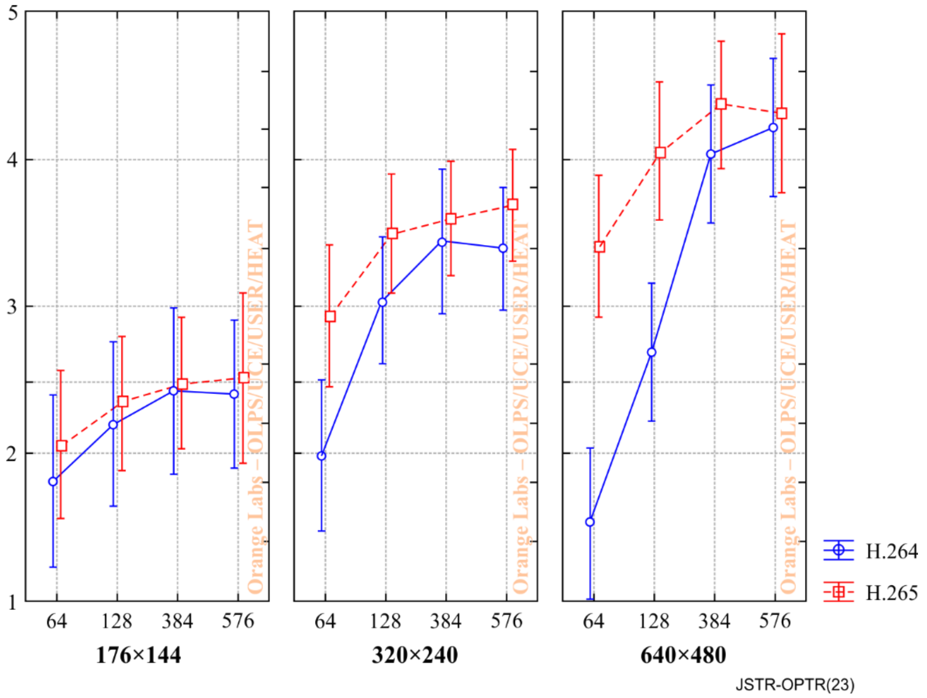


Figure 6 – Example of test results showing the benefit of using ITU-T H.265 at low bit rates

Furthermore, such a comparison has been found to be highly dependent on the implementation of the video decoder. Figure 7 illustrates this very well, where for a video coding bit rate at 128 kbits/s, the difference between [ITU-T H.265] and [ITU-T H.264] varies from 0.5 to 1.2 MOS depending on the implementation of ITU-T H.264 decoding.

