ETSI TR 126 906 V15.0.0 (2018-07)

TECHNICAL REPORT

**Universal Mobile Telecommunications System (UMTS);
LTE;
Evaluation of High Efficiency Video Coding (HEVC)
for 3GPP services
(3GPP TR 26.906 version 15.0.0 Release 15)**

Reference

RTR/TSGS-0426906vf00

Keywords

LTE,UMTS

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00   Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Important notice

The present document can be downloaded from:
http://www.etsi.org/standards-search

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the only prevailing document is the print of the Portable Document Format (PDF) version kept on a specific network drive within ETSI Secretariat.

Users of the present document should be aware that the document may be subject to revision or change of status. Information on the current status of this and other ETSI documents is available at
https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

If you find errors in the present document, please send your comment to one of the following services:
https://portal.etsi.org/People/CommiteeSupportStaff.aspx

Copyright Notification

ETSI

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Report (TR) has been produced by ETSI 3rd Generation Partnership Project (3GPP).

The present document may refer to technical specifications or reports using their 3GPP identities, UMTS identities or GSM identities. These should be interpreted as being references to the corresponding ETSI deliverables.

The cross reference between GSM, UMTS, 3GPP and ETSI identities can be found under http://webapp.etsi.org/key/queryform.asp.

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Contents

Foreword

This Technical Report has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x   the first digit:

   1   presented to TSG for information;

   2   presented to TSG for approval;

   3   or greater indicates TSG approved document under change control.

y   the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z   the third digit is incremented when editorial only changes have been incorporated in the document.

1   Scope

The present document reports the evaluation of the High Efficiency Video Coding (HEVC) codec in 3GPP services. It provides an overview of the codec and a comparison to H.264 (AVC) codec. The support of H.264 (AVC) is mandated for 3GP-DASH (TS 26.247[18]), PSS (TS 26.234 [19]), MBMS (TS 26.346 [20]), 3GPP file format (TS 26.244 [21]), MTSI (TS 26.114 [22]) and MMS (TS 26.140 [23]) in Release 11. The present document reports on the performance of H.265 (HEVC) when used in 3GPP services for video coding in comparison to H.264 (AVC) and the performance of H.265 (HEVC) when used in 3GPP services for image coding in comparison to JPEG. Performance is evaluated in typical 3GPP service environments taking into account bandwidth and coding efficiency, user experience and complexity. Based on the performance results, recommendations are provided for the proper inclusion of H.265 (HEVC) in 3GPP services.

2   References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1]   3GPP TR 21 905: "Vocabulary for 3GPP Specifications".

[2]   Sullivan, G. J.; Ohm, J.-R.; Han, W.-J.; Wiegand T., "Overview of the High Efficiency Video Coding (HEVC) Standard," IEEE Trans. Circuits and Systems for Video Technology, vol.22, no.12, pp.1649-1668, Dec. 2012.

[3]   Bossen, F.; Bross, B.; Suhring, K.; Flynn D., "HEVC Complexity and Implementation Analysis," IEEE Trans. Circuits and Systems for Video Technology, vol.22, no.12, pp.1685-1696, Dec. 2012.

[4]   Vanne, J.; Viitanen, M.; Hamalainen, T. D.; Hallapuro A., "Comparative Rate-Distortion-Complexity Analysis of HEVC and AVC Video Cdecs," IEEE Trans. Circuits and Systems for Video Technology, vol.22, no.12, pp.1885-1898, Dec. 2012.

[5]   F. Bosen, "On software complexity," document JCTVC-G757 of JCT-VC, Geneva, Switzerland, Nov. 2011.

[6]   F. Bosen, "On software complexity: decoding 720p content on a tablet," document JCTVC-J0128 of JCT-VC, Stockholm, Sweden, Jul. 2012.

[7]   K. McCann, J.-Y. Choi, e al, "HEVC software player demonstration on mobile devices," document JCTVC-G988 of JCT-VC, Geneva, Switzerland, Nov. 2011.

[8]   K. Veera, R. Ganguly, e al, "A real-time ARM HEVC decoder implementation," document JCTVC-H0693 of JCT-VC, José, CA, USA, Feb. 2012.

[9]   J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards – including high efficiency video coding (HEVC)," IEEE Trans. Circuits and Systems for Video Technology, December 2012.

[10]   V. Baroncini, G. J. Sullivan, and J.-R. Ohm, "Report on preliminary subjective testing of HEVC compression capability," document JCTVC-H1004 of JCT-VC, San José, USA, Feb. 2012.

[11]   Y. Zhao, et al, "Coding efficiency comparison between HM5.0 and JM16.2 based on PQI, PSNR and SSIM," document JCTVC-H0063 of JCT-VC, San José, USA, Feb. 2012.

[12]   T. K. Tan, A. Fujibayashi, Y. Suzuki and J. Takiue, "Objective and subjective evaluation of HM5.0," document JCTVC-H0116 of JCT-VC, San José, USA, Feb. 2012.

[13]   B. Li, G. J. Sullivan, and J. Xu, "Comparison of compression performance of HEVC working draft 7 with AVC high profile", document JCTVC-J0236 of JCT-VC, Stockholm, Sweden, July 2012.

[14]   Recommendation ITU-T P.910 (04-08): "Subjective video quality assessment methods for multimedia applications".

[15]   Recommendation ITU-R BT.500 (01-12): "Methodology for the subjective assessment of the quality of television pictures".

[16]   ISO/IEC PDTR 29170-1: "Information technology -- Advanced image coding and evaluation methodologies -- Part 1: Guidelines for codec evaluation".

[17]   3GPP TS 26.273: "ANSI-C code for the fixed-point Extended Adaptive Multi–Rate - Wideband (AMR-WB+) speech codec".

[18]   3GPP TS 26.247: "Transparent end-to-end Packet-switched Streaming Service (PSS); Progressive Download and Dynamic Adaptive Streaming over HTTP (3GP-DASH)".

[19]   3GPP TS 26.234: "Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs".

[20]   3GPP TS 26.346: "Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs".

[21]   3GPP TS 26.244: "Transparent end-to-end packet switched streaming service (PSS); 3GPP file format (3GP)".

[22]   3GPP TS 26.114: "IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction".

[23]   3GPP TS 26.140: "Multimedia Messaging Service (MMS); Media formats and codecs".

[24]   Recommendation ITU-T P.800 (08/1996): "Methods for subjective determination of transmission quality".

3   Definitions and abbreviations

3.1   Definitions

For the purposes of the present document, the terms and definitions given in TR 21.905 [1] apply.

3.2   Abbreviations

For the purposes of the present document, the abbreviations given in TR 21.905 [1] and the following apply.

AVC	Advanced Video Coding
BD	Bjontegaard Delta
BLA	Broken Link Access
CABAC	Context Adaptive Binary Arithmetic Coding
CAVLC	Context Adaptive Variable Length Coding
CPB	Coded Picture Buffer
CRA	Clean Random Access
CTU	Coding Tree Unit
DASH	Dynamic Adaptive Streaming over HTTP
DCT	Discrete Cosine Transform
DPB	Decoded Picture Buffer
DST	Discrete Sine Transform
GDR	Gradual Decoding Refresh
GOP	Group of Pictures
HEVC	High Efficiency Video Coding
HRD	Hypothetical Reference Decoder
IDR	Instantaneous Decoding Refresh
IRAP	Intra Random Access Point
ISOBMFF	ISO Base Media File Format

MANE	Media Aware Network Element
MOS	Mean Opinion Score
NAL	Network Abstraction Layer
JCT-VC	Joint Collaborative Team on Video Coding
MBMS	Multimedia Broadcast Multicast Service
MMCO	Memory Management Control Operation
MMS	Multimedia Messaging Service
MTSI	Multimedia Telephony Service for IMS
MTU	Maximum Transmission Unit
PPS	Picture Parameter Set
QP	Quantization Parameter
PSNR	Peak Signal Noise Ratio
PSS	Packet-switched Streaming
RADL	Random Access Decodable Leading
RAP	Random Access Period
RASL	Random Access Skipped Leading
RPLC	Reference Picture List Construction
RPS	Reference Picture Set
SAO	Sample Adaptive Offset
SEI	Supplemental Enhancement Information
SAP	Stream Access Point
SPS	Sequence Parameter Set
TSA	Temporal Sub-layer Access
STSA	Stepwise Temporal Sub-layer Access
VPS	Video Parameter Set
WPP	Wavefront Parallel Processing

4 Introduction

The present document provides a study on evaluation of the H.265 (HEVC) video codec for 3GPP services. Use cases and technical solutions are investigated regarding a variety of setups using 3GPP's streaming, multicast/broadcast, download and progressive download as well as conversational services. Clause 5 provides a technical overview of the H.265 (HEVC) video codec, highlighting its key differences compared to H.264 (AVC). Clause 6 provides a description of the test cases and experimental setup for evaluation of H.265 (HEVC) versus H.264 (AVC). Clauses 7 and 8 provide experimental results for video and image coding, respectively. Clause 9 provides the conclusions for evaluation of HEVC for each of the 3GPP services.

5 Overview of H.265 (HEVC)

5.1 Key coding-tool features of H.265 (HEVC) and differences versus H.264 (AVC)

Similar to earlier hybrid-video-coding based standards, including H.264 (AVC), the following basic video coding design is employed by H.265 (HEVC). Prediction signal is first formed either by intra or motion compensated prediction, and the residual (the difference between the original and the prediction) is then coded. The gains in coding efficiency are achieved by redesigning and improving almost all parts of the codec over earlier designs. In addition, H.265 (HEVC) includes several tools to make the implementation on parallel architectures easier. Below is a summary of key H.265 (HEVC) coding-tool features, and a more elaborate list can be found in [2]:

- **Quadtree block and transform structure:** One of the major tools that contribute significantly to the coding efficiency of H.265 (HEVC) is the usage of flexible coding blocks and transforms, which are defined in a hierarchical quad-tree manner. Unlike H.264 (AVC), where the basic coding block is a macroblock of fixed size 16x16, H.265 (HEVC) defines a Coding Tree Unit (CTU) of a maximum size of 64x64. Each CTU can be divided into smaller units in a hierarchical quad-tree manner and can represent smaller blocks of size 4x4. Similarly, the transforms used in H.265 (HEVC) can have different sizes, starting from 4x4 and going up to 32x32.

Utilizing large blocks and transforms contribute to the major gain of H.265 (HEVC), especially at high resolutions:

- **Entropy coding:** H.265 (HEVC) uses a single entropy coding engine, which is based on Context Adaptive Binary Arithmetic Coding (CABAC), whereas H.264 (AVC) uses two distinct entropy coding engines. CABAC in H.265 (HEVC) shares many similarities with CABAC of H.264 (AVC), but contains several improvements. Those include improvements in coding efficiency and lowered implementation complexity, especially for parallel architectures.

- **In-loop filtering:** H.264 (AVC) includes an in-loop adaptive deblocking filter, where the blocking artefacts around the transform edges in the reconstructed picture are smoothed to improve the picture quality and compression efficiency. In H.265 (HEVC), a similar deblocking filter is employed but with somewhat lower complexity. In addition, pictures undergo a subsequent filtering operation called Sample Adaptive Offset (SAO), which is a new design element in H.265 (HEVC). SAO basically adds a pixel level offset in an adaptive manner and usually acts as a de-ringing filter. It is observed that SAO improves the picture quality, especially around sharp edges contributing substantially to visual quality improvements of H.265 (HEVC).

- **Motion prediction and coding:** There have been a number of improvements in this area that are summarized as follows:

  - Merge and Advanced Motion Vector Prediction (AMVP) modes: The motion information of a prediction block can be inferred from the spatially or temporally neighbouring blocks. This is similar to the DIRECT mode in H.264 (AVC) but includes new aspects to incorporate the flexible quad-tree structure and methods to improve the parallel implementations. In addition, the motion vector predictor can be signalled for improved efficiency.

  - High precision interpolation: The interpolation filter length is increased to 8-tap from 6-tap, which improves the coding efficiency but also comes with increased complexity. In addition, interpolation filter is defined with higher precision without any intermediate rounding operations to further improve the coding efficiency.

- **Intra prediction and intra coding:** Similar to motion prediction, intra prediction has many improvements, which can be summarized as:

  - Compared to 8 intra prediction modes of H.264 (AVC), H.265 (HEVC) supports angular intra prediction with 33 directions. This increased flexibility improves both objective coding efficiency and visual quality as the edges can be better predicted and ringing artefacts around the edges are reduced.

  - The reference samples are adaptively smoothed based on the prediction direction. In addition, to avoid contouring artefacts, a new interpolative prediction generation is included to improve the visual quality.

  - Discrete Sine Transform (DST) is utilized instead of traditional Discrete Cosine Transform (DCT) for 4x4 intra transform blocks.

- **Other coding-tool features:** H.265 (HEVC) includes some tools for lossless coding and efficient screen content coding:

  - Lossless coding: H.265 (HEVC) allows certain part of the coded picture to be coded in a lossless manner by setting a dedicated flag equal to 1.

  - Screen content coding: H.265 (HEVC) includes some tools to better code computer generated screen content, such as skipping the transform coding for certain blocks. These tools are particularly useful for example when streaming the user-interface of a mobile device to a large display.

5.2   Complexity of H.265 (HEVC)

Measuring the complexity of a video codec is a difficult task, due to different constraints of different architectures. For example, for hardware implementations CABAC might not be very problematic but for software implementations it could become a bottleneck, especially at higher bitrates. Nevertheless, there had been several studies that analyses the complexity of H.265 (HEVC), and the conclusions could be roughly summarized as follows (see also [3] and [4]):

- H.265 (HEVC) Decoder: Even though many parts of H.265 (HEVC) are more complex than their counterparts in H.264 (AVC) (e.g. motion compensation, intra prediction), some parts are easier to implement (e.g. CABAC, deblocking filter). Therefore, the additional complexity of H.265 (HEVC) decoder over H.264 (AVC) decoder is not expected to be substantial.

- H.265 (HEVC) Encoder: As well known, the standard does not define how the encoding is performed, which means there will be various encoders with different complexity-quality trade-offs. However, it is estimated that

the encoder complexity of H.265 (HEVC) needs to be higher than that of H.264 (AVC), in order to achieve the coding efficiency gains of H.265 (HEVC). The main reason for that is that there exists higher number of combinations to be tested during the rate-distortion optimization as H.265 (HEVC) supports more flexible partitioning of blocks and transforms. It should be noted that the parallel processing tools are mostly useful for encoders and their efficient utilization is expected to improve the complexity aspects of H.265 (HEVC) encoders. It is also expected that there will be significant efforts over the coming years to develop efficient methods for H.265 (HEVC) encoding.

Some more existing complexity analyses of H.265 (HEVC) and H.264 (AVC) can be found in [3] to [8], where [3] and [5] to [8] reported real-time H.265 (HEVC) decoding by H.265 (HEVC) decoder implementations based on ARM platforms.

5.3 Systems and transport interfaces of H.265 (HEVC) and differences versus H.264 (AVC)

H.265 (HEVC) inherited the basic systems and transport interfaces designs, such as parameter sets and network abstraction layer (NAL) units based syntax structure, the hierarchical syntax and data unit structure from sequence-level parameter sets, multi-picture-level or picture-level parameter sets, slice-level header parameters, lower-level parameters, supplemental enhancement information (SEI) message mechanisms, hypothetical reference decoder (HRD) based video buffering model, and so on.

In the following, a list of differences in these aspects compared to H.264 (AVC) is summarized:

- **Video parameter set:** A new type of parameter set, called video parameter set (VPS), was introduced. The VPS provides a "big picture" of a bitstream, including what types of operation points are provided, the profile, tier, and level of the operation points, and some other high-level properties of the bitstream that can be used as the basis for session negotiation and content selection, etc.

- **Profile, tier and level:** The profile, tier and level syntax structure that can be included in both VPS and sequence parameter set (SPS) includes 12 bytes data for the entire bitstream, and possibly include more profile, tier and level information for temporal scalable layers, which are referred to as sub-layers in the H.265 (HEVC) specification:

  - The profile indicator indicates the "best viewed as" profile when the bistream conforms to multiple profiles, like the major brand as in 3GPP file format and other ISO base media file format (ISOBMFF) based file formats.

  - The profile, tier and level syntax structure also includes the indications of whether the bitstream is free of frame-packed content, whether the bitstream is free of interlaced source and free of field pictures, i.e. contains only frame pictures of progressive source, such that clients/players with no special support of post-processing functionalities for handling of frame-packed contents, or contents with interlaced source or field pictures can stay away from those contents.

- **Bitstream and elementary stream:** H.265 (HEVC) includes a definition of elementary stream, which is new compared to H.264 (AVC). An elementary stream consists of a sequence of one or more bitstreams. An elementary stream that consists of two or more bitstreams would typically have been formed by splicing together two or more bitstreams (or parts thereof). When an elementary stream contains more than one bitstream, the last NAL unit of the last access unit of a bitstream (except the last bitstream in the elementary stream) contains an end of bitstream NAL unit and the first access unit of the subsequent bitstream is an intra random access point (IRAP) access unit. This IRAP access unit may be a clean random access (CRA), broken link access (BLA), or instantaneous decoding refresh (IDR) access unit.

- **Improved random accessibility support:** H.265 (HEVC) includes signalling in NAL unit header, through NAL unit types, of IRAP pictures beyond IDR pictures. Three types of IRAP pictures, namely IDR, CRA, and BLA pictures, are supported, wherein IDR pictures are conventionally referred to as closed group-of-pictures (closed-GOP) random access points, while CRA and BLA pictures are those conventionally referred to as open-GOP random access points.

  - BLA pictures usually originate from splicing of two bitstreams or part thereof at a CRA picture, e.g. during stream switching.

  - To enable better systems usage of IRAP pictures, altogether six different NAL units are defined to signal the properties of the IRAP pictures, which can be used to better match the stream access point (SAP) types as

defined in the ISOBMFF, which are utilized for random access support in both 3GP-DASH and MPEG DASH.

- Pictures following an IRAP picture in decoding order and preceding the IRAP picture in output order are referred to as leading pictures associated with the IRAP picture. There are two types of leading pictures, namely random access decodable leading (RADL) pictures and random access skipped leading (RASL) pictures. RADL pictures are decodable when random access starts at the associated IRAP picture, and RASL pictures are not decodable when random access starts at the associated IRAP picture and are usually discarded.

- H.265 (HEVC) provides mechanisms to enable the specification of conformance of bitstreams with RASL pictures being discarded, thus to provide a standard-complaint way to enable systems components to discard RASL pictures when needed.

- **Improved temporal scalability support:** H.265 (HEVC) includes an improved support of temporal scalability, by inclusion of the signalling of temporal ID in the NAL unit header, the restriction that pictures of a particular temporal sub-layer cannot be used for inter prediction reference by pictures of a higher temporal sub-layer, the sub-bitstream extraction process, and the requirement that each sub-bitstream extraction output be a conforming bitstream. Media-aware network elements (MANEs) can utilize the temporal ID in the NAL unit header for stream adaptation purposes based on temporal scalability.

- **Improved temporal layer switching support:** H.265 (HEVC) specifies, through NAL unit types present in the NAL unit header, the signalling of temporal sub-layer access (TSA) and stepwise temporal sub-layer access (STSA):

  - A TSA picture and pictures following the TSA picture in decoding order do not use pictures prior to the TSA picture in decoding order with TemporalId greater than or equal to that of the TSA picture for inter prediction reference. A TSA picture enables up-switching, at the TSA picture, to the sub-layer containing the TSA picture or any higher sub-layer, from the immediately lower sub-layer.

  - An STSA picture does not use pictures with the same TemporalId as the STSA picture for inter prediction reference. Pictures following an STSA picture in decoding order with the same TemporalId as the STSA picture do not use pictures prior to the STSA picture in decoding order with the same TemporalId as the STSA picture for inter prediction reference. An STSA picture enables up-switching, at the STSA picture, to the sub-layer containing the STSA picture, from the immediately lower sub-layer.

- **Sub-layer reference or non-reference pictures:** The concept and signalling of reference/non-reference pictures in H.265 (HEVC) are different from H.264 (AVC). In H.264 (AVC), if a picture may be used by any other picture for inter prediction reference, it is a reference picture; otherwise it is a non-reference picture, and this is signalled by two bits in the NAL unit header. In H.265 (HEVC), a picture is called a reference picture only when it is marked as "used for reference". In addition, the concept of sub-layer reference picture was introduced. If a picture may be used by another other picture with the same TemporalId for inter prediction reference, it is a sub-layer reference picture; otherwise it is a sub-layer non-reference picture. Whether a picture is a sub-layer reference picture or a sub-layer non-reference picture is signalled through NAL unit type values.

- **Improved extensibility:** Besides the temporal ID in the NAL unit header, H.265 (HEVC) also includes the signalling of six-bit layer ID in the NAL unit header, which is equal to 0 for a single-layer bitstream. Extension mechanisms have been included in VPS, SPS, PPS, SEI NAL unit, slice headers, and so on. All these extension mechanisms enable future extensions in a backward compatible manner, such that bitstreams encoded according to potential future H.265 (HEVC) extensions can be fed to then-legacy decoders (e.g. H.265 (HEVC) version 1 decoders) and the then-legacy decoder can decode and output the base layer bitstream.