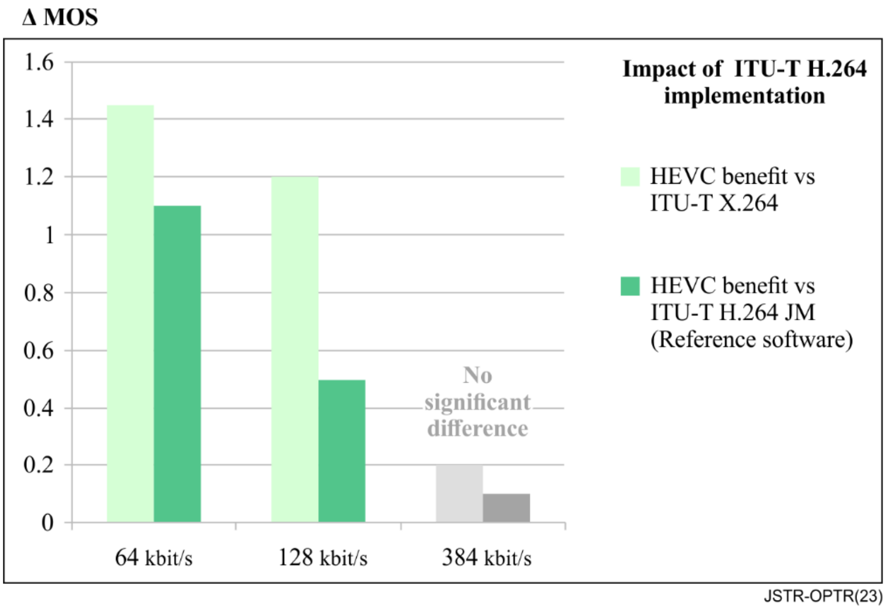


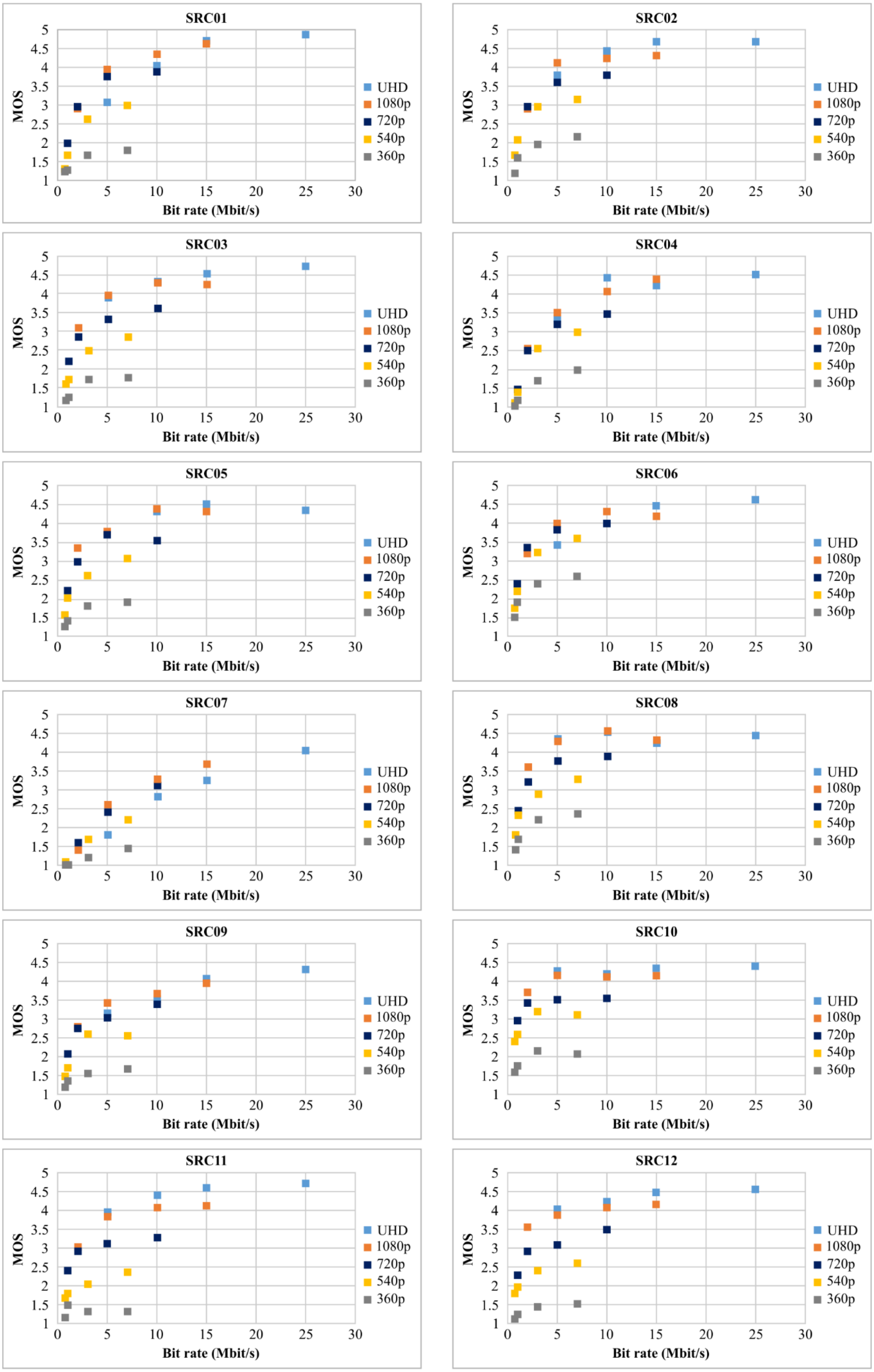
Figure 7 – Example of test result showing the impact of ITU-T H.264 implementation on perceived quality

The various lessons learned that are reported above are practical illustrations of the various factors that can impact video quality experience in videotelephony contexts; they may be of interest for optimizing bit rates and transmission resolution by considering display characteristics and available bandwidth.

9 Perceptual video quality of UHD signals at various bit rates and transmission resolutions

In the subjective test, 12 high quality UHD resolution video clips were used. The video clips were reduced to 1080p, 720p, 540p and 360p. Then, the reduced video clips were encoded [ITU-T H.264] at various bit rates and the decoded video clips were enlarged back to their original UHD resolutions using an interpolation method. The original and enlarged lower resolution video clips were displayed on a UHD TV display (75"). A subjective test was performed (ACR, 24 viewers after screening).