- **Bitstream extraction:** H.265 (HEVC) includes bitstream extraction process as an integral part of the overall decoding process, as well as specification of the use of the bitstream extraction process in description of bitstream conformance tests as part of the hypothetical reference decoder (HRD) specification.

- **Improved reference picture management:** H.265 (HEVC) includes a different way of reference picture management, including reference picture marking and removal from the decoded picture buffer (DPB) as well as reference picture list construction (RPLC). Instead of the sliding window plus adaptive memory management control operation (MMCO) based reference picture marking mechanism in H.264 (AVC), H.265 (HEVC) specifies a reference picture set (RPS) based reference picture management and marking mechanism, and the RPLC is consequently based on the RPS mechanism.

- A reference picture set consists of a set of reference pictures associated with a picture, consisting of all reference pictures that are prior to the associated picture in decoding order, that may be used for inter prediction of the associated picture or any picture following the associated picture in decoding order. The reference picture set consists of five lists of reference pictures; RefPicSetStCurrBefore, RefPicSetStCurrAfter, RefPicSetStFoll, RefPicSetLtCurr and RefPicSetLtFoll. RefPicSetStCurrBefore, RefPicSetStCurrAfter and RefPicSetLtCurr contains all reference pictures that may be used in inter prediction of the current picture and that may be used in inter prediction of one or more of the pictures following the current picture in decoding order. RefPicSetStFoll and RefPicSetLtFoll consists of all reference pictures that are not used in inter prediction of the current picture but may be used in inter prediction of one or more of the pictures following the current picture in decoding order.

- RPS provides an "intra-coded" signalling of the DPB status, instead of an "inter-coded" signalling, mainly for improved error resilience.

- The RPLC process in H.265 (HEVC) is based on the RPS, by signalling an index to an RPS subset for each reference index. The RPLC process has been simplified compared to that in H.264 (AVC), by removal of the reference picture list modification (also referred to as reference picture list reordering) process.

- **Ultralow delay support:** H.265 (HEVC) specifies a sub-picture-level HRD operation, for support of the so-called ultralow delay. The mechanism specifies a standard-complaint way to enable delay reduction below one picture interval. Sub-picture-level coded picture buffer (CPB) and DPB parameters may be signalled, and utilization of these information for the derivation of CPB timing (wherein the CPB removal time corresponds to decoding time) and DPB output timing (display time) is specified. Decoders are allowed to operate the HRD at the conventional access-unit-level, even when the sub-picture-level HRD parameters are present.

- **Parallel processing support:** H.265 (HEVC) is the first video coding standard that includes some features that are specifically to enable parallel coding, particularly parallel encoding. These tools are tiles and wavefront parallel processing (WPP), which cannot be applied at the same time within a coded video sequence (as defined in the H.265 (HEVC) specification).

  - In WPP, the picture is partitioned into single rows of CTUs. Entropy decoding and prediction are allowed to use data from CTUs in other partitions. Parallel processing is possible through parallel decoding of CTU rows, where the start of the decoding of a CTU row is delayed by two CTUs, so to ensure that data related to a CTU above and to the right of the subject CTU is available before the subject CTU is being decoded. Using this staggered start (which appears like a wavefront when represented graphically), parallelization is possible with up to as many processors/cores as the picture contains CTU rows. Because in-picture prediction between neighbouring CTU rows within a picture is permitted, the required inter-processor/inter-core communication to enable in-picture prediction can be substantial. The WPP partitioning does not result in the production of additional NAL units compared to when it is not applied, thus WPP is not a tool for MTU size matching. However, if MTU size matching is required, slices and dependent slice segments can be used with WPP, with certain coding overhead.

  - Tiles define horizontal and vertical boundaries that partition a picture into tile columns and rows. The scan order of CTUs is changed to be local within a tile (in the order of a CTU raster scan of a tile), before decoding the top-left CTU of the next tile in the order of tile raster scan of a picture. Similar to slices, tiles break in-picture prediction dependencies as well as entropy decoding dependencies. However, they do not need to be included into individual NAL units (same as WPP in this regard); hence tiles cannot be used for MTU size matching, though slices and dependent slice segments can be used in combination for that purpose. Each tile can be processed by one processor/core, and the inter-processor/inter-core communication required for in-picture prediction between processing units decoding neighbouring tiles is limited to conveying the shared slice header in cases a slice is spanning more than one tile, and loop filtering related sharing of reconstructed samples and metadata. When more than one tile or WPP segment is included in a slice, the entry point byte offset for each tile or WPP segment other than the first one in the slice is signalled in the slice header.

- **New SEI messages:** H.265 (HEVC) inherits many SEI messages from H.264 (AVC) with changes in syntax and/or semantics to make them applicable to H.265 (HEVC). Additionally, H.265 (HEVC) includes some new SEI messages; some of them are summarized below.

  - The display orientation SEI message signals the recommended anticlockwise rotation of the decoded picture (after applying horizontal and/or vertical flipping when needed) prior to display. This SEI message was also agreed to be included into H.264 (AVC).

- The active parameter sets SEI message includes the IDs of the active video parameter set and the active sequence parameter set, and can be used to activate VPSs and SPSs. In addition, the SEI message includes the following indications:

  - An indication of whether "full random accessibility" is supported (when supported, all parameter sets needed for decoding of the remaining of the bitstream when random accessing from the beginning of the current coded video sequence by completely discarding all access units earlier in decoding order are present in the remaining bitstream and all coded pictures in the remaining bitstream can be correctly decoded).

  - An indication of whether there is any parameter set within the current coded video sequence that updates another parameter set of the same type preceding in decoding order. An update of a parameter set refers to the use of the same parameter set ID but with some other parameters changed. If this property is true for all coded video sequences in the bitstream, then all parameter sets can be sent out-of-band before session start.

- The region refresh information SEI message can be used together with the recovery point SEI message (present in both H.264 (AVC) and H.265 (HEVC)) for improved support of gradual decoding refresh (GDR). This supports random access from inter-coded pictures, wherein complete pictures can be correctly decoded or recovered after an indicated number of pictures in output/display order.

- The decoding unit information SEI message provides coded picture buffer removal delay information for a decoding unit. The message can be used in very-low-delay buffering operation.

- The structure of pictures SEI message provides information on the NAL unit types, picture order count values and prediction dependencies of a sequence of pictures. The SEI message can be used for example for concluding which impact a lost picture has on other pictures.

- The decoded picture hash SEI message provides a checksum derived from the sample values of a decoded picture. It can be used for detecting whether a picture was correctly received and decoded.

5.4   H.265 (HEVC) for image coding

H.265 (HEVC) includes a Main Still Picture profile to efficiently code still images. This profile utilizes the same coding tools as the Main Profile of H.265 (HEVC) but can be used for encoding/decoding of still images. H.265 (HEVC) Main Still Picture profile is believed to be very useful for coding still images because of the following reasons:

- High coding efficiency: Compared to legacy still picture codecs, H.265 (HEVC) provides significant benefits in compression capability.

- Tile support: H.265 (HEVC) includes mechanism to divide a picture into regions called Tiles and to code those independently. This "spatial random access" provides various useful functionalities, such as easy browsing of extremely large pictures.

- Using the same coding engine as for video coding: H.265 (HEVC) Main Still Picture profile uses the same tools as the Main profile for video coding. This means that all the H.265 (HEVC) implementations will most likely come with a support for the Main Still Picture profile as well, because no extra codec implementation is needed, thus it makes the deployment of this image codec relatively easy.

6   Test case definitions

6.1   Introduction

For the evaluation of H.265 (HEVC) for different 3GPP multimedia services (3GP-DASH, MMS, PSS, MBMS, MTSI, and IMS Messaging and Presence), coding efficiency tests comparing H.265 (HEVC) and H.264 (AVC) for video coding as well as comparing H.265 (HEVC) and JPEG for image coding need to be performed. Besides, some analysis of complexity impacts should be made. The decision on whether to support H.265 (HEVC) for a particular 3GPP multimedia service should be made based on both coding efficiency test results and complexity analysis.

It should also be noted that it is expected that the decision would be made separately for each service.

The present document describes test cases and test procedures for evaluation of H.265 (HEVC) for 3GPP multimedia services in general as well as for specific 3GPP services.

For reference, some existing coding performance analyses of H.265 (HEVC) and H.264 (AVC) can be found in [4] and [9] to [13].

6.2 Test cases for evaluation of H.265 (HEVC) for video coding

6.2.1 Generic test cases

The generic test cases discussed in this clause, except for the random access point (RAP) period, apply for coding efficiency evaluation of H.265 (HEVC) in all 3GPP video services. The RAP period parameter applies to 3GP-DASH, RTP/RTSP based streaming as specified in PSS, MBMS, and MMS, but not MTSI.

Specific test cases for a particular service are specified based on the generic test cases specified here. For example, the test cases for 3GP-DASH are specified in clause 6.2.5.

The test cases included here are expected to target mainly two aspects:

- Improvements in quality for the same bitrate compared to H.264 (AVC)

- Bitrate savings for the same quality compared to H.264 (AVC)

In order to generate relevant test results, the characteristics of 3GPP streaming service environments, especially DASH should be taken into account. These include, but are not limited to target bitrates (e.g. in the range from about a hundred kbit/s up to 8 MBit/s), spatial resolutions (such as 240p, 480p, 720p, and 1080p) and temporal resolutions (such as 24 fps, 30 fps, 50 fps, and 60 fps), maximum random access points distance (1 or 2 seconds).

Specifically the test case parameters in Table 1 are recommended. Note that it is not expected to produce combinations of all parameters below and more work is necessary to produce relevant test cases with suitable parameter combinations.

**Table 1: Parameters and Parameter Settings for evaluations of H.265 (HEVC)
compared to H.264 (AVC)**

Parameter	Settings
Bitrates	Ranging from 100 kbit/s to 8 Mbps
Spatial resolutions	240 p, 480 p, 720 p, 1080 p
Frame rates	24 fps, 30 fps, 50 fps, 60 fps
RAP distance	1 s, 2 s

6.2.2 Test sequences and codec software

6.2.2.1 Test sequences

The test sequences used by JCT-VC for development of H.265 (HEVC) are used for the evaluation. Additional test sequences could be included in the tests if they become available. The test sequences and their characteristics are described in Table 2.

Table 2: Test sequences and their characteristics

Class	Sequence	Spatial resolution	Frame rate
Class B	Kimono	1920x1080	24 fps
	ParkScene	1920x1080	24 fps
	Cactus	1920x1080	50 fps
	BasketballDrive	1920x1080	50 fps
	BQTerrace	1920x1080	60 fps
Class C	BasketballDrill	832x480	50 fps
	BQMall	832x480	60 fps
	PartyScene	832x480	50 fps
	RaceHorses	832x480	30 fps
	Kimono_480p	832x480	24 fps
	ParkScene_480p	832x480	24 fps
	Cactus_480p	832x480	50 fps
	BasketballDrive_480p	832x480	50 fps
	BQTerrace_480p	832x480	60 fps
Class D	BasketballPass	416x240	50 fps
	BQSquare	416x240	60 fps
	BlowingBubbles	416x240	50 fps
	RaceHorses	416x240	30 fps
	Kimono_240p	416x240	24 fps
	ParkScene_240p	416x240	24 fps
	Cactus_240p	416x240	50 fps
	BasketballDrive_240p	416x240	50 fps
	BQTerrace_240p	416x240	60 fps
Class E	Kimono_720p	1280x720	24 fps
	ParkScene_720p	1280x720	24 fps
	Cactus_720p	1280x720	50 fps
	BasketballDrive_720p	1280x720	50 fps
	BQTerrace_720p	1280x720	60 fps
NOTE 1	The Class-C test sequences Kimono_480p, ParkScene_480p, Cactus_480p, BasketballDrive_480p, and BQTerrace_480p were generated by firstly down-sampling the corresponding Class-B test sequences using the down-sampling filter used by the JCT-VC for the SHVC work with 2x down-sampling ratio in each dimension (from 1920x1080 to 960x540), followed by cropping 64 luma samples from both left and right, and 30 luma samples from both top and bottom.		
NOTE 2	The Class-D test sequences Kimono_240p, ParkScene_240p, Cactus_240p, BasketballDrive_240p, and BQTerrace_240p were generated by down-sampling the corresponding Class-C test sequences using the down-sampling filter used by the JCT-VC for the SHVC work with 2x down-sampling ratio in each dimension (from 832x480 to 416x240), with no cropping.		
NOTE 3	The Class-E test sequences Kimono_720p, ParkScene_720p, Cactus_720p, BasketballDrive_720p, and BQTerrace_720p were generated by down-sampling the corresponding Class-B test sequences using the down-sampling filter used by the JCT-VC for the SHVC work with 1.5x down-sampling ratio in each dimension (from 1920x1080 to 1 280x720), with no cropping.		

The down-sampling filter used for generation of the test sequences is described in clause 6.2.2.1.1.

6.2.2.1.1 Down-sampling filter

The filters used to generate Class C/D/E test sequences support both 1.5x and 2x down-sampling ratio. The filters are Cosine windowed Sinc function with cut-off frequency at 0.9π in the down-sampling domain to preserve high frequency details. The filters were designed in odd-length symmetric such that the down-sampled videos have zero phase shift compared with the original videos.

The coefficients of the 1.5x-down-sampling filter are shown in Table 3. The corresponding impulse and frequency responses are shown in Figure 1. The coefficients of the 2x-down-sampling filter are shown in Table 4. The corresponding impulse and frequency responses are shown in Figure 2.

Table 3: 1.5x down-sampling filter

Phase	Filter Coefficients
Integer	[0 5 -6 -1 0 37 76 37 -10 -6 5 0 0] /128
1	[-1 5 -3 -1 2 29 75 45 -7 -8 5 0 0] /128
2	[-1 4 -1 -1 3 22 73 52 -3 -10 4 1 0] /128
3	[-1 4 1 -1 3 14 70 59 2 -12 3 2 -1] /128
4	[-1 3 2 -1 3 8 65 65 8 -13 2 3 -1] /128
5	[-1 2 3 -1 2 2 59 70 14 -13 1 4 -1] /128
6	[0 1 4 -1 0 -3 52 73 22 -13 -1 4 -1] /128
7	[0 0 5 -0 8 -7 45 75 29 -12 -3 5 0] /128

Table 4: 2.0x down-sampling filter

Phase	Filter Coefficients
Integer	[2 -3 -9 6 39 58 39 6 -9 -3 2 0] /128
1	[1 -1 -8 -1 31 57 47 13 -7 -5 1 0] /128
2	[1 0 -7 -5 22 53 53 22 -5 -7 0 1] /128

Figure 1: (a) Impulse response and (b) frequency response of the 1.5x-down-sampling filter: void

Figure 2: (a) Impulse response and (b) frequency response of the 2x-down-sampling filter: void

Class E test sequences were generated by directly applying the 1,5x down-sampling filter on Class B sequences. Class C sequences were generated by additionally applying a cropping process to maintain the original picture aspect ratio. For example, the original 1920x1080 test sequences were firstly down-sampled 2x into 960x540 sequences, and then the additional cropping is applied to crop 64 luma samples from left/right and 30 luma samples from top/bottom evenly to get 832x480 Class C test sequences.

Class D (416x240) test sequences can be generated by further applying the 2x-down-sampling filters on the corresponding Class C (832x480) test sequences without additional cropping process.

6.2.2.2 Codec software

For coding efficiency tests, HM version 10 is used for H.265 (HEVC) and JM version 18.4 is used for H.264 (AVC). For all submitted results the exact version and the configuration files from the test software should be provided.

Companies that would like to report test results should also be allowed to use other implementations of H.265 (HEVC) and H.264 (AVC).

6.2.3 Quality evaluation metrics

To evaluate the quality of H.265 (HEVC) within the test cases, Bjontegaard Delta Bitrate (BD rate) metric is used. BD rate measures the difference between two Rate Distortion curves and it is a widely used and established measure for comparing the performances of different video coding algorithms.

For each test sequence, several encodings are performed at 10 different QPs ranging from very low quality to high quality. The QP settings for H.265 (HEVC) are given as follows: 19, 22, 25, 28, 31, 34, 37, 40, 43 and 46. From this data, the following information is gathered:

- Coding efficiency improvement of H.265 (HEVC) over H.264 (AVC) for different bitrates and resolutions

- Suitable bitrate range for H.265 (HEVC) for different video resolutions

- Gains of H.265 (HEVC) over H.264 (AVC) for sequences with different characteristics (texture / motion complexity)

6.2.4 Complexity analysis

For MTSI, analyses of both encoding and decoding complexities are required. For other services, encoding complexity is not so much relevant, thus only decoding complexity analysis is required.

Both algorithmic and numerical analyses are encouraged to be reported.

6.2.5   Test conditions for 3GP-DASH, PSS, and MBMS

6.2.5.1   General testing settings

The general testing parameters as listed in Table 1 are recommended for evaluations of H.265 (HEVC) for video coding in 3GP-DASH, PSS, and MBMS.

6.2.5.2   Test sequences

The test sequences as described in Table 2 are used. Results based on additional test sequences are welcome but not required.

6.2.5.3   Encoding settings

- Profile

    H.265 (HEVC) Main profile and H.264 (AVC) High profile are used.

- QP configuration

    Fixed QP configuration is used without rate control to avoid uncertainty due to different rate control algorithms. Cascaded QP setting (e.g. higher QP for P pictures than I pictures, higher QP for B pictures than P pictures, and higher QP for higher temporal level than lower temporal level in hierarchical coding structures) is allowed and similar QP cascading strategy is used for both H.265 (HEVC) and H.264 (AVC).

- GOP structures

    Hierarchical B coding structures with GOP size of 8 is used for both H.265 (HEVC) and H.264 (AVC).

- IRAP pictures

    Two types of tests are performed that uses open GOP or closed-GOP configuration for random access. For closed-GOP test, IRAP pictures are IDR pictures for both H.265 (HEVC) and H.264 (AVC). For open-GOP test, IRAP pictures are clean random access (CRA) pictures for H.265 (HEVC) and open-GOP intra pictures (indicated by recovery point SEI messages) for H.264 (AVC). The first picture is an IDR picture for both H.265 (HEVC) and H.264 (AVC) for both tests.

- RAP distance

    RAP periods of 1 and 2 seconds are tested. In cases when the GOP structure and the frame rate combination is not convenient to generate exact RAP periods of 1 or 2 seconds, the RAP period is required to be adjusted to be as close as possible to the target RAP period. For example, for GOP size 8 and 30 fps, the RAP period is required to be of 4 GOPs for the target RAP period of 1 second, and 8 GOPs for the target RAP period of 2 seconds.

- Temporal scalability

    Temporal scalability (with 4 temporal sub-layers) is enabled for both H.265 (HEVC) and H.264 (AVC).

6.2.6   Test conditions for MMS

6.2.6.1   General testing settings

The general testing parameters as listed in Table 1 are recommended for evaluations of H.265 (HEVC) for video coding in MMS.

6.2.6.2   Test sequences

The test sequences as described in Table 2 are used.

6.2.6.3   Encoding settings

- Profile

    H.265 (HEVC) Main profile and H.264 (AVC) High profile are used.

- QP configuration

  Fixed QP configuration is used without rate control to avoid uncertainty due to different rate control algorithms. Cascaded QP setting (e.g. higher QP for P pictures than I pictures, higher QP for B pictures than P pictures, and higher QP for higher temporal level than lower temporal level in hierarchical coding structures) is allowed and similar QP cascading strategy is used for both H.265 (HEVC) and H.264 (AVC).

- GOP structures

  Hierarchical B coding structures with GOP size of 8 is used for both H.265 (HEVC) and H.264 (AVC).

- IRAP pictures

  Two types of tests are performed that uses open GOP or closed-GOP configuration for random access. For closed-GOP test, IRAP pictures are IDR pictures for both H.265 (HEVC) and H.264 (AVC). For open-GOP test, IRAP pictures are clean random access (CRA) pictures for H.265 (HEVC) and open-GOP intra pictures (indicated by recovery point SEI messages) for H.264 (AVC). The first picture is an IDR picture for both H.265 (HEVC) and H.264 (AVC) for both tests.

- RAP distance

  RAP periods of 1 and 2 seconds are tested. In cases when the GOP structure and the frame rate combination is not convenient to generate exact RAP periods of 1 or 2 seconds, the RAP period is adjusted to be as close as possible to the target RAP period. For example, for GOP size 8 and 30 fps, the RAP period is required to be of 4 GOPs for the target RAP period of 1 second, and 8 GOPs for the target RAP period of 2 seconds.

- Temporal scalability

  Temporal scalability (with 4 temporal sub-layers) is enabled for both H.265 (HEVC) and H.264 (AVC).

- Number of reference pictures

  The number of reference pictures in each reference picture list is set equal to 1.

- Motion vector search range

  The motion vector search range, in units of integer luma samples, is restricted to 32.

- Rate-distortion optimized quantization

  Rate-distortion optimized quantization is disabled for both H.265 (HEVC) and H.264 (AVC).