Figure 8 shows MOS distributions of the 12 source sequences at various bit rates and resolutions. Although there are some variations, it appears that using 720p or lower resolutions is not desirable for UHD signals when the bit rate is larger than or equal to 5 Mbit/s. Also, when the bit rate is lower than 5 Mbit/s [ITU-T H.264], the MOS values of UHD resolution signals were always lower than some of the other lower resolutions. For some UHD source sequences, the differences between 1080p and UHD resolutions were very small and increasing the bit rates did not produce improved perceptual video quality. Also, 540p or 360p resolutions may not be used for UHD programmes since their perceptual video quality is considerably lower. Thus, by carefully selecting the bit rate and transmission resolution based on content characteristics, it may be possible to optimize the bandwidth resources while providing good or acceptable perceptual video quality.



JSTR-OPTR(23)

Figure 8 – MOS distribution at various bit rates and resolutions

10 Usability of video quality models for deciding optimal bit rate and transmission resolution

In the following experiment, 12 UHD source sequences were chosen and the resolution reduced to 1080p, 720p, 540p and 360p. Then, the UHD sequences were encoded using [ITU-T H.264], and a subjective test performed using the 75" display at 1.5H distance, where H is the display height. The ACR method was used for subjective tests. Figure 9 shows the average MOS values for the processed video sequences and Figure 10 shows the MOS values for the UHD and 1080p resolution video sequences for each source sequence. Figure 11 shows enlarged images of a goat's hair at various bit rates.

Next, it the use objective video quality models to select optimal transmission resolutions and bit rates was explored. Several objective models were applied to the encoded UHD video signals at various resolutions. Four objective models were tested: PSNR, SSIM, VMAF and [ITU-T J.342]. Figure 12 shows the scatter plots between the objective scores and the MOS values for the various objective models. ITU-T J.342 and VMAF showed good performance with high correlations (ITU-T J.342: 0.936, VMAF: 0.895). The conventional PSNR shows rather poor performance (correlation coefficient: 0.694).

These results indicate that some of these objective video quality models may be used to determine the optimal bit rates and transmission resolutions for a given UHD programme. For example, Figure 12 shows MOS distribution at various bit rates and resolutions. Table 13 shows the MOS values and the corresponding ITU-T J.342 values along with resolutions and bit rates (UHD source 1 at 2160p). If the target bit rate is 10 Mbit/s, the objective model correctly suggests 1080p since the 1080p is the highest value (3.832) and the corresponding MOS is also the highest (3.96). If the target bit rate is 2-3 Mbit/s, the objective model recommends 720p at 2 Mbit/s or 540p at 3 Mbit/s. Although the ITU-T J.342 value of 540p is higher than that of 720p, the difference is very small. Furthermore, the 540p clip is encoded at 3 Mbit/s whereas the 720p clip is encoded at 2 Mbit/s. The operator can take into account this difference and slightly adjust the bit rate of 720p and find the optimal bit rate by repeatedly using the objective model [ITU-T J.342].

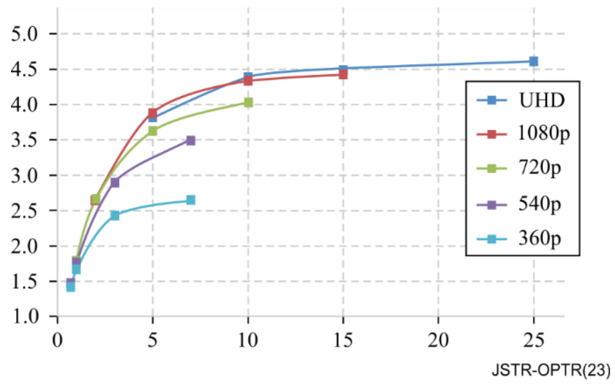


Figure 9 – MOS values for the various resolutions and bit rates (75-inch TV)