6.2.7   Test conditions for MTSI

6.2.7.1   General testing settings

The general testing parameters as listed in Table 1, excluding the RAP distance parameters, are recommended for evaluations of H.265 (HEVC) for video coding in MTSI.

6.2.7.2   Test sequences

The test sequences as described in Table 2 and Table 5 are used. Results based on additional test sequences are welcome but not required.

Table 5: Additional test sequences for tests for MTSI

Class	Sequence	Spatial resolution	Frame rate
Class VC-E	FourPeople	1280x720	60 fps
	Johnny	1280x720	60 fps
	KristenAndSara	1280x720	60 fps

6.2.7.3   Encoding settings

- Profile

H.265 (HEVC) Main profile, H.264 (AVC) Constrained Baseline profile (thus CAVLC is used while CABAC cannot be used, and 8x8 transform cannot be used), and H.264 (AVC) High profile are used.

- QP configuration

  Fixed QP configuration is used without rate control to avoid uncertainty due to different rate control algorithms. Cascaded QP setting (e.g. higher QP for P pictures than I pictures, higher QP for B pictures than P pictures, and higher QP for higher temporal level than lower temporal level in hierarchical coding structures) is allowed. Similar QP cascading strategy is used for both H.265 (HEVC) and H.264 (AVC).

- Number of reference pictures

  The number of reference pictures in the reference picture list is set equal to 2.

- GOP structures

  The IPPP coding structure, wherein the first picture in the bitstream is an IDR picture and the rest are P pictures, and the decoding order equals the output order, is used for both H.265 (HEVC) and H.264 (AVC). The prediction structure for the case with temporal scalability support is illustrated in Figure 3.
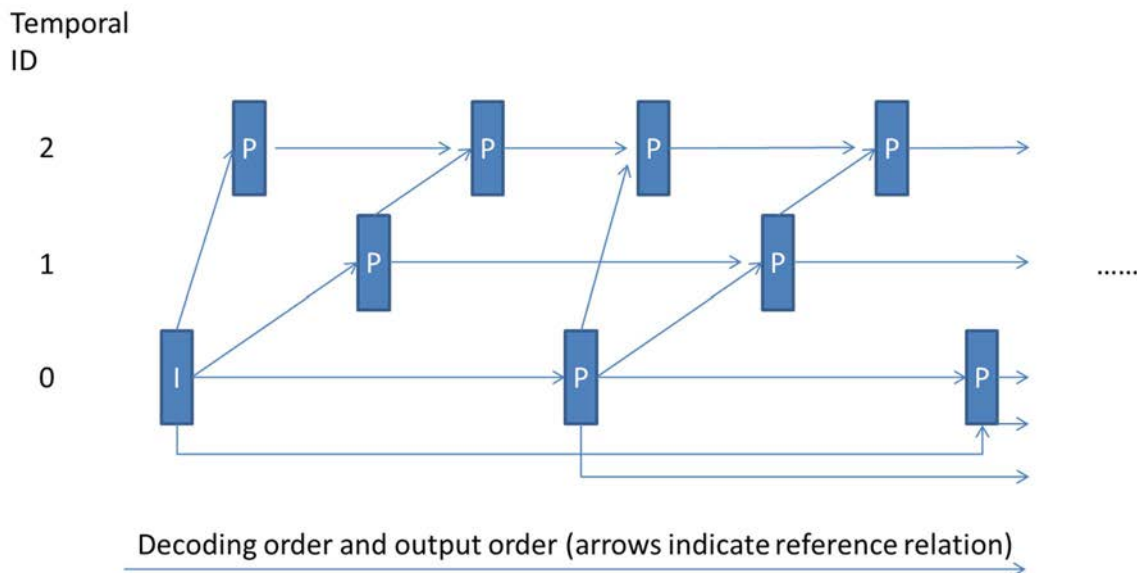
Figure 3: IPPP prediction structure with temporal layered prediction under low-delay conditions

- Temporal scalability.

  Both cases with temporal scalability not enabled and enabled with 3 temporal sub-layers are tested, for both H.265 (HEVC) and H.264 (AVC). See the "prediction structure" item below.

- Prediction structure.

  Different prediction structures are used to test different conditions.

  - Case 1: Temporal scalability is not supported. The previous two pictures in decoding order are always used for prediction.

  - Case 2: To test conditions with packet losses. Temporal scalability with 3 temporal sub-layers is supported. Each pi19icae picA occurring at (or immediately after) the end of one second intervals (the first interval begins from the first picture, which is an IDR picture), uses the picture that precedes the output ti19icaf picA by roughly 300 ms and that belongs to the same or a lower temporal sub-layer as the reference picture for prediction. For all other pictures, the two pictures preceding in decoding order that belong to the same or a lower temporal sub-layer are used for prediction.

- MTU size matching

  Multiple slices are allowed. The size of each slice in a picture is set to 1200 bytes, with the exception that the last slice in each picture is allowed to have a smaller size.

- Motion vector search range.

  The motion vector search range, in units of integer luma samples, is restricted to 32.

- Rate-distortion optimized quantization.

  Rate-distortion optimized quantization is disabled for both H.265 (HEVC) and H.264 (AVC).

6.3 Test cases for evaluation of H.265 (HEVC) for image coding

6.3.1 Codec software

For coding efficiency tests, HM version 10 is used for H.265 (HEVC) and the ImageMagick software is used for JPEG.

6.3.2 Test sequences

The first pictures of the JCT-VC test sequences as described in Table 2 are used. Results based on additional test pictures are welcome but not required.

6.3.3 Encoding settings

Still pictures are coded at three different quality levels with H.265 (HEVC) and JPEG. The quality levels are defined with PSNR and they correspond to:

High quality: 40 dB

Medium quality: 36 dB

Low quality: 32 dB

For JPEG, ImageMagick is configured to code pictures specified in the 3GPP services (as baseline DCT, non-differential, Huffman coding, as defined in Table B.1, s'mbol"SOF0' in 3GPP TS 26.273 [17]).

6.3.4 Evaluation metrics

For each picture and quality level, the file size of H.265 (HEVC) picture is compared with the corresponding JPEG picture and the file size saving H.265 (HEVC) brings is measured.

7 Test results for video coding

7.1 Introduction

This clause documents simulation results for evaluation of HEVC for video coding in 3GPP multimedia services according to the test conditions specified in subclause 6.2.

For evaluation of HEVC for video coding in 3GP-DASH, PSS, and MBMS, three sets of objective test results are summarized in subclauses 7.2 to 7.4. Detailed results can be found in the attached Excel sheets in S4-130708, S4-130790, and S4-130747, respectively. [Ed. (YK): In the finalized version to be published, consider including these Excel sheets as direct attachments of the present document] These objective test results were generated per the test conditions described above, except that for the third set of objective tests different test sequences were used.

Additionally, a set of subjective test results is reported for evaluation of HEVC for video coding in 3GP-DASH, PSS, and MBMS, as summarized in subclause 7.5.

Objective simulation results for evaluation of HEVC for video coding in MMS and MTSI are summarized in subclause 7.6. Detailed results for these simulations can be found in the attached Excel sheets in S4-131197 [Ed. (YK): In the finalized version to be published, consider including these Excel sheets as direct attachments of the present document].

7.2 Summaries of the first set of objective test results for 3GP-DASH, PSS, and MBMS

The first set of summaries were extracted from the full results by firstly selecting the H.264 (AVC) encodings with bitrates roughly matching to 2 Mbps, 1,5 Mbps, 1 Mbps, and 250 kbps for 1080p, 720p, 480p, and 240p respectively. Then the corresponding H.265 (HEVC) sequence with roughly the same objective quality as measured by PSNR was selected. The results were then averaged for different resolutions.

The summaries are provided in the tables below.

Table 6: RAP period = 1 second, closed GOP

	HEVC		AVC		
	Bitrate (kbit/s)	Y-PSNR (dB)	Bitrate (kbit/s)	Y-PSNR (dB)	HEVC Gain
Average (1080p)	1169.5	33.7	1960.1	33.6	40.3%
Average (720p)	1087.9	35.2	1693.6	35.3	35.8%
Average (480p)	639.9	34.8	970.9	35.0	34.1%
Average (240p)	202.2	33.2	292.3	33.6	30.8%

Table 7: RAP period = 2 second, closed GOP

	HEVC		AVC		
	Bitrate (kbit/s)	Y-PSNR (dB)	Bitrate (kbit/s)	Y-PSNR (dB)	HEVC Gain
Average (1080p)	1040.8	33.7	1724.9	33.4	39.7%
Average (720p)	971.6	35.2	1489.7	35.1	34.8%
Average (480p)	577.0	34.7	860.0	34.8	32.9%
Average (240p)	180.3	33.2	255.4	33.3	29.4%

Table 8: RAP period = 1 second, open GOP

	HEVC		AVC		
	Bitrate (kbit/s)	Y-PSNR (dB)	Bitrate (kbit/s)	Y-PSNR (dB)	HEVC Gain
Average (1080p)	1145.4	33.8	1921.9	33.6	40.4%
Average (720p)	1064.2	35.3	1658.8	35.3	35.8%
Average (480p)	626.3	34.9	951.1	35.1	34.1%
Average (240p)	197.7	33.4	285.9	33.7	30.9%

Table 9: RAP period = 2 second, open GOP

	HEVC		AVC		
	Bitrate (kbit/s)	Y-PSNR (dB)	Bitrate (kbit/s)	Y-PSNR (dB)	HEVC Gain
Average (1080p)	1029.0	33.7	1706.1	33.4	39.7%
Average (720p)	960.2	35.2	1472.6	35.1	34.8%
Average (480p)	570.2	34.8	850.4	34.8	33.0%
Average (240p)	178.0	33.3	252.2	33.4	29.4%

As can be seen from the above tables:

- H.265 (HEVC) achieves roughly similar PSNR using about 30-40% less bitrate compared to H.264 (AVC).

- The coding efficiency gains of H.265 (HEVC) are larger for higher resolutions (e.g. 720p and 1080p) compared to smaller resolutions (e.g. 240p and 480p).

- The coding efficiency gains of H.265 (HEVC) are consistent along different random access periods and also prediction structures (open GOP and closed GOP)

7.3 Summaries of the second set of objective test results for 3GP-DASH, PSS, and MBMS

In the second set of summaries, four sets of overlapping QP value ranges, as described below in Table 10, were used to compute the BD-rate values.

Table 10: QP values used for computing BD-rate values for different rate conditions

Bit rate	QP values used for BD-rate computation
High bit rate	19, 22, 25, 28
Medium bit rate	28, 31, 34, 37
Low bit rate	37, 40, 43, 46
Overall	19, 28, 37, 36

The summarized BD-rate results are presented in Table 11 to Table 14. The results for various prediction structures and RAP periods are presented in separate tables.

Table 11: BD-rate of H.265 (HEVC) compared to H.264 (AVC) for open GOP structure with 1 sec RAP period

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-34.2%	-36.0%	-34.5%	-42.2%	-31.8%	-30.3%	-53.9%	-59.6%	-61.2%	-43.6%	-39.9%	-38.9%
720p	-29.3%	-27.0%	-26.1%	-34.9%	-26.3%	-25.6%	-47.4%	-54.6%	-56.4%	-36.7%	-32.8%	-32.2%
480p	-28.3%	-28.3%	-27.0%	-31.9%	-26.6%	-24.2%	-42.0%	-51.7%	-53.8%	-33.5%	-32.0%	-31.4%
240p	-25.2%	-27.4%	-26.3%	-27.1%	-21.5%	-18.6%	-32.4%	-36.9%	-43.0%	-28.0%	-25.8%	-25.6%
Overall	**-29.0%**	**-29.8%**	**-28.4%**	**-33.0%**	**-25.9%**	**-23.6%**	**-42.1%**	**-48.7%**	**-51.9%**	**-34.4%**	**-31.8%**	**-31.1%**

Table 12: BD-rate of H.265 (HEVC) compared to H.264 (AVC) for open GOP structure with 2 sec RAP period

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-35.0%	-38.6%	-38.5%	-43.6%	-34.7%	-33.3%	-56.3%	-60.8%	-63.3%	-45.0%	-42.3%	-41.6%
720p	-30.4%	-29.4%	-28.7%	-35.7%	-27.7%	-26.7%	-48.8%	-55.3%	-57.6%	-37.7%	-34.5%	-33.9%
480p	-29.8%	-31.9%	-30.4%	-33.5%	-29.3%	-26.4%	-43.9%	-52.5%	-54.9%	-35.1%	-34.7%	-33.8%
240p	-26.8%	-31.3%	-29.5%	-28.7%	-23.6%	-21.2%	-34.3%	-38.1%	-42.5%	-29.6%	-28.8%	-28.0%
Overall	**-30.4%**	**-33.2%**	**-31.9%**	**-34.5%**	**-28.4%**	**-26.0%**	**-44.2%**	**-49.8%**	**-52.9%**	**-36.0%**	**-34.5%**	**-33.6%**

Table 13: BD-rate of H.265 (HEVC) compared to H.264 (AVC) for closed GOP structure with 1 sec RAP period

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-33.1%	-33.4%	-31.4%	-41.0%	-28.8%	-27.4%	-53.2%	-58.6%	-60.3%	-42.5%	-37.7%	-36.7%
720p	-28.1%	-24.4%	-23.7%	-33.5%	-22.9%	-22.4%	-46.7%	-53.2%	-55.6%	-35.6%	-30.4%	-30.2%
480p	-27.4%	-26.3%	-25.0%	-30.8%	-23.7%	-21.5%	-41.2%	-50.4%	-52.9%	-32.5%	-30.0%	-29.6%
240p	-24.3%	-25.4%	-24.3%	-25.9%	-18.5%	-15.6%	-31.5%	-34.6%	-41.3%	-27.0%	-23.5%	-23.4%
Overall	**-28.1%**	**-27.6%**	**-26.1%**	**-31.8%**	**-22.9%**	**-20.7%**	**-41.4%**	**-47.1%**	**-50.8%**	**-33.4%**	**-29.6%**	**-29.1%**

**Table 14: BD-rate of H.265 (HEVC) compared to H.264 (AVC)
for closed GOP structure with 2 sec RAP period**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-34.5%	-37.5%	-37.1%	-43.0%	-33.2%	-31.8%	-55.9%	-60.3%	-62.8%	-44.5%	-41.2%	-40.6%
720p	-29.9%	-28.3%	-27.5%	-35.0%	-26.0%	-25.1%	-48.4%	-54.6%	-57.2%	-37.2%	-33.3%	-32.9%
480p	-29.4%	-30.9%	-29.5%	-33.0%	-28.0%	-24.9%	-43.5%	-51.8%	-54.5%	-34.7%	-33.7%	-32.8%
240p	-26.4%	-30.3%	-28.6%	-28.1%	-22.1%	-19.6%	-33.8%	-37.1%	-41.4%	-29.1%	-27.7%	-27.1%
Overall	**-30.0%**	**-32.2%**	**-30.8%**	**-34.0%**	**-26.9%**	**-24.4%**	**-43.7%**	**-49.1%**	**-52.2%**	**-35.5%**	**-33.4%**	**-32.6%**

Figure 4 to Figure 7 show the plots of PSNR versus bit rate for a typical sequence (BasketballDrive) under open GOP structure with 2-sec RAP period for various picture resolutions (240p, 480p, 720p, and 1080p). The attached Excel file also provides means to plot the PSNR-versus-rate curves for all the test conditions and sequences.
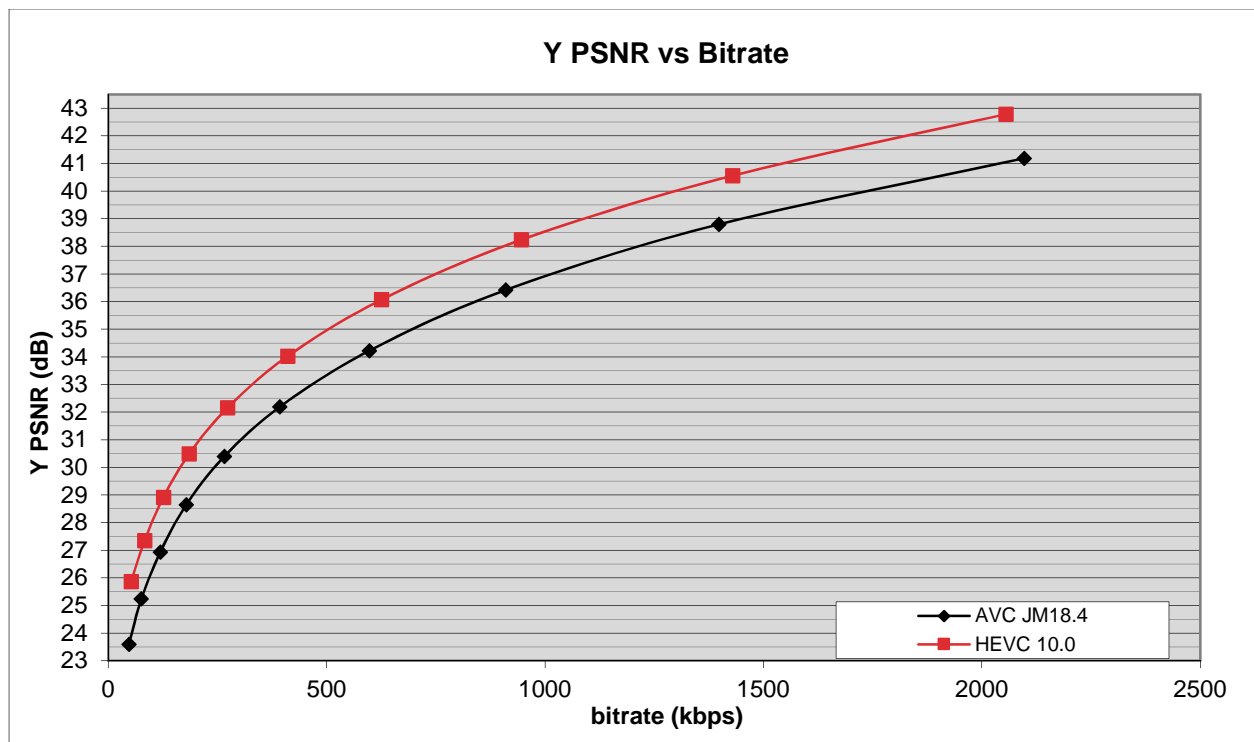
Figure 4: BasketballDrive 240p sequence under open GOP structure and 2 sec RAP period
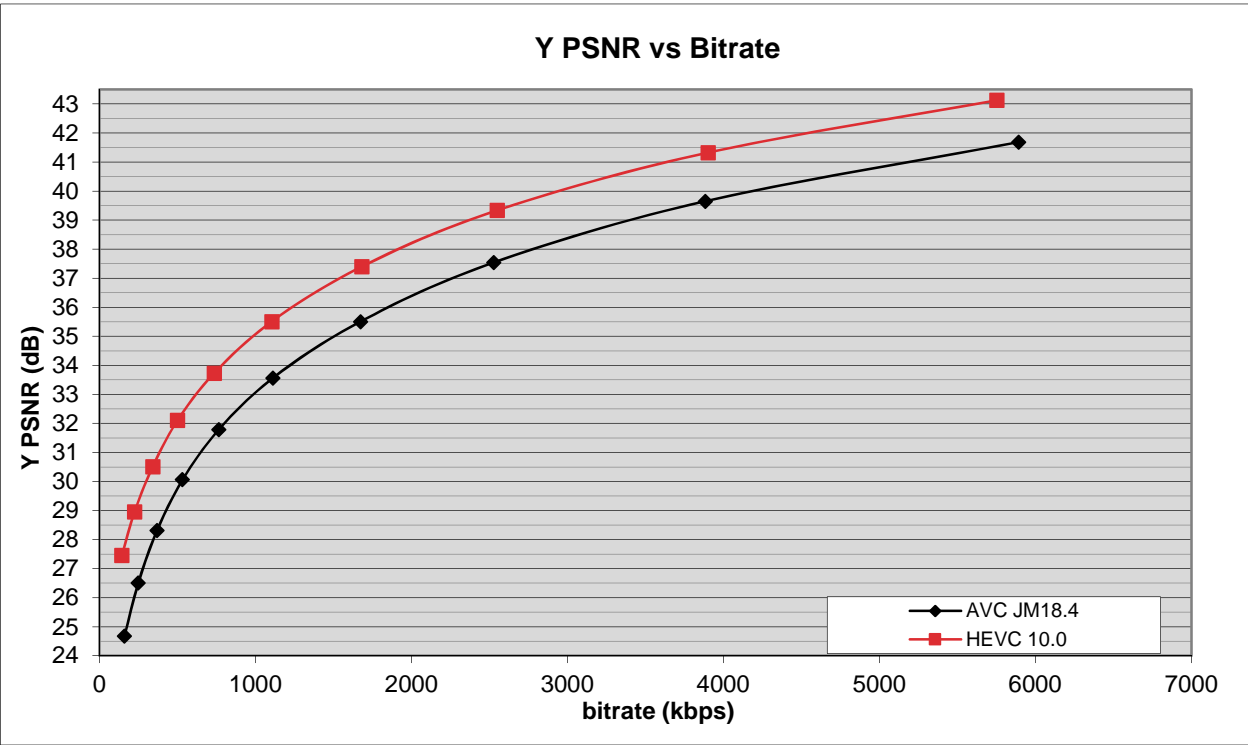
Figure 5: BasketballDrive 480p sequence under open GOP structure and 2 sec RAP period
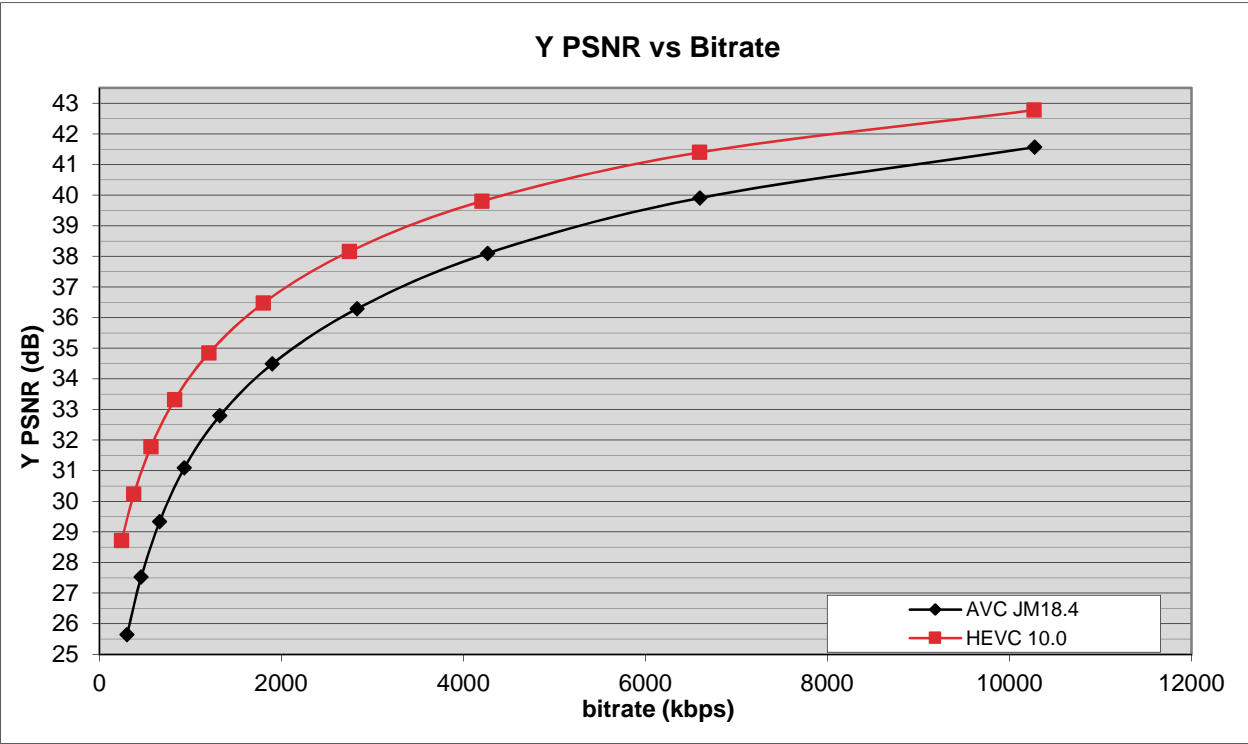

Figure 6: BasketballDrive 720p sequence under open GOP structure and 2 sec RAP period
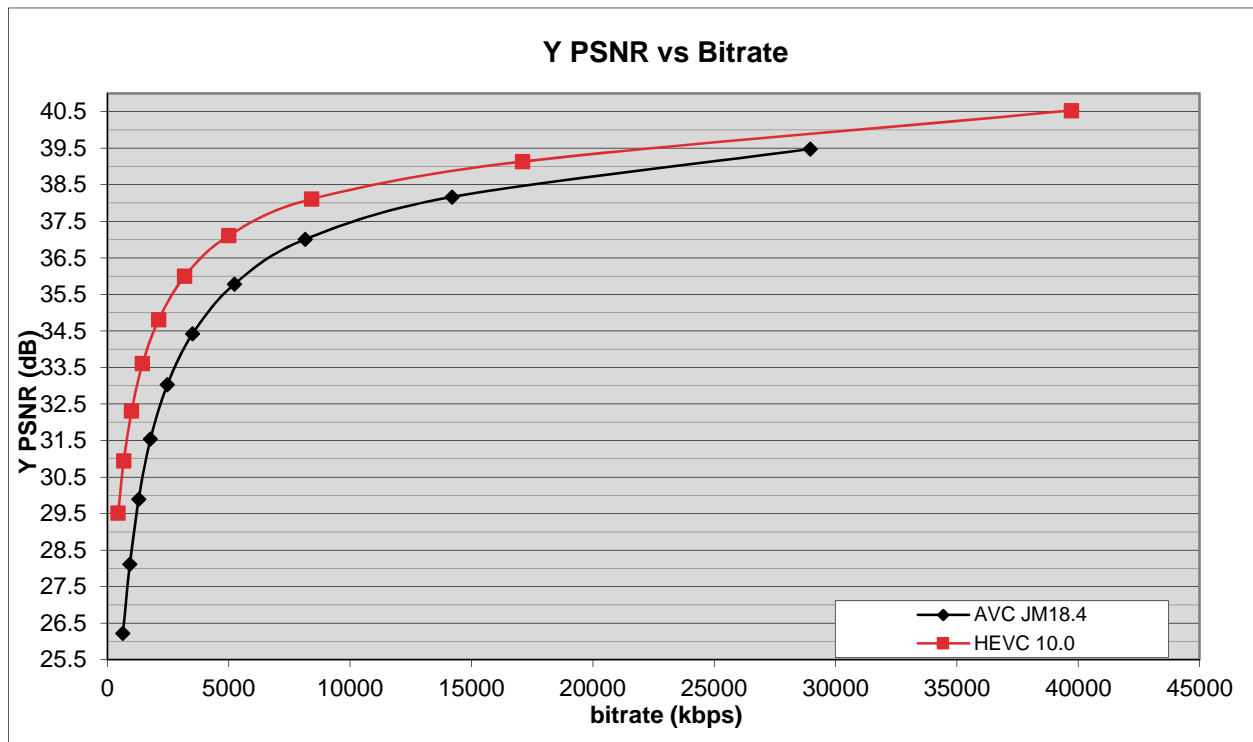
Figure 7: BasketballDrive 1080p sequence under open GOP structure and 2 sec RAP period

As can be seen from the above tables and figures, the average decrease in BD-rate of H.265 (HEVC) when compared to H.264 (AVC) is 30 – 40%. More specifically:

- The average decrease in BD-rate values for H.265 (HEVC) when compared to H.264 (AVC) –s 30 - 40% for different prediction (open and closed GOP) structures.

- The results are consistent across different RAP periods (1 sec and 2 sec).

- The performance gap is bigger for higher resolutions than lower spatial resolutions.

- Within each spatial resolution, the performance gap is bigger for lower bit rates than higher bit rates. For example, the gap at 1080p resolution was around 35% for higher bit rate range and 50% to 55% for lower bit rate range.

7.4   Summaries of the third set of objective test results for 3GP-DASH, PSS, and MBMS

7.4.1   Test setup

In the third set of summaries, five different test sequences than listed in the test conditions described above were used. Two of them came from the "The Big BuckBunny" animation mIe ((c) copyright 2008, Blender Foundation / www.bigbuckbunny.org) and were originally available in 1080p25, and the other three were provided with authorization by the European Broadcast Union (EBU), which were provided in 1080p50. A snapshot of each of these sequences is provided below.

Figure 8: Bunny sequences #1 and #2

Figure 9: Opening, ESC and IceDance source sequences from the EBU

From the sequences at 50fps, a version of 25 frames per second was obtained by temporally sub-sampling the original source, for which the sub-sampling process is as described in clause 6.2.2.1.1.

Table 15: List of source formats

Full HD	1080p50
	1080p25
HD	720p50
	720p25
SD	480p
Quarter SD	240p

Only the open-GOP configuration with 2s of RAP period was tested. For each sequence in each of the source formats listed in Table 15, a set of 10 quantization steps was generated using the following QP values for both H.264 (AVC) and H.265 (HEVC): 16, 19, 22, 25, 28, 31, 34, 37, 40 and 43.

7.4.2 Test summaries

Three sets of overlapping QP value ranges, as described in the table below, were used to compute the BD-rate values. Although the QP 16 and 43 configurations were generated and documented in the Excel sheet, they were not taken into

account due to the fact that they are not realistic in terms of service implementation (bitrate too high, out of level limits, or quality very degraded).

Table 16: QP values used for computing BD-rate values for different rate conditions

Bit rate	QP values used for BD-rate computation
High bit rate	19, 22, 25, 28
Medium bit rate	25, 28, 31, 34
Low bit rate	31, 34, 37, 40

The summarized BD-rate results are presented in the following table. There is no overall average gain presented due to the fact that the variation is too important and it was considered that 3GPP services should focus only on the performances from the medium and low bit-rate ranges.

Table 17: BD-rate results summary for the third set of objective tests

	High bit-rate			Medium bit-rate			Low bit-rate		
	Y	U	V	Y	U	V	Y	U	V
1080p50	-27,9%	-21,1%	-26,1%	-39,5%	-42,8%	-44,4%	-45,3%	-57,6%	-58,9%
1080p25	-25,4%	-21,0%	-22,5%	-35,2%	-33,6%	-33,9%	-41,8%	-47,4%	-47,2%
720p50	-30,5%	-29,6%	-31,5%	-34,8%	-39,3%	-40,3%	-39,9%	-53,0%	-54,7%
720p25	-23,3%	-21,5%	-21,5%	-27,3%	-26,0%	-26,3%	-33,6%	-38,3%	-39,2%
480p	-25,0%	-24,9%	-25,1%	-28,4%	-30,2%	-30,5%	-33,5%	-42,3%	-42,7%
240p	-21,3%	-23,2%	-23,1%	-23,7%	-27,3%	-27,3%	-27,4%	-38,1%	-37,5%
Overall	-25,6%	-23,9%	-25,1%	-31,3%	-33,0%	-33,4%	-36,7%	-45,8%	-46,1%

When comparing these results with the ones in the first and second sets of objective tests (based on different test sequences), it can be noted that the H.265 (HEVC) gain over H.264 (AVC) is lower by 5% in average. Nevertheless, for the low-to-medium bit-rate ranges, H.265 (HEVC) significantly outperforms H.264 (AVC) for this set of tests for an average decrease in BD-rate in the range–of 27,4 - 45,3%.

7.5 Subjective test results for 3GP-DASH, PSS, and MBMS

The video quality when displayed on a smartphone and a tablet was evaluated by naïve test subjects. No formal test methods on how to do tests on mobile terminal exist, but the test followed Recommendation ITU-T P.910 [14] as close as possible.

7.5.1 Test setup

7.5.1.1 Test material

The original source sequences used here are:

- Kimono1 1920x1080@24fps

- Park Scene 1920x1080@24fps

- Cactus 1920x1080@50fps

- BQTerrace 1920x1080@60fps

- BasketBallDrive 1920x1080@50fps

All sequences are 10 seconds in length.

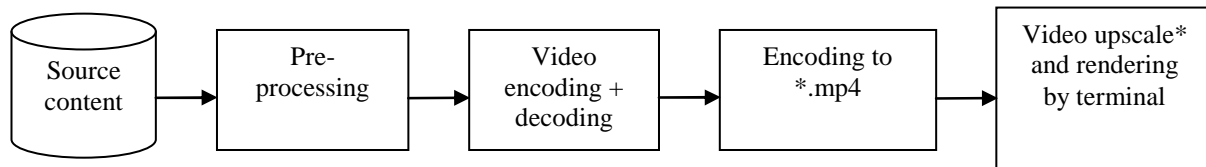The original source sequences were processed according to Figure 10.

Figure 10: Complete processing chain. * Video upscale to full-screen in terminal (no cropping)

The processing steps were:

- Pre-processing: Resizing to 1280x720 and 832x480. The 832x480 files were also cropped.

- Video encoding & decoding.
  All encodings were performed with open GOP, an Intra picture interval of one second, hierarchical B pictures with a length of 8, with an increase of QP with 1 for each hierarchical level and non-reference pictures at the highest level. Temporal layers were not used. QPs were selected to span a quality range from low to high subjective quality. The QPs were set so that each H.265 (HEVC) bit stream has a corresponding H.264 (AVC) bit stream with a slightly higher bit rate. The QP was kept static during each encoding except for QP offsets depending on the GOP position of each picture.

  - H.264 (AVC): High Profile @ original framerate
    The JM 18.4 encoder was used using HM-like configuration for random access. The encodings were based on the "encoder_JM_RAB_HE.cfg" configuration file which is part of the JM18.4 software package.

  - H.265 (HEVC): Main profile @ original framerate
    The HM-10.0 encoder was used using random access configuration. The encodings were based on the "encoder_randomaccess_main.cfg" configuration file which is part of HM-10.0 software package.

- To be able to display the H.265 (HEVC) encoded clips, transcoding of each reconstructed H.265 (HEVC) video to H.264 (AVC) was applied at around 10 Mbps.

- Video upscale and rendering by terminals. The upscale should keep the Pixel Aspect Ratio (PAR) format i.e. the 16/9 format of the video.

7.5.1.2   Display by terminal

Since terminals normally upscale videos to full screen this was used also in this test. The upscale is done by the respective terminal. The files were played out on smartphones and tablets having a screen resolution of 1920x1080. The format ratio of the video was not affected during play-out on the screen.

7.5.1.3   Test conditions

The test conditions contain variations of following parameters:

- Content

- Encoding bit rate

- Picture formats:

  a) 1920x1080

  b) 1280x720

  c) 832x480 (smartphone only)

- The frame rates were 24, 50 and 60 Hz.

All videos were displayed in full screen, up-scaled by the device (to maximum possible picture size for respective screen). Reasons for doing this upscale in the device are:

- The quality normally decreases with upscaling e.g. in the terminal at display in full screen mode, and to cover this effect of potentially introduced artifacts the videos were displayed in full-screen mode in the test.

- When watching longer clips large/full-screen picture format might be more common than watching in native/small format.

- Native formats is probably used when several windows are open and the person also does something else (looking for other clips, edit in document, etc.), however then the quality might not be in focus and use of full screen is thus more applicable for a quality assessment.

- The test is easier to perform if all clips have the same format. Clips of different formats are normally not tested in the same session.

7.5.1.4 Subjective test procedure

The test procedure followed Recommendation ITU-T P.910 [14] as closely as possible. The evaluation was done according to the Absolute Category Rating method (ACR). The test subjects performed evaluation of both the smartphone and the tablet.

7.5.1.5 Test methodology

The tests were performed according to the Absolute Category Rating (ACR) method [14]. Figure 11 illustrates the voting procedure; each clip is shown only once to each viewer and a grey background is shown as the viewer rates the clip. To avoid bias due to clip order the order is randomized for each viewer.
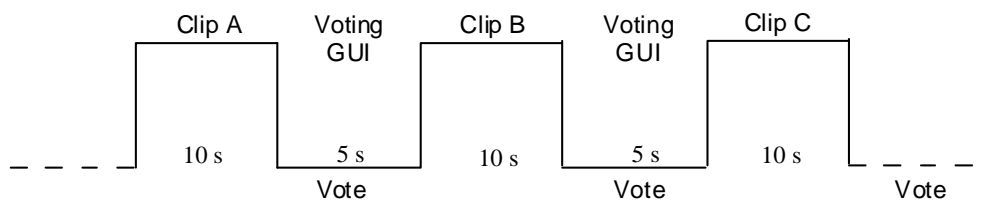

Figure 11: Voting procedure

A continuous 5-grade scale as defined in [14], Annex B was used for the voting. The scale had labels in the native language (Swedish) with the following translations; mycket dålig (bad), dålig29ica29tablecceptabel (fair), bra (good), utmärkt (excellent).
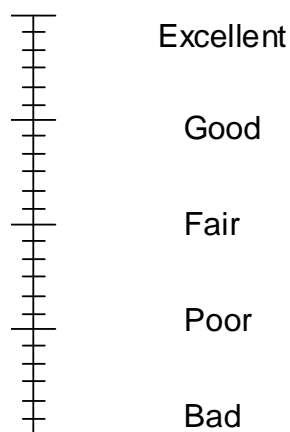

Figure 12: The continuous 5-grade Video Quality scale.

7.5.1.6 Test design

The test was executed on two smartphones and two tablets, tested in separate rooms. The test subjects performed the test on one device type. Four test subjects performed the test in parallel. Test design:

- Introduction: 15 min

- Pretest (10 sequences): 3 min

- Test session Smartphone (120 sequences): 2×16 minutes

- Test session Tablet (70 sequences): ~20 min

- Visual test: 5 min

Total test time: ~30 and 50 minutes respectively.

The test subjects were divided in 13 groups of four persons each, each group having a unique play out orders (14 in total). The test was executed in 3 working days.

7.5.1.7 Test environment

The test was performed in four small test rooms at the multimedia lab at Ericsson Research. Four test persons at a time performed the test in different rooms, using a Smartphone and tablet respectively. The test subjects distance to small screens are recommended to be $6\text{-}10 \times H$ (H = the screen height) [14] and [15] (normal reading distance is 25-30 cm).

- The smartphone screen size is 4,7-5,0", ca 3x5 cm. The test subjects distance to the screens is then ~6-8 x H (fullscreen) respectively.

- The tablet screen size is 10,1", ca 14x22 cm. The test subjects distance to the screen is then ~3-4 x H (fullscreen) respectively.

Room illumination (see note): ~20 Lux measured at terminal position and test subject face position.

   NOTE: This value indicates a setting allowing maximum detectability of distortions, for some applications higher values are allowed or they are determined by the application.

The screen luminance was adjusted to be as equal as possible, ~200 cd/m2. The luminance is measured when a white test signal is played.

Room noise: ≤ 30 dBA. The level is not defined in [14] but same level as for Recommendation ITU-T P.800 [24] MOS tests was strived for to achieve a quiet environment. Any Hoth noise was not activated.

7.5.1.8 MOS test tool

An in-house MOS Test tool application was used handling both video playout and voting on the same device.

The scoring time between play out of two files was of six seconds.

7.5.1.9 Test devices

Two smartphones (Sony Xperia Z™, HTC One™) and two tablets (ASUS Transformer Pad Infinity TF700™, Google Nexus 10™) were used during the test.

7.5.1.10 Test subjects

28 non-expert viewers employed at Ericsson performed the test. A non-expert viewer is here defined as a person not having good knowledge about video coding and video coding artifacts. The test subjects were compensated for their effort. Test instructions are available on request.

A near-viewing acuity test was performed. Noticeable is that two test subjects with "Not OK" performance had very low correlations, ~0,5 during the tablet test. Post screening of the results took place and the scores from the two test subjects with low correlation in the tablet test were removed.

7.5.2 Subjective test results

7.5.2.1 Smartphone results

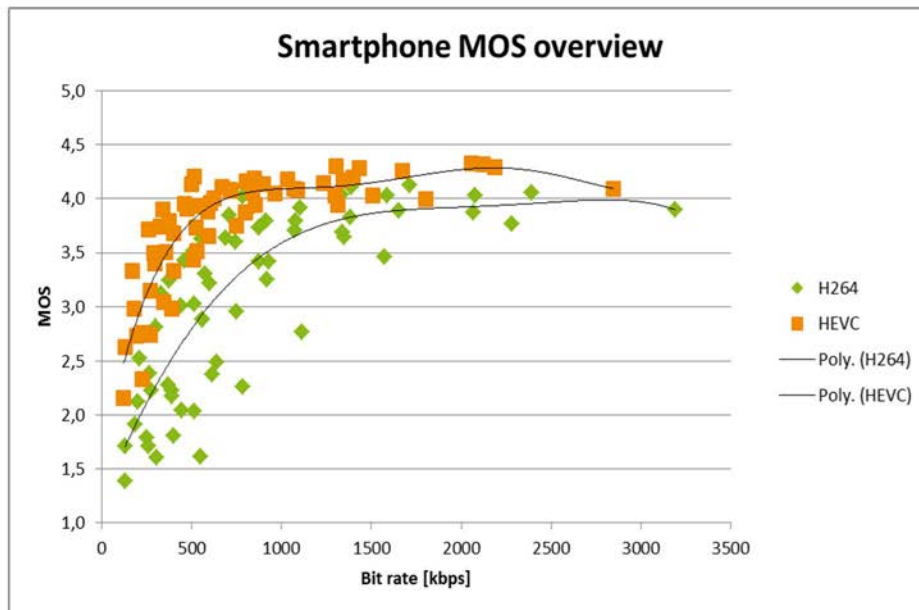All smartphone MOS (per condition) are displayed in Figure 13.

Figure 13: Smartphone MOS. Trend lines are in^{cl}uded (5th order polynomial)

The average 95% confidence interval is 0,31, i.e. less than the average MOS difference.