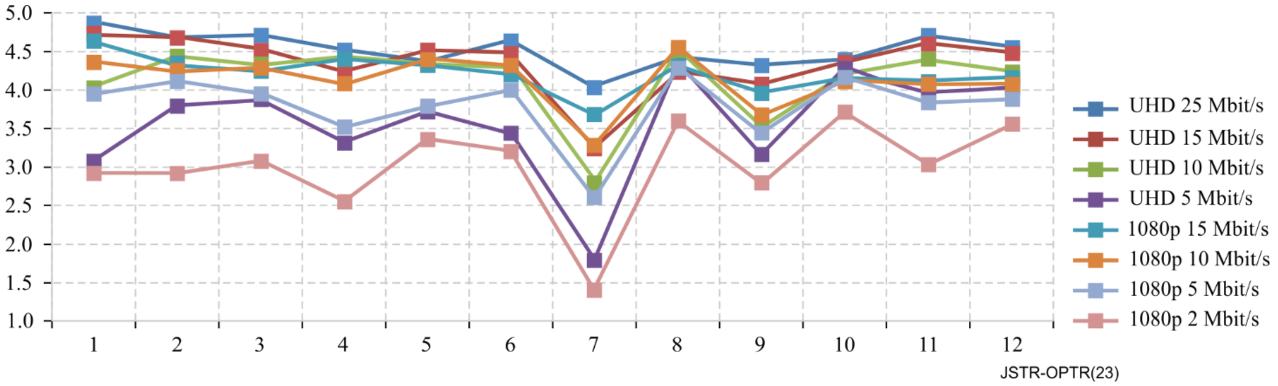


Figure 10 – MOS values for UHD and 1080p resolutions for each source sequence

Figure 11 – Enlarged sub-images of a goat hair

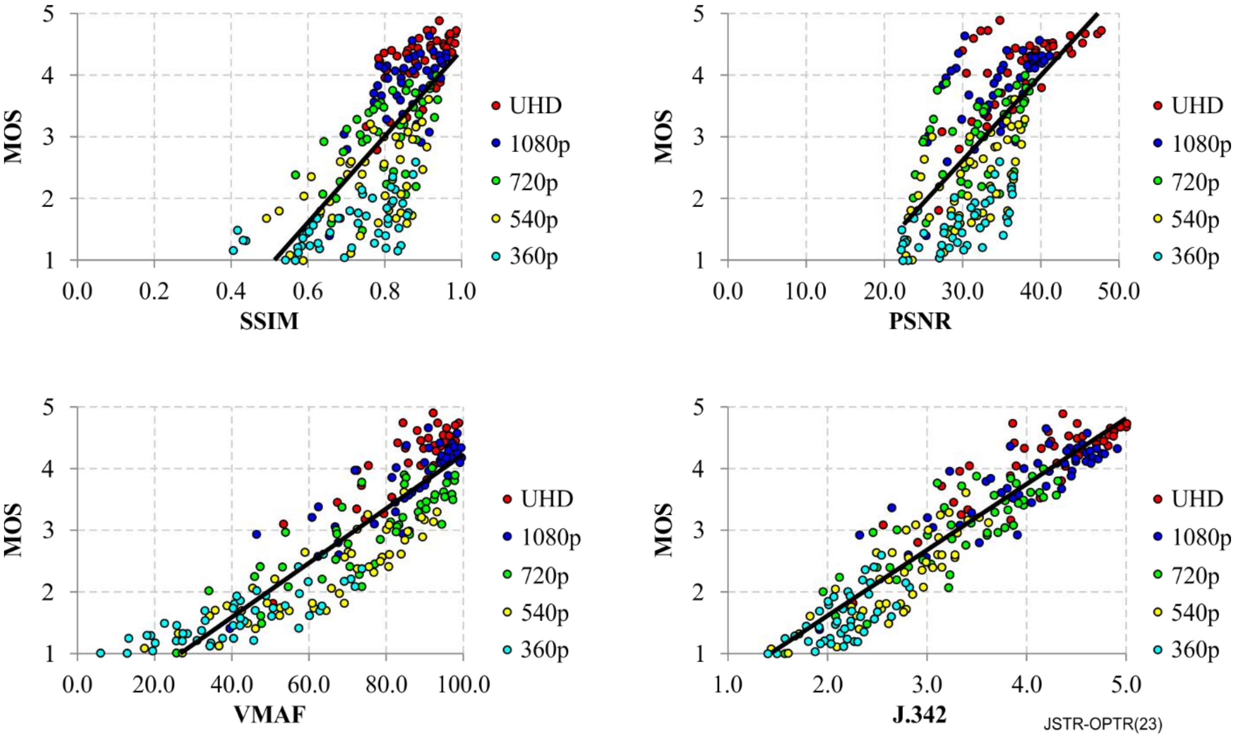


Figure 12 – Scatter plots between objective scores and MOS values for the various objective models

Table 13 – MOS values and corresponding predicted MOS values (ITU-T J.342) at various bit rates and resolutions (UHD source 1)

Resolution	Bit rate (Mbit/s)	MOS	J.342
540p	0.7	1.32	1.708
360p	0.7	1.24	1.566
720p	1	2	1.945
540p	1	1.68	1.913
360p	1	1.28	1.67
720p	2	2.96	2.459
1080p	2	2.92	2.316
540p	3	2.64	2.485
360p	3	1.68	1.909
1080p	5	3.96	3.223
720p	5	3.76	3.095
UHD	5	3.08	2.553
540p	7	3	2.785
360p	7	1.8	2.003
1080p	10	4.36	3.832
720p	10	3.88	3.492
UHD	10	4.04	3.417
1080p	15	4.64	4.192
UHD	15	4.72	3.859
UHD	25	4.88	4.362