As expected, the MOS are clearly higher for H.265 (HEVC) than for H264. The gain in MOS for using H.265 (HEVC) is larger for lower bit rates than for higher bit rates. Approximate figures using the trend lines result in gains of ~1 MOS for 500 kbps, and ~0,5 MOS for 1000 kbps.

According to the analysis, the ranking of the conditions should not be affected by the different smartphones.
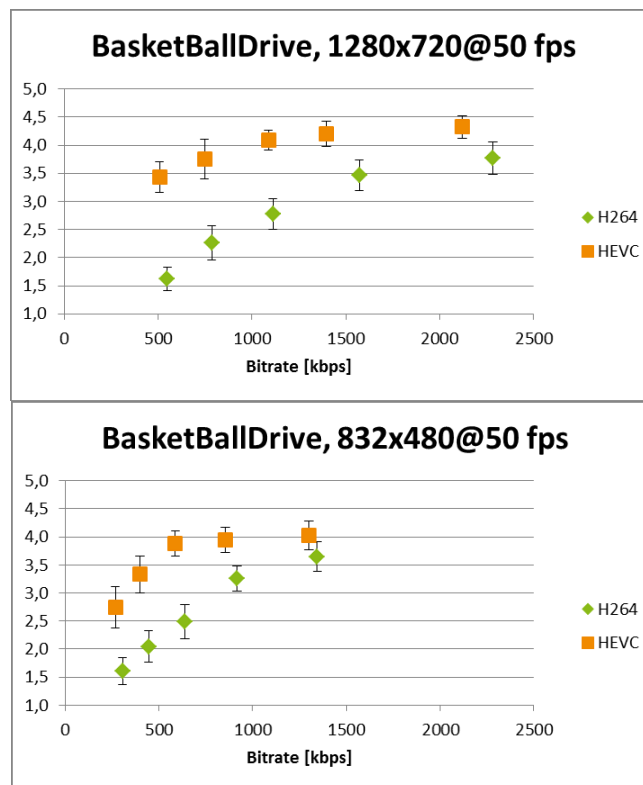
Figure 14: Quality vs. bit rate for BasketBallDrive
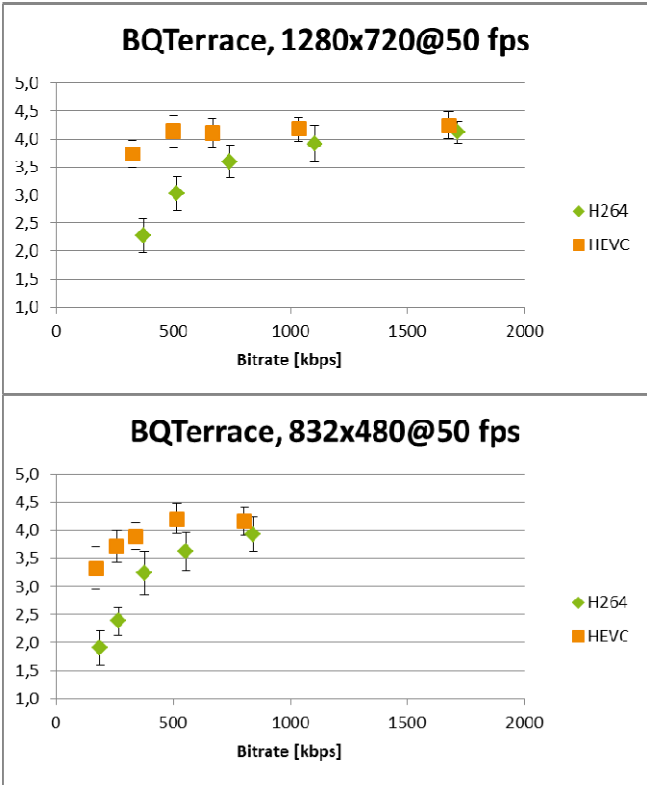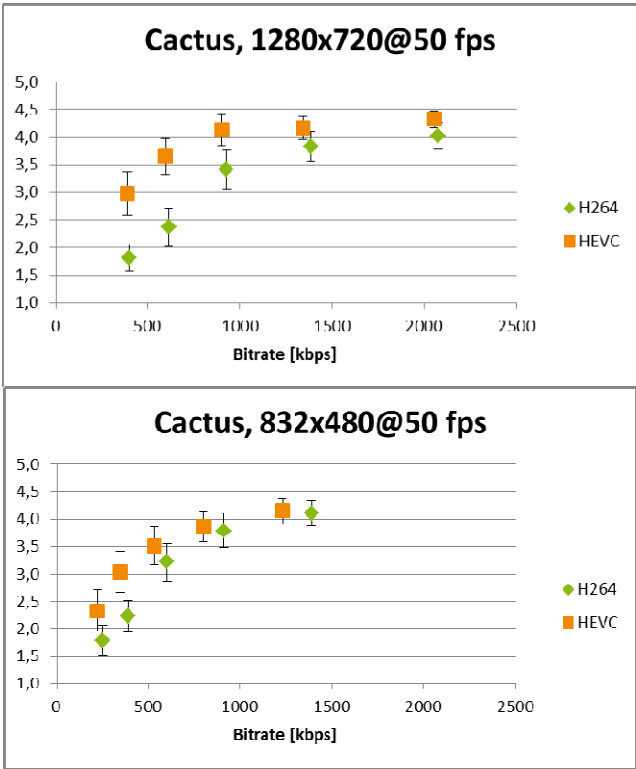
Figure 15: Quality vs. bit rate for BQTerrace

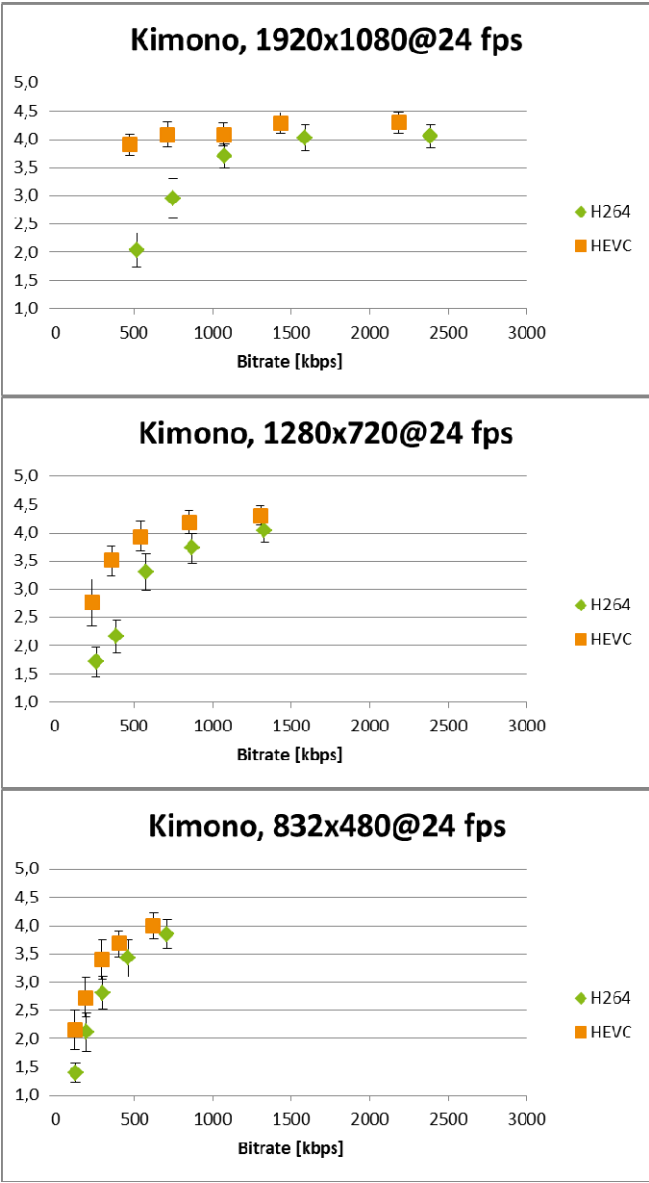
Figure 16: Quality vs. bit rate for Cactus

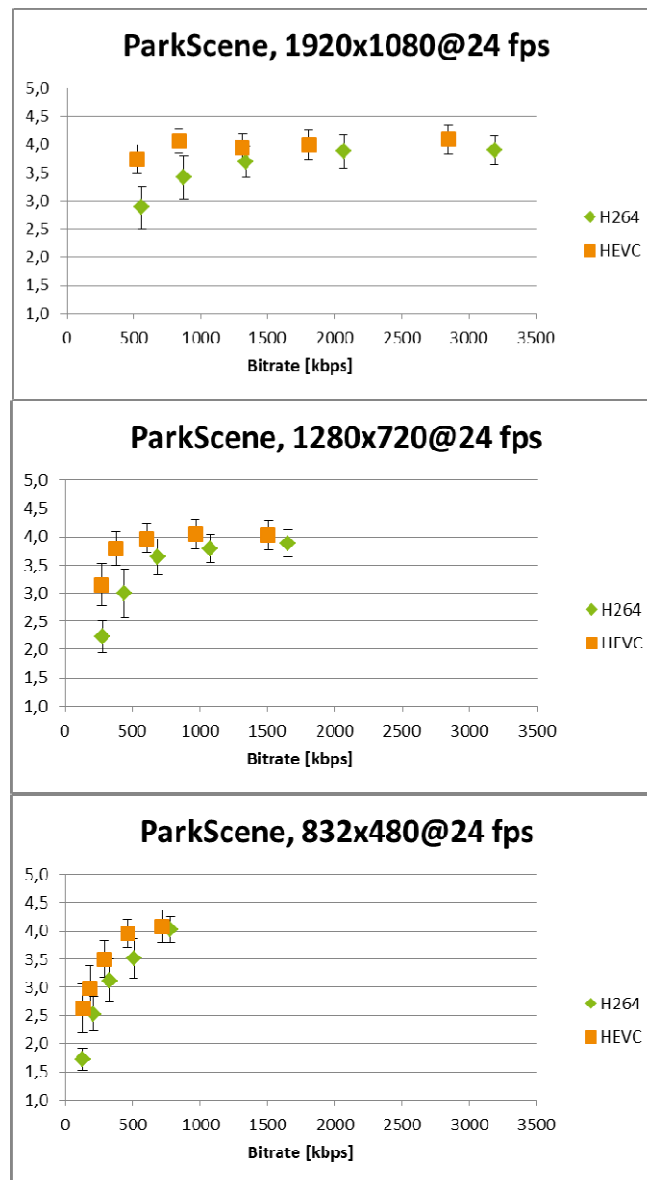**Figure 17: Quality vs. bit rate for Kimono**

Figure 18: Quality vs. bit rate for ParkScene

An overview of the bit rates required to achieve MOS = 3.5 ("good quality") for some content types is displayed in the following table for smartphones:

Table 18: Minimum bit rates [kbps] to achieve MOS = 3,5 ("Good Quality") for smartphones, displayed at full-screen format (1920x1080)

Resolution	HEVC		H264	
	Low motion	High motion	Low motion	High motion
1920x1080	< 500	-	1000	-
1280x720	~300	600	600	900
832x480	290	510	500	1200

The bit rates to achieve MOS=3,5 using H.265 (HEVC) is ~50% of the bit rate using H264 for 1280x720 and 832x480, and potentially less than 50% for 1920x1080.

Table 19: Relationship between H.265 (HEVC) and H264 bit rates to achieve MOS = 3.5 for smartphones

Content	HEVC/H264 bit rate		
	1920x1080	1280x720	832x480
BasketBallDrive		0,33	0,42
BQTerrace		0,43	0,40
Cactus		0,50	0,40
Kimono	<0,50	0,50	0,60
ParkScene	<0,50	0,53	0,57
Average		0,46	0,48

7.5.2.2   Tablet results

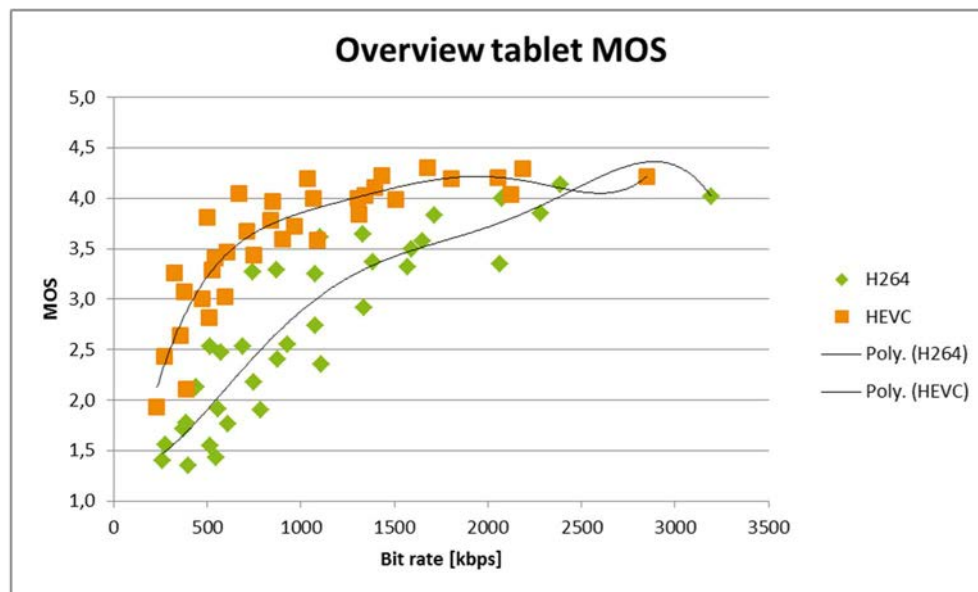All tablet MOS (per condition) are displayed in Figure 19.

Figure 19: Tablet MOS: Trend lines are included (5th order polynomial)

The average 95% confidence interval is 0,26 (0,31 for smartphones).

As expected, the MOS are clearly higher for H.265 (HEVC) than for H264. The gain in MOS for using H.265 (HEVC) is larger for lower bit rates than for higher. Approximate figures using the trend lines result in gains of ~1 MOS for 750 kbps, and ~0,5 MOS for 1500 kbps.

According to the analysis, the ranking of the conditions should not be affected by the different tablets.
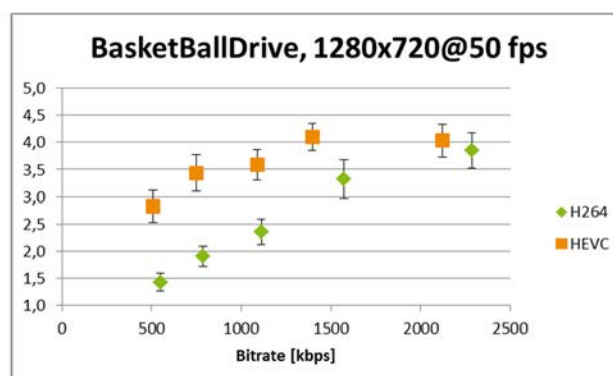

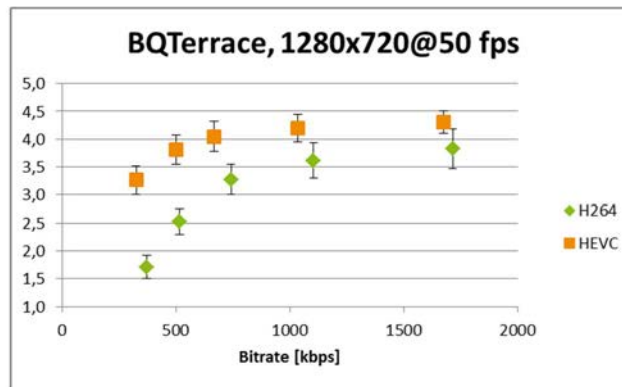**Figure 20: Quality vs. bit rate for BasketBallDrive**
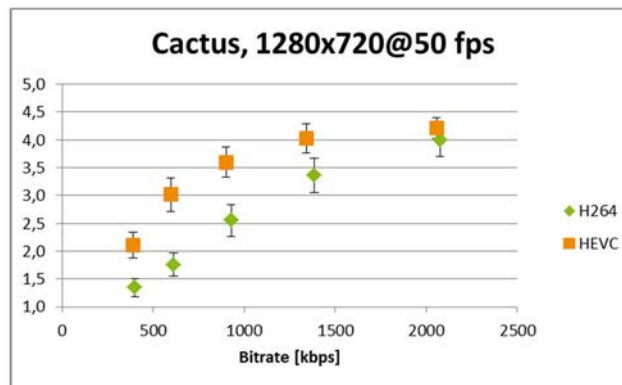
Figure 21: Quality vs. bit rate for BQTerrace

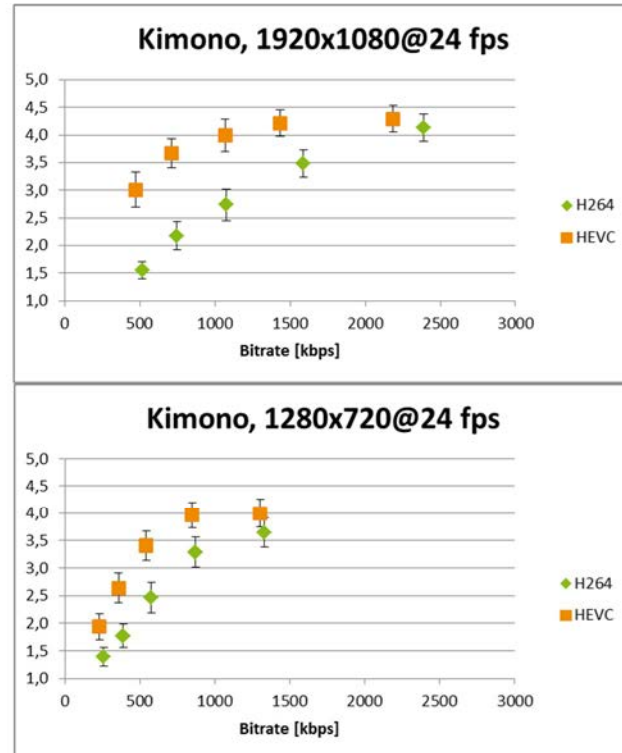
Figure 22: Quality vs. bit rate for Cactus

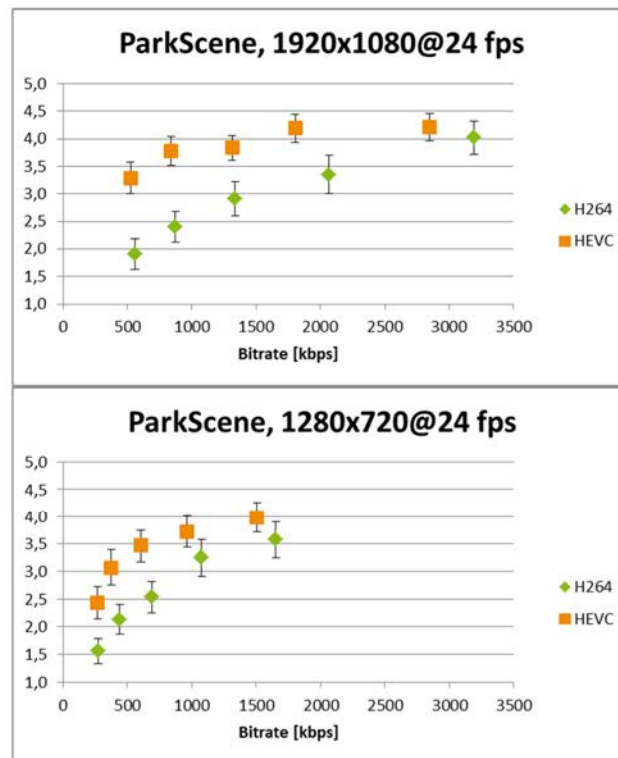**Figure 23: Quality vs. bit rate for Kimono**

Figure 24: Quality vs. bit rate for ParkScene

An overview of the bit rates required to achieve MOS=3,5 for some content types is displayed in the following table for tablets.

**Table 20: Minimum bit rate [kbps] to achieve MOS=3.5 ("Good Quality")
for tablets, when displayed at full-screen format (1920x1080).**

Resolution	HEVC		H264	
	Low motion	High motion	Low motion	High motion
1920x1080	650	-	1600	-
1280x720	550	900	1100	1800

The bit rates to achieve MOS=3,5 on tablets using H.265 (HEVC) is ~50% of the bit rate using H264 for 1280x720, but only ~30-40% for 1920x1080.

Table 21: Relationship between H.265 (HEVC) and H264 bit rates to achieve MOS = 3,5 for tablets

Content	HEVC/H264 bit rate	
	1920x1080	1280x720
BasketBallDrive		0,50
BQTerrace		0,56
Cactus		0,53
Kimono	0,41	0,50
ParkScene	0,28	0,43
Average	0,34	0,50

7.5.3 Summary of the subjective tests

The subjective tests show that the bit rates to achieve a certain MOS are clearly lower for H.265 (HEVC) than for H264. This is valid for all content, format, and devices in the test.

The bit rates to achieve MOS=3.5 using H.265 (HEVC) is *about 50%* of the bit rate using H264, but differ for different content types, resolutions, and device types. The gain is larger at lower bit rates than at higher rates.

7.6 Summaries of subjective test results for MTSI

7.6.1 Introduction

The video quality when displayed on a smartphone and a tablet was evaluated by naïve test subjects. No formal test methods on how to do tests on mobile terminal exist, but the test followed Recommendation ITU-T P.910 [14] as close as possible. The main task was to evaluate the video quality of H.265 (HEVC) vs. H.264 (AVC) on mobile devices. Other tasks considered were:

- The quality at different bit rates

- The quality of different format (854x480 and 1280x720 were included)

- The quality on different device types (smartphone and tablets)

7.6.2 Test setup

7.6.2.1 Test material

The source content for this test was selected to be typical examples of real-time communication content. The original source sequences used here are:

- C2hoffice (1920x1080@25fps)

- Johnny (1280x720@60fps)

- FourPeople (1280x720@60fps)

- Konf1_Jacob (1920x1080@50fps)

- Konf3_Beatriz (1920x1080@50fps)

- Konferans3Sofa (1920x1080@25fps)

10 seconds of the original source sequences are selected and used as source files in this test. "Johnny" and "FourPeople" are standardization sequences, while the other four are Ericsson sequences.

The processing of the content was selected for a fair comparison of performances regarding average bit rate, and to have the same quality within the whole clip.
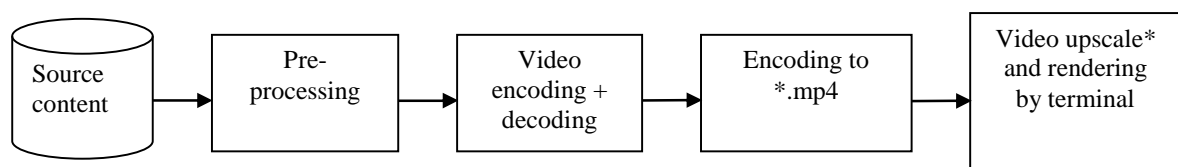

Figure 25: Complete processing chain. Video was upscaled to full-screen by the respective terminal (no cropping)

Examples of processing steps are:

- Pre-processing: Resizing to 1280x720 and 854x480 and frame rate conversion to 25 fps

  - Scaling of resolution and frame rate was done using ffmpeg. Command line as follows:

    - ff–pe–.exe -y -vcodec –awvideo -pix_fmt–yuv420p -r <original frame rate> -s <originaI size> -i <input file.yuv> -vcodec–rawvideo -pix_fm– yuv420p -vf–ames –50 -r 25 -s <1280x720 alt. 854x480> <output file.yuv>

NOTE: No other pre-processing than resizing and frame rate conversion took place.

Video encoding & decoding. The ~10 sec contents are mirrored to 20 sec before encoding to avoid initial quality variations in the test clip, where the last ~10 sec is used as test clip. The bit rates were calculated by summing the bits spent for coding the last 10 seconds of each clip:

- H.264 (AVC) (x264): Constrained Baseline Profile level 1.2

- H.265 (HEVC) (HM): Main profile

- Encoding to <H264>.mp4: H.264-encoding at ~20 Mbps in *.mp4-for play-out on mobile device.

- Video upscale and rendering by terminal. The upscale should keep the Pixel Aspect Ratio (PAR) format i.e. the 16/9 format of the video.

7.6.2.2 Video codecs

H.264 (AVC) encodings were performed with x264. H.265 (HEVC) encodings were performed with HM. Both using low-delay settings without picture reordering which also were real-time settings. The x264 rate-control was turned on. HM does not have a rate-control so those streams were encoded using fixed QPs.

NOTE: Strictly speaking it is not fair to compare a real-time H.264 (AVC) encoder (x264) and a several times slower H.265(HEVC) encoder (HM). The purpose of this is to evaluate the potential improved coding efficiency of using H.265(HEVC) in 3GPP MTSI. It is reasonable to assume that HW implementation will eventually be close to HM performances in the future.

7.6.2.3 Display by terminal

Since terminals normally upscale videos to full screen this was used also in this test. The upscale is performed by the respective terminal.

The files were played out on smartphones and tablets having a screen resolution of at least 1920x1080. The video aspect ratio (16:9) was not affected during play-out on the screen.

7.6.2.4 Test conditions

The parameters varied in the test were:

- Codec

- Content

- Bit rate

- Resolution

Two different types of devices were used, smartphones and tablets, and two different brands per device type. The screens were capable to display video in Full HD (1920x1080) resolution.

The same test files were used on the two different device types.

Five bit rates were selected for each video resolution to span the quality range from "Excellent" to "Bad".

The frame rates were 25 fps only.

The average encoding bit rate for a file is used in this document if nothing else is stated.

7.6.2.5 Subjective test procedure

Same procedure as in clause 7.5.1.4.

7.6.2.6 Test methodology

Same methodology as in clause 7.5.1.5.

7.6.2.7 Test environment

Same environment as in clause 7.5.1.7.

7.6.2.8 MOS test tool

Same tool as in clause 7.5.1.8.

7.6.2.9 Test devices

The test was performed using one Sony Xperia Z™ and one HTC One™ smartphone, and one ASUS TF701™ and a Google Nexus 10™ tablet. All devices have a screen resolution of at least 1920x1080 i.e. FullHD.

7.6.2.10 Test subjects

38 naïve viewers employed at Ericsson performed the test, 19 on each device type. A naïve viewer is here defined as a person not having good knowledge or experience about video coding and video coding artifacts. The test subjects were compensated for their effort.

No test subject results were post screened.

7.6.3 Subjective test results

7.6.3.1 Smartphone results

All smartphone MOS (per condition) are displayed in Figure 26.
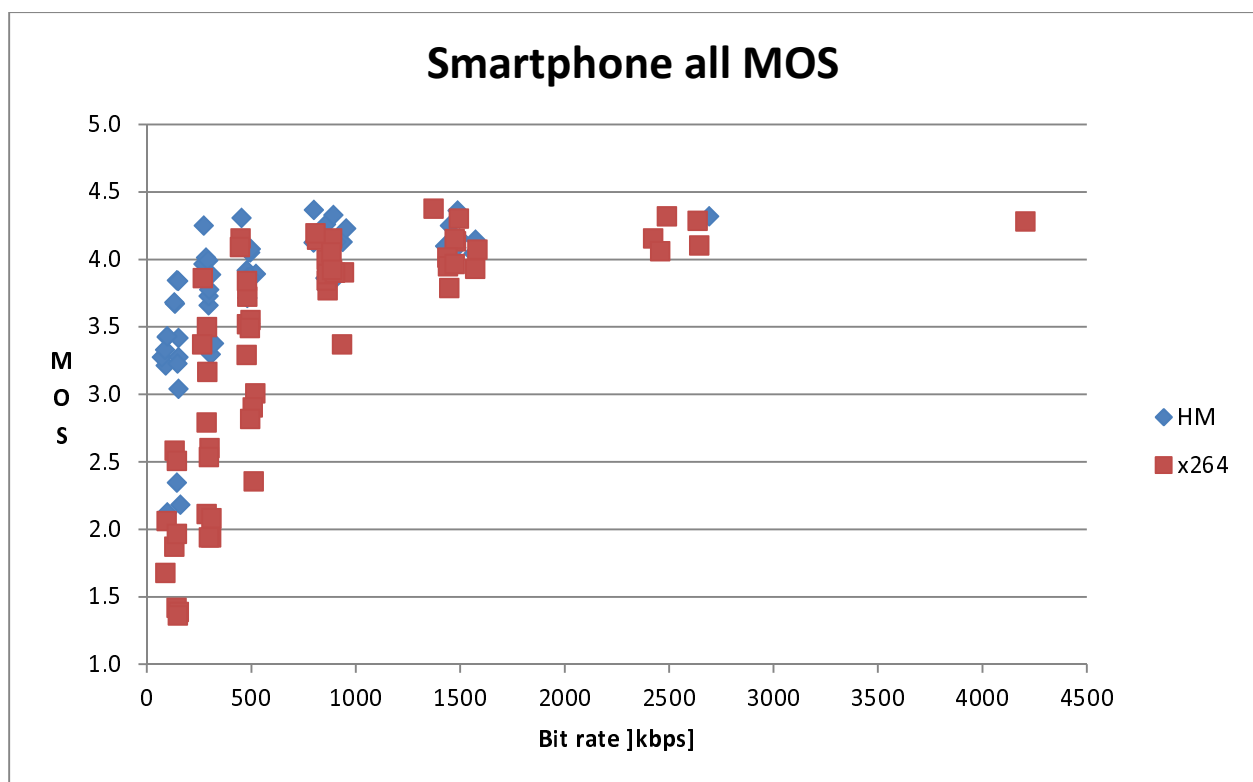

Figure 26: Smartphone all MOS

Figure 27: Quality vs. bit rate for c2hoffice

Figure 28: Quality vs. bit rate for FourPeople

Figure 29: Quality vs. bit rate for Johnny

Figure 30: Quality vs. bit rate for Jacob

Figure 31: Quality vs. bit rate for Beatriz

Figure 32: Quality vs. bit rate for Sofa

The bit rates required to achieve MOS = 3.5 (in between "Fair" and "Good") are estimated (estimated using a log-log estimation) and corresponding average bit rate gains are summarized in Table 22:

Table 22: Minimum bit rate [kbps] to achieve MOS = 3.5 for smartphones.

Content	Format	Estimated bit rate (kbps) to achieve MOS 3.5	
		HM	x264
Conf2_Home_office	854x480	137	330
FourPeople	854x480	444	693
Johnny	854x480	113	413
Conf1_Jacob	854x480	237	675
Conf3_Beatriz	854x480	185	691
Conf3_Sofa	854x480	554	1062
Conf2_Home_office	1280x720	107	465
FourPeople	1280x720	501	912
Johnny	1280x720	93	577
Conf1_Jacob	1280x720	212	959
Conf3_Beatriz	1280x720	299	1048
Conf3_Sofa	1280x720	430	1446
Average	**854x480**	**278**	**644**
Average	**1280x720**	**274**	**901**

7.6.3.2   Tablet results

All tablet MOS (per condition) are displayed in Figure 33.
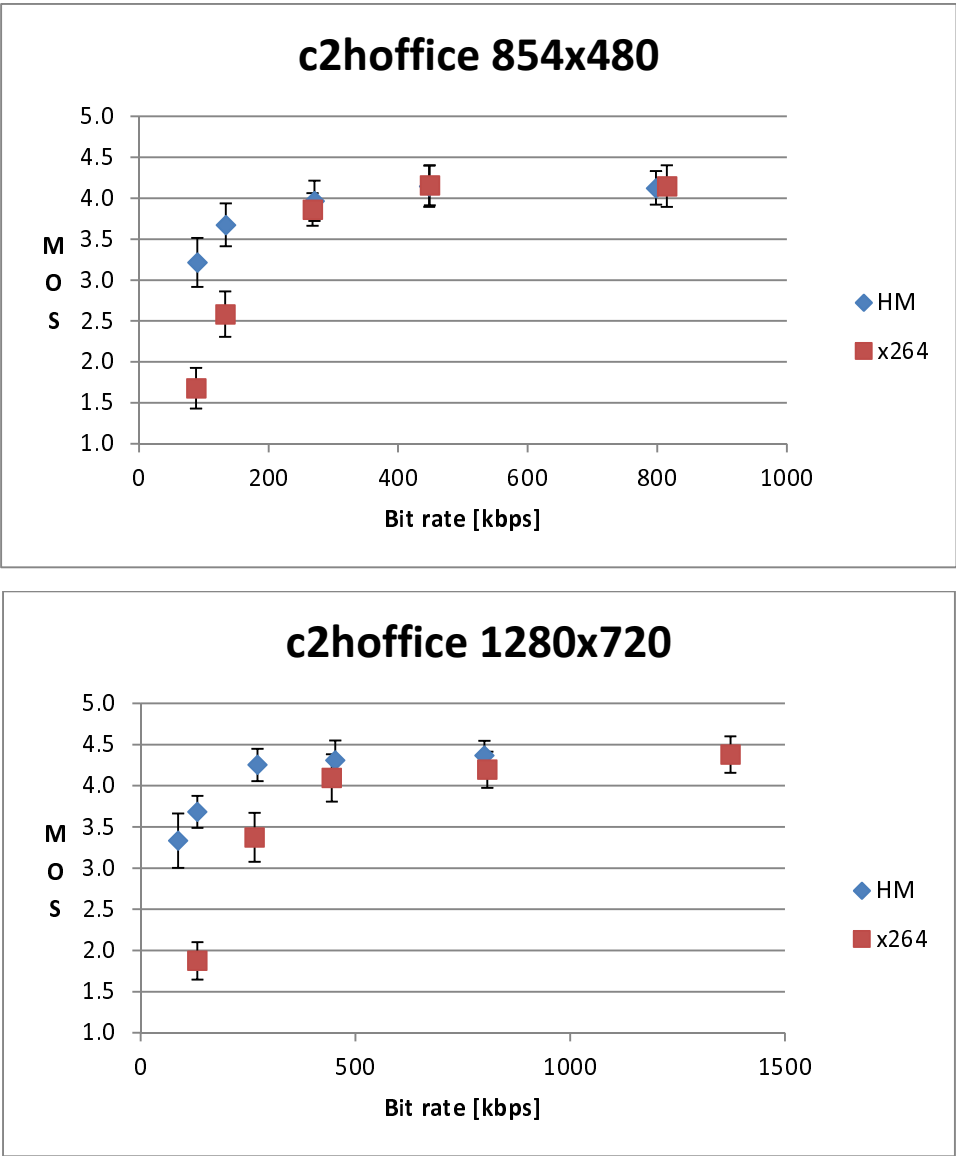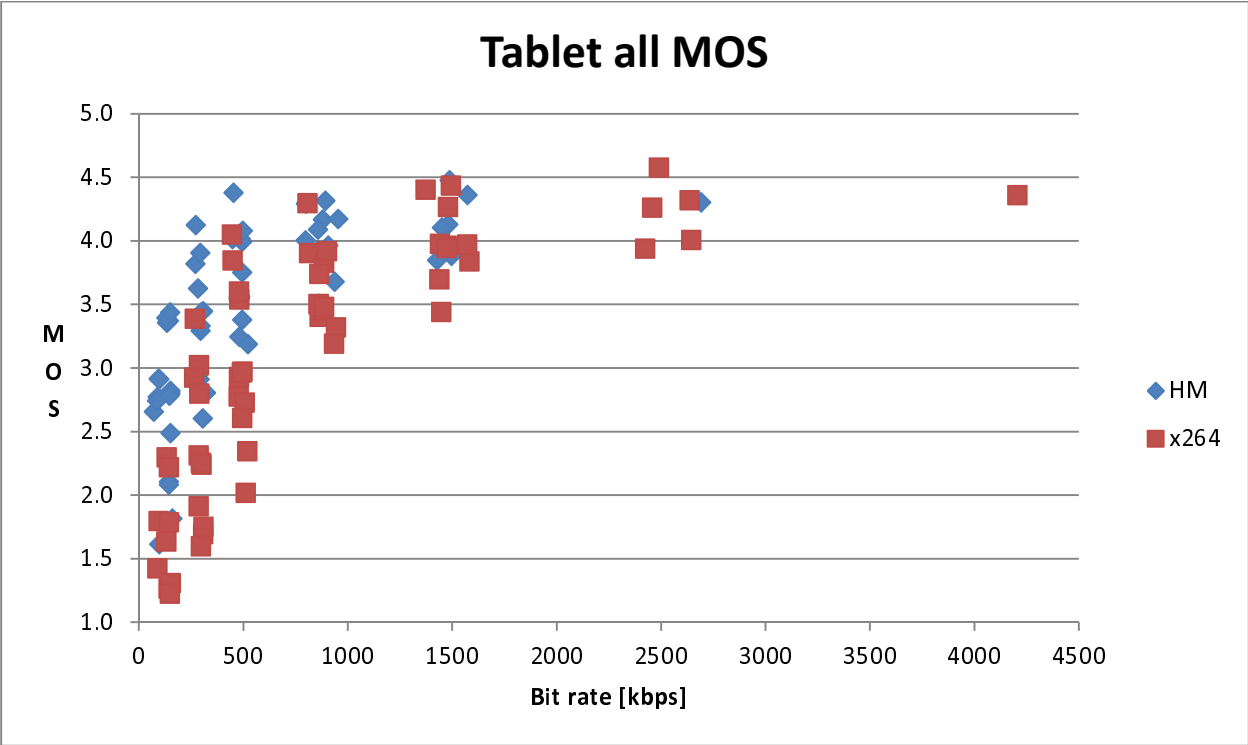

Figure 33: Tablet all MOS
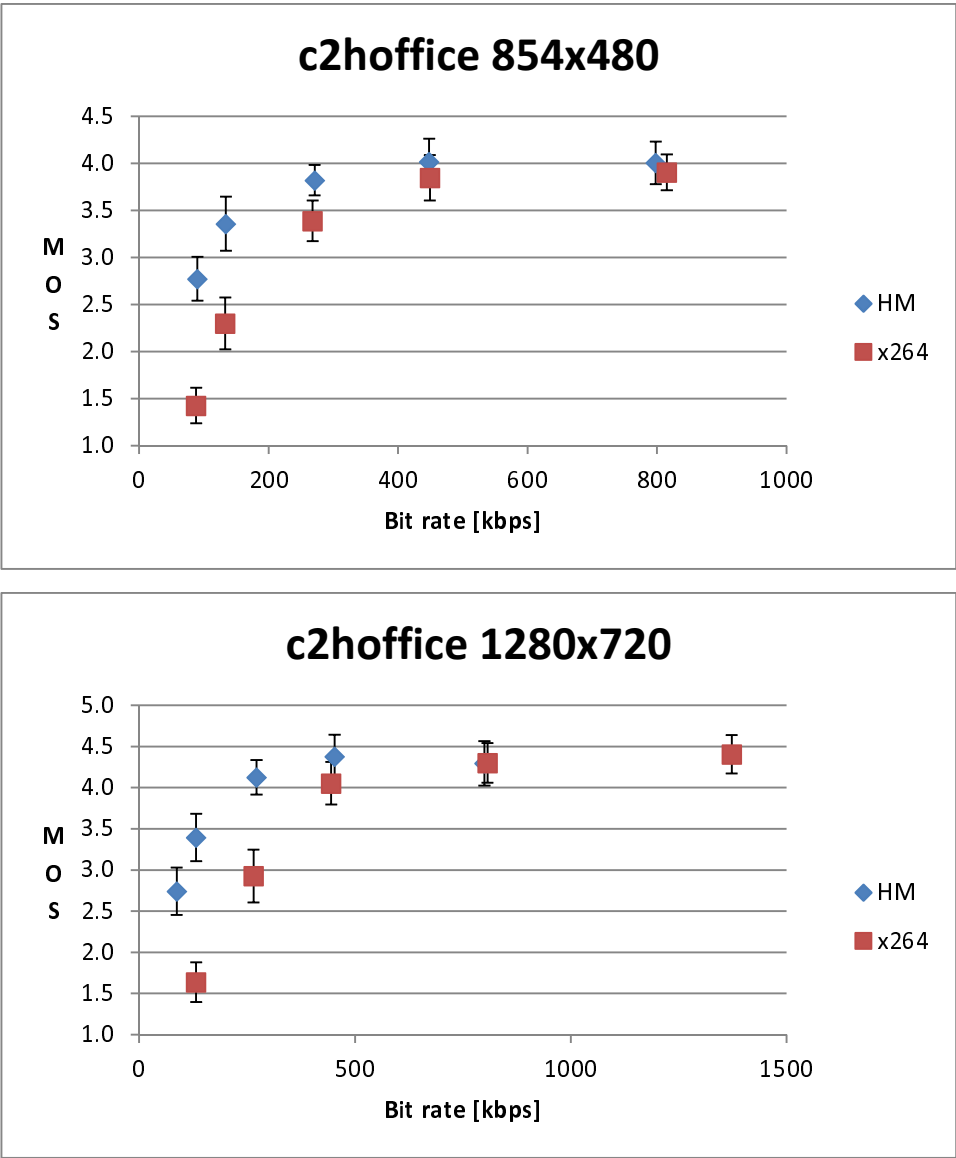
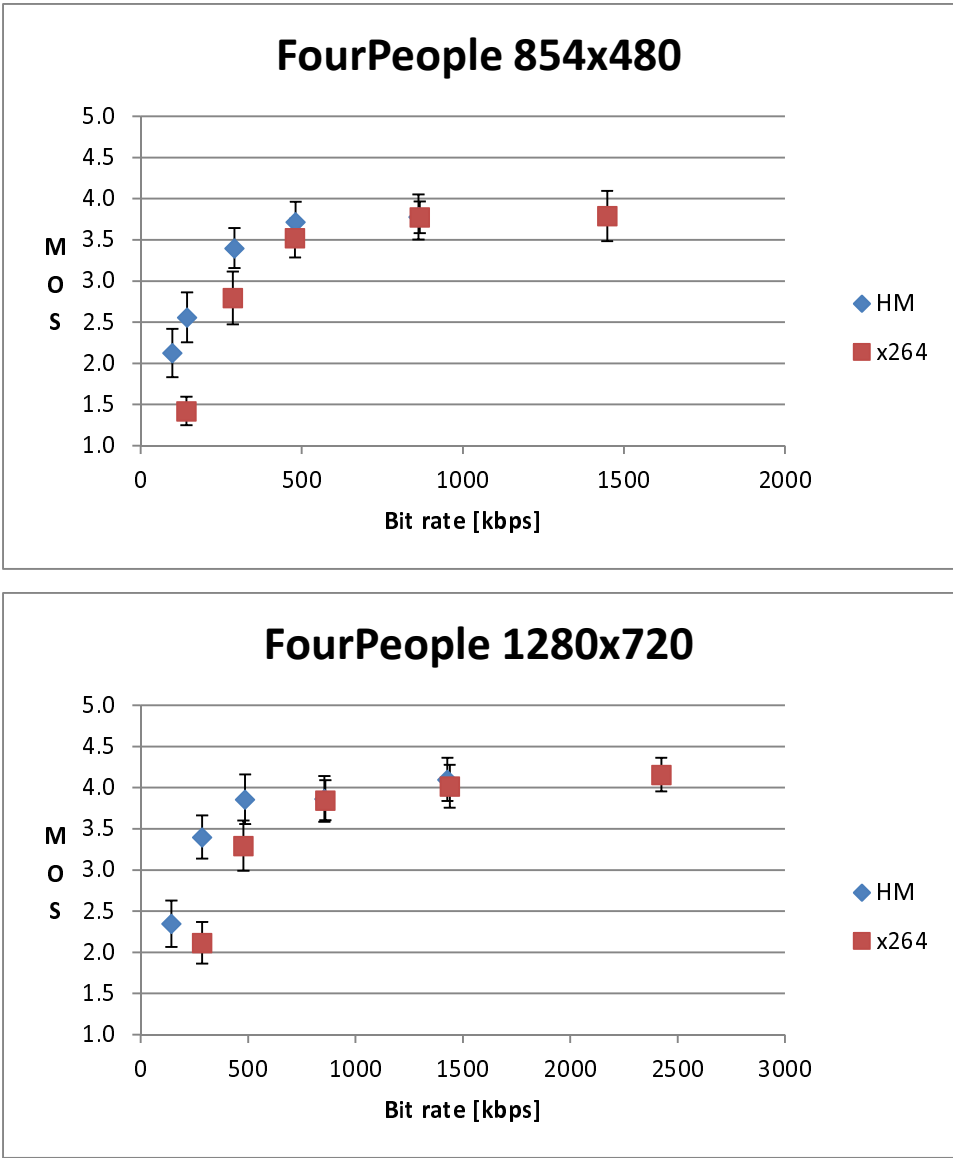**Figure 34: Quality vs. bit rate for c2hoffice**
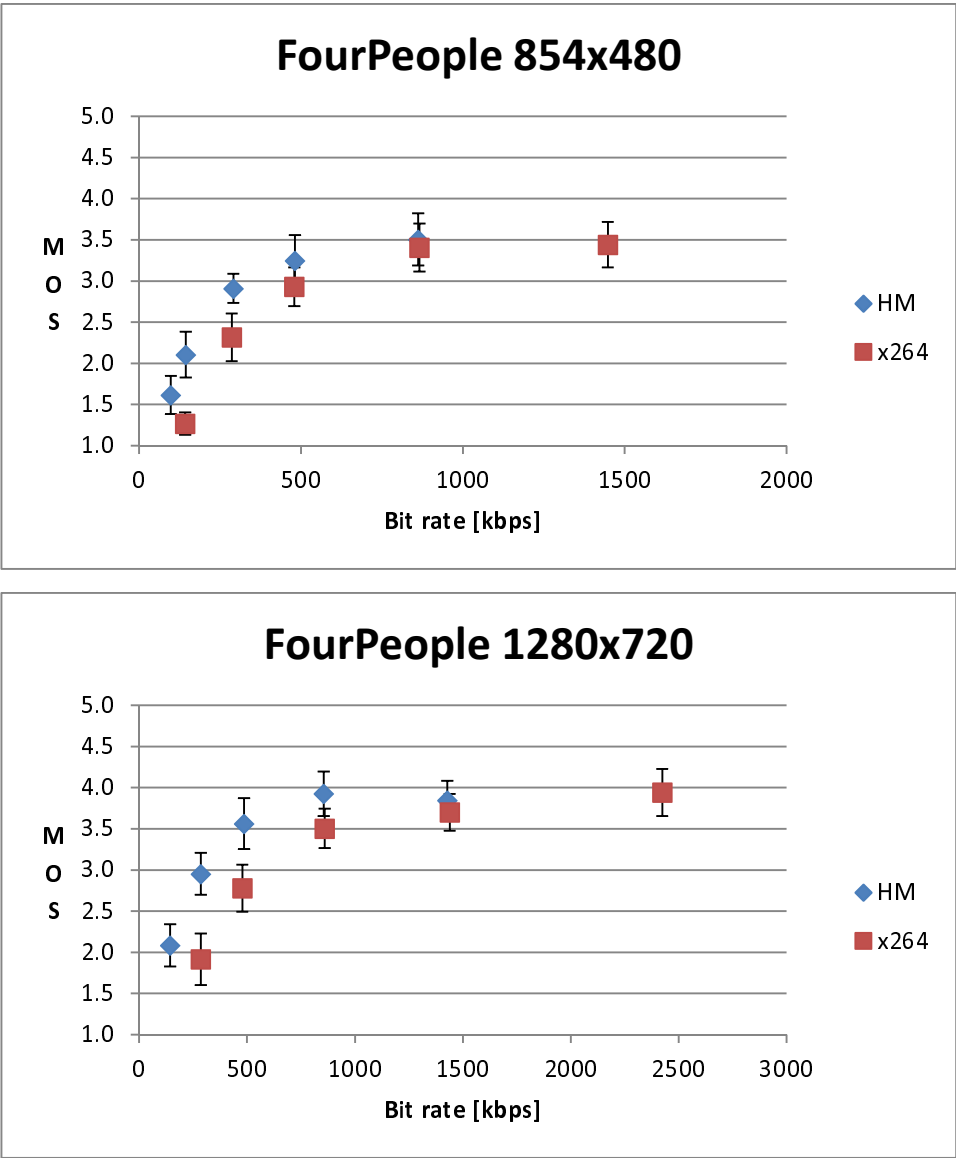
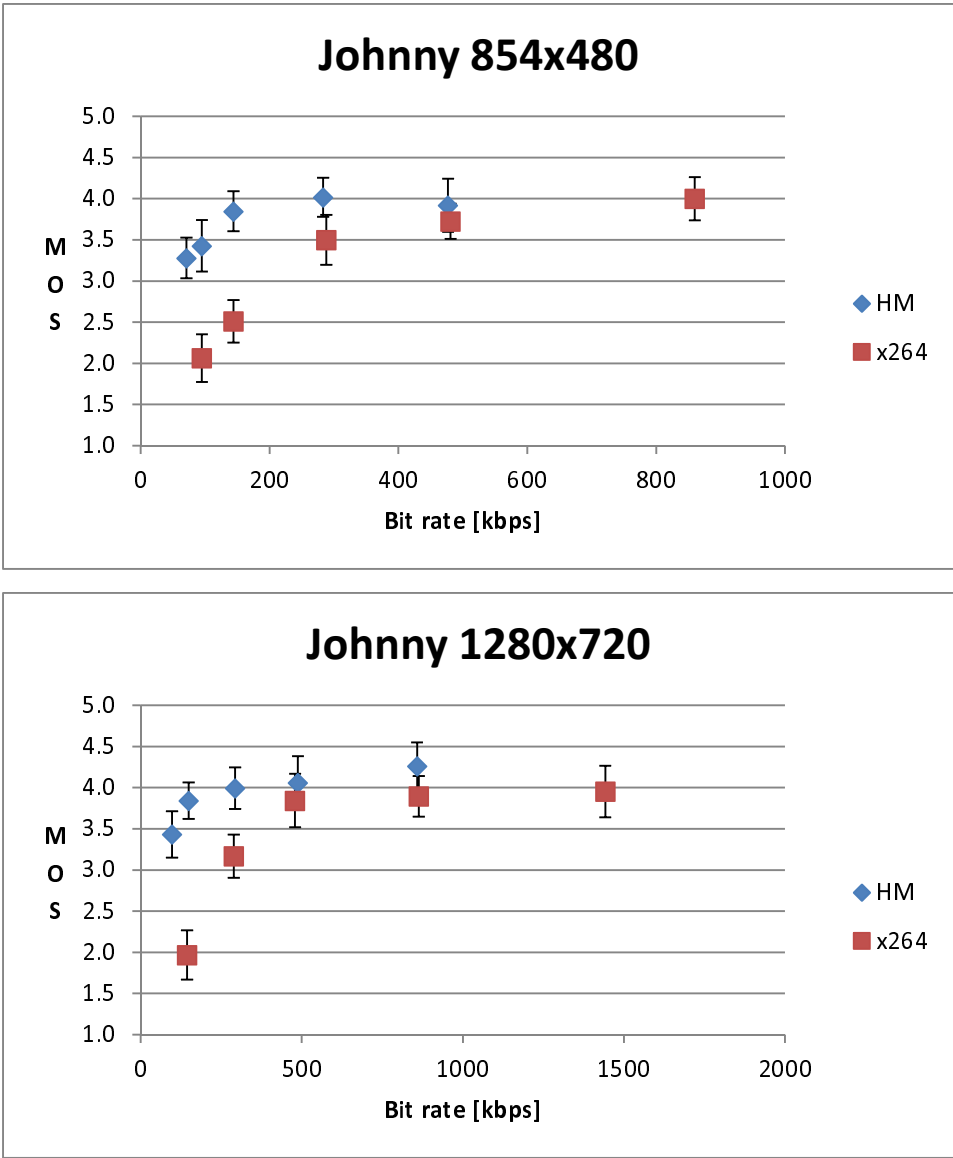**Figure 35: Quality vs. bit rate for FourPeople**
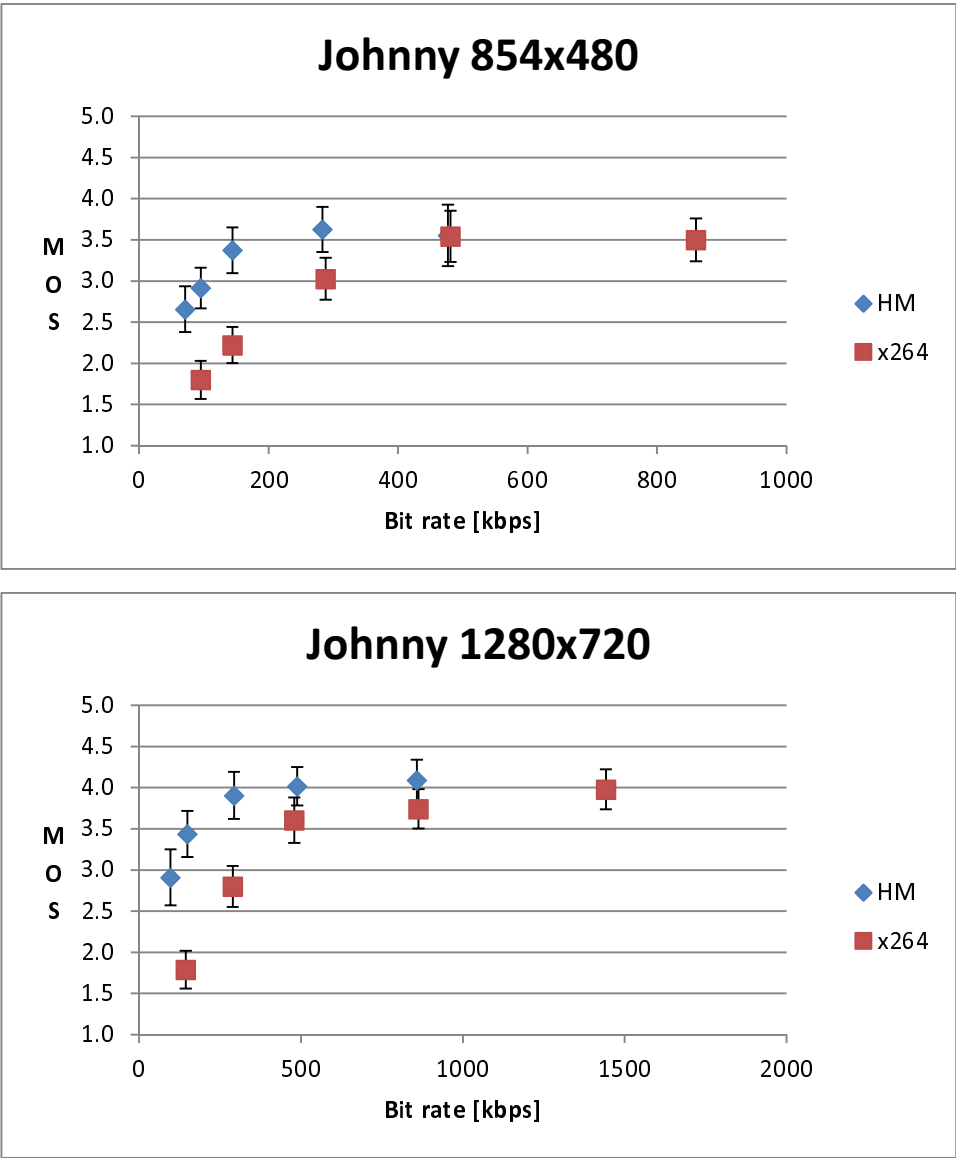
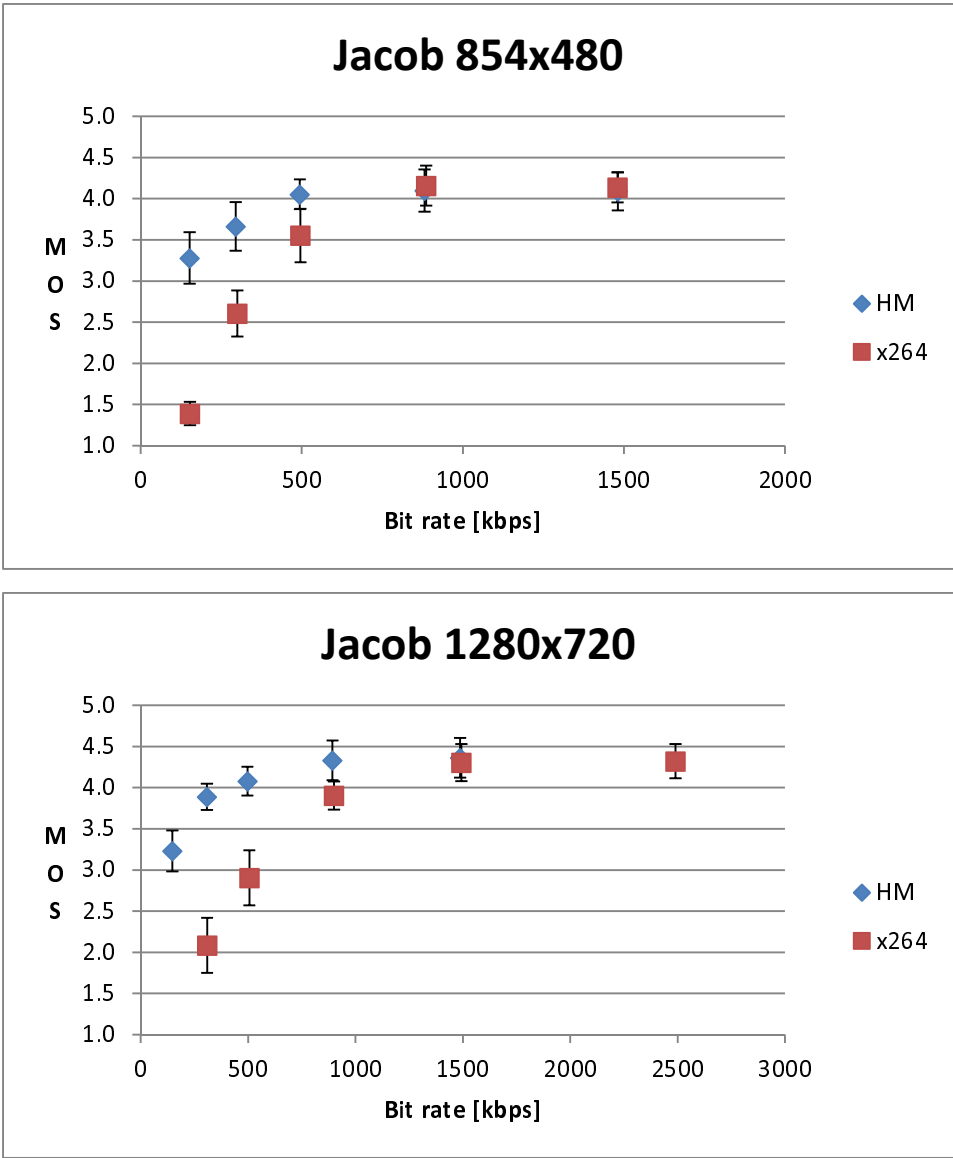**Figure 36: Quality vs. bit rate for Johnny**
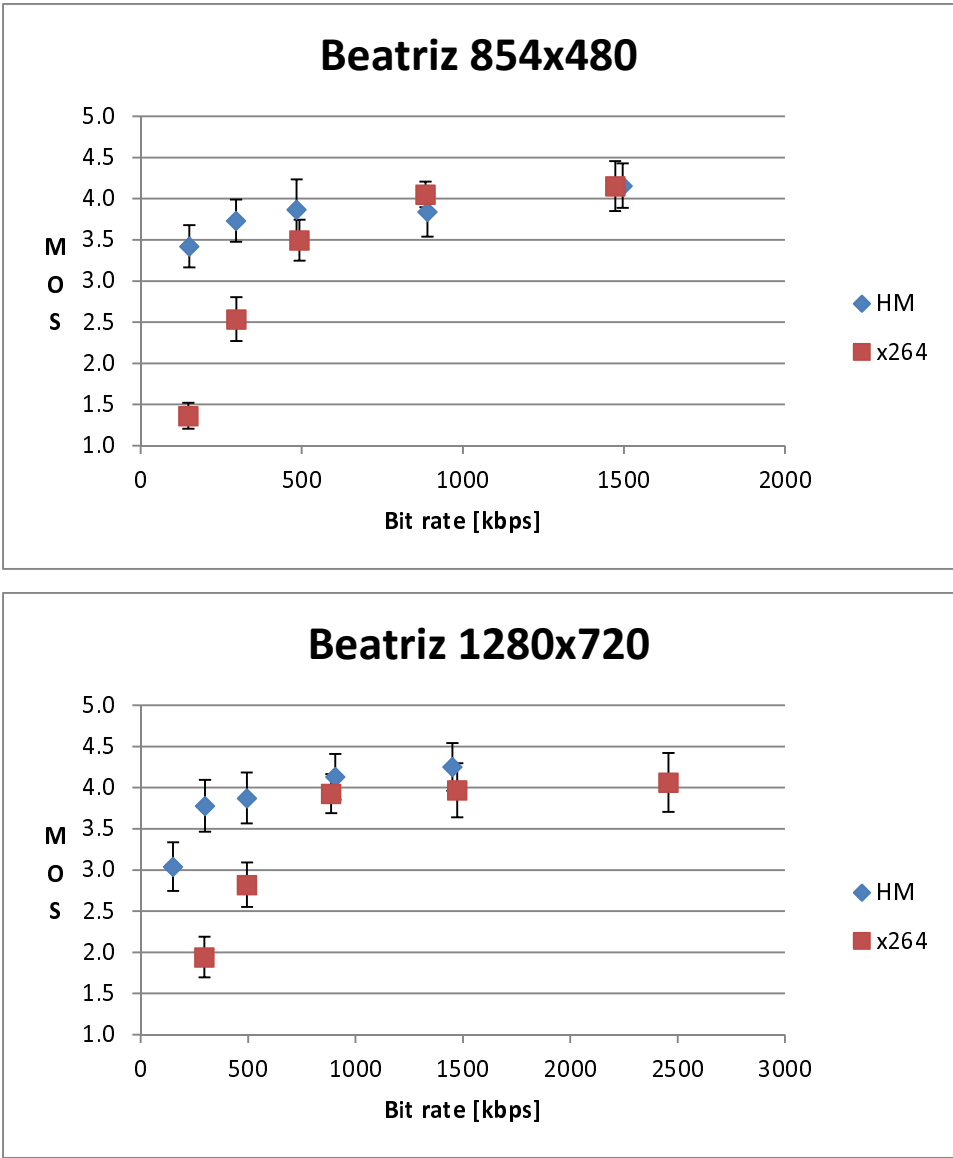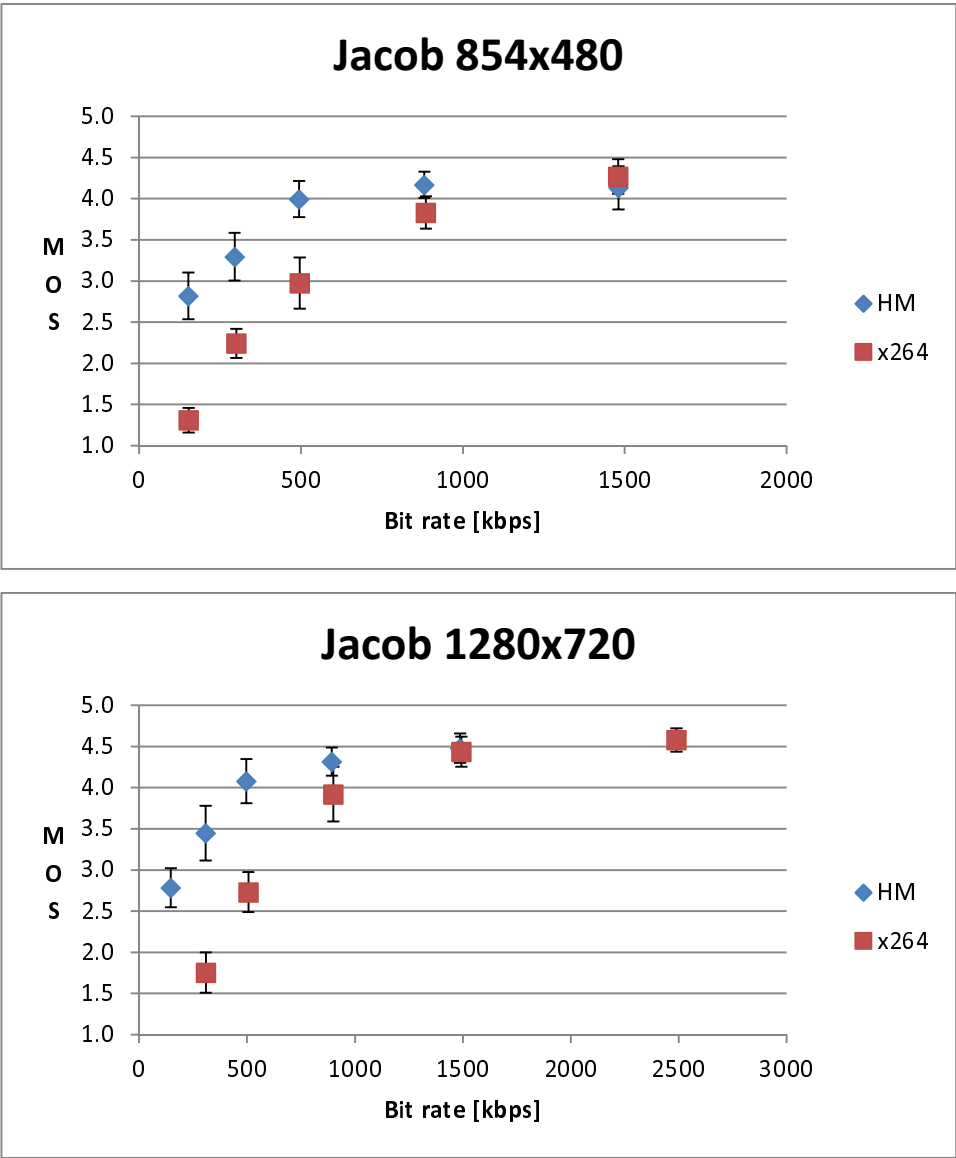
Figure 37: Quality vs. bitrate for Jacob

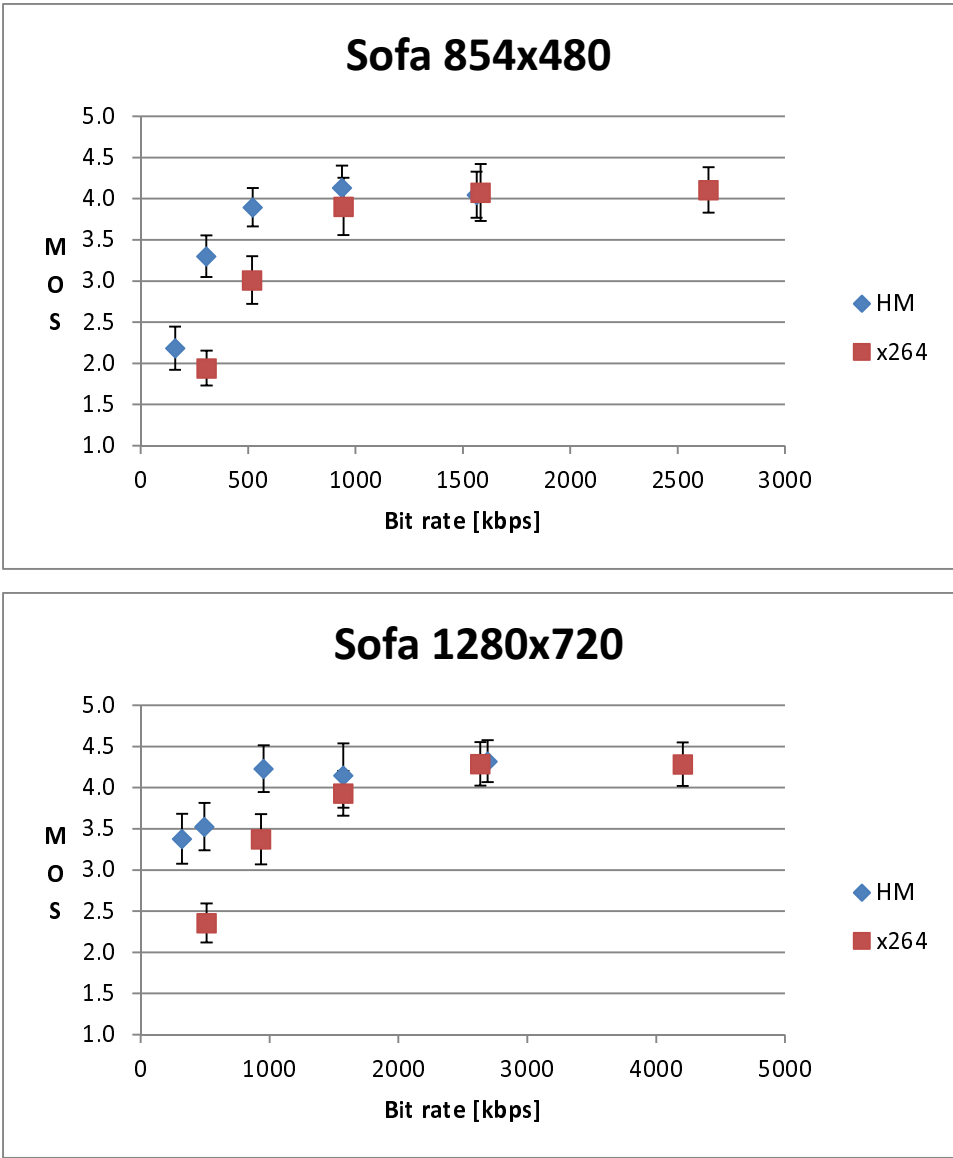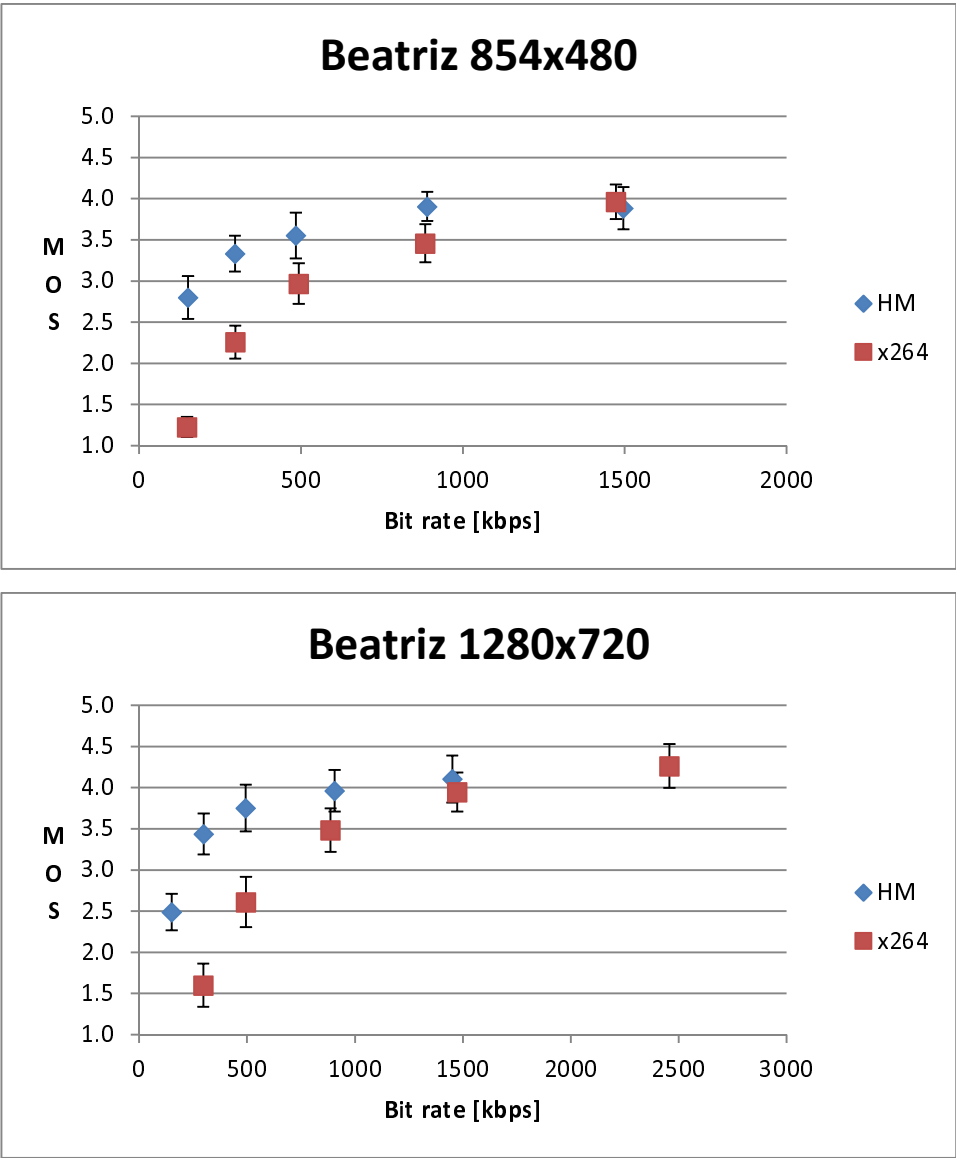**Figure 38: Quality vs. bit rate for Beatriz.**
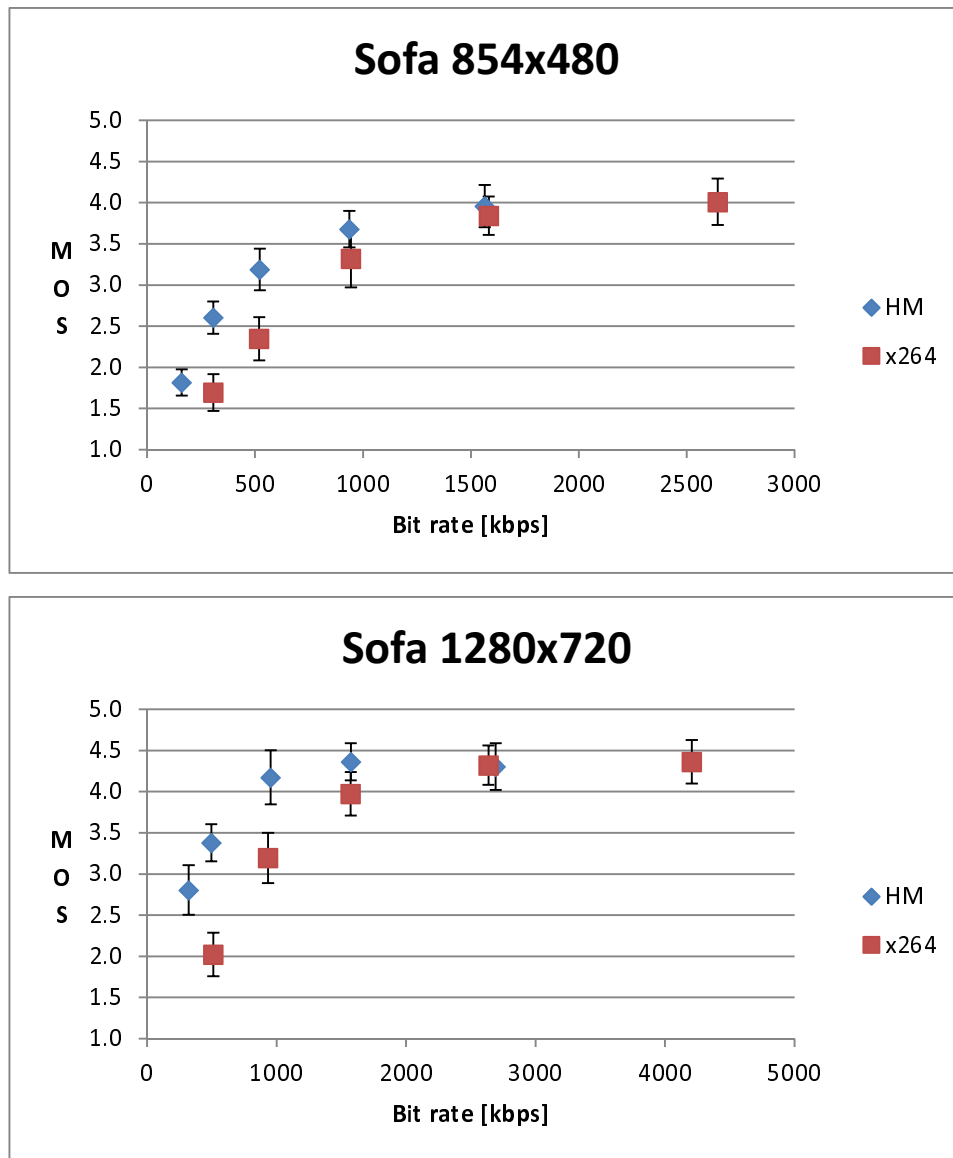
Figure 39: Quality vs. bit rate for Sofa.

The bit rates required to achieve MOS = 3.5 (in between "Fair" and "Good") are estimated (estimated using a LOG-LOG estimation) and average gains are summarized in Table 23.

Table 23: Minimum bit rate [kbps] to achieve MOS = 3.5 for tablets.

Content	Format	Estimated bit rate (kbps) to achieve MOS 3.5	
		HM	x264
Conf2_Home_office	854x480	237	401
FourPeople	854x480	624	931
Johnny	854x480	266	564
Conf1_Jacob	854x480	407	798
Conf3_Beatriz	854x480	520	869
Conf3_Sofa	854x480	842	1398
Conf2_Home_office	1280x720	198	512
FourPeople	1280x720	637	1191
Johnny	1280x720	225	671
Conf1_Jacob	1280x720	359	993
Conf3_Beatriz	1280x720	496	1166
Conf3_Sofa	1280x720	690	1588
Average	**854x480**	483	827
Average	**1280x720**	434	1020

Note that at low bit rates, HM related MOS scores are higher for 480p than for 720p resolutions. This is because un-encoded 720p has 55% more pixels than 480p and therefore looks sharper and better. Therefore, at a given bit rate, 720p has to be compressed harder than 480p, i.e. each pixel is represented with fewer bits. The encoding process introduces an amount of encoding noise that increases with decreasing amount of bits per pixel. Although 720p has inherently better quality than 480p, sufficiently low bit rate will introduce an amount of encoding noise that reduces the 720p quality to the extent that 480p becomes better, because the encoding noise per pixel for 480p is lower due to fewer pixels.

7.6.4   Summary of the subjective tests

In summary, everything else being equal, to achieve a MOS = 3.5 (i.e. in between "Fair" and "Good"):

- on smartphones:

  - At 480p: H.265 (HEVC) (HM) requires 278 kbps on average whereas H.264 (AVC) (x264) requires 644 kbps on average.

  - At 720p: H.265 (HEVC) (HM) requires 274 kbps on average whereas H.264 (AVC) (x264) requires 901 kbps on average.

On tablets:

  - At 480p: H.265 (HEVC) (HM) requires 483 kbps on average and H.264 (AVC) (x264) requires 827 kbps on average.

  - At 720p: H.265 (HEVC) (HM) requires 434 kbps on average and H.264 (AVC) (x264) requires 1020 kbps on average.

Based on those test results we expect the H.265 (HEVC) video codec to enable implementations providing significant efficiency gains in 3GPP real-time video communication services like e.g. video over LTE.

7.7   Summaries of objective test results for MMS and MTSI

Four sets of overlapping QP value ranges, as described in Table 6, were used to compute the BD-rate values. Subclauses 7.6.1 and 7.6.2 provide a detailed list of results for all the bitrate ranges and for various sequences ranging from 240p to 1080p.

For the MTSI service, video conferencing test sequences (FourPeople, Johnny, and KristenAndSara) encoded with the lossy encoding configuration at the medium bitrate range (corresponding to bit rates between 0,2 to 1,3 Mbps) form the most relevant test case. For this test case, the BD-rate decrease for H.265 (HEVC) over H.264 (AVC) Constrained Baseline Profile is 60% (i.e. H.265 (HEVC) uses 2/5th of the bit rate of H.264 (AVC) Baseline Profile). When compared with H.264 (AVC) High Profile, the BD-rate decrease for H.265 (HEVC) over H.264 (AVC) is 50% (i.e. H.265 (HEVC) uses half of the bit rate of H.264 (AVC) High Profile).

When all the results are averaged, the following conclusions can be made. The average BD-rate decrease for H.265 (HEVC) when compared to H.264 (AVC) was around 35% for MMS. For MTSI, the average BD-rate decrease of H.265 (HEV–) was 40 – 4% and 30 – 35%, when compared to H.264 (AVC) Constrained Baseline profile and H.264 (AVC) High profile, respectively. The results are further summarized below:

- The BD-rate decrease for H.265 (HEVC) under MMS test conditions is roughly 35%, and is consistent across various RAP periods and GOP structures (open and closed GOP).

- The performance gap of H.265 (HEVC) compared to H.264 (AVC) under MMS conditions increases with decreasing bit rates. At low and medium bit rates, the performance gap of H.265 (HEVC) compared to H.264 (AVC) under MMS conditions increases with increasing resolution.

- The BD-rate decrease for H.265 (HEVC) under MTSI test conditions is higher when compared to H.264 (AVC) Constrained Baseline profile than H.264 (AVC) High Profile.

- Enabling the temporal prediction and QP cascading for H.265 (HEVC) and H.264 (AVC) results in higher improvement in the BD-rate performance of H.265 (HEVC) when compared to H.264 (AVC). The BD-rate improvement (of H.265 (HEVC) compared to H.264 (AVC) Constrained Baseline Profile) was around 40% when no temporal prediction was used and around 50% when 3 temporal layers were used.

7.7.1   Results for MMS

The summarized BD-rate results are presented in Table 24 to Table 27. The results for various prediction structures and RAP periods are presented in separate tables. Figure 40 to Figure 43 show the plots of PSNR versus bit rate for a typical sequence (BasketballDrive) under closed GOP structure with 2-second RAP period for various picture resolutions (240p, 480p, 720p and 1080p).

**Table 24: BD-rate of H.265 (HEVC) compared to H.264 (AVC)
for open GOP structure with 1 sec RAP period**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-31.7%	-43.8%	-45.3%	-42.0%	-44.1%	-42.8%	-53.4%	-67.5%	-70.1%	-43.1%	-49.5%	-49.5%
720p	-32.0%	-39.1%	-37.8%	-36.8%	-39.3%	-37.1%	-47.6%	-62.4%	-66.4%	-38.4%	-43.9%	-43.1%
480p	-29.5%	-37.0%	-36.2%	-33.5%	-38.6%	-36.8%	-42.1%	-58.9%	-62.2%	-34.6%	-42.1%	-41.7%
240p	-27.4%	-36.3%	-34.5%	-28.8%	-32.5%	-29.7%	-32.6%	-46.9%	-50.0%	-29.5%	-36.8%	-35.1%
Overall	**-30.1%**	**-37.6%**	**-37.0%**	**-34.2%**	**-37.6%**	**-35.9%**	**-41.3%**	**-57.8%**	**-61.4%**	**-35.1%**	**-41.9%**	**-41.3%**

**Table 25: BD-rate of H.265 (HEVC) compared to H.264 (AVC)
for open GOP structure with 2 sec RAP period**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-33.0%	-46.1%	-48.6%	-44.1%	-48.7%	-48.3%	-56.4%	-68.7%	-71.5%	-45.1%	-52.6%	-52.9%
720p	-33.9%	-42.1%	-41.5%	-39.3%	-44.2%	-42.4%	-50.4%	-64.4%	-67.8%	-40.7%	-47.4%	-46.6%
480p	-31.3%	-40.2%	-39.6%	-35.9%	-43.6%	-41.8%	-44.7%	-60.3%	-62.7%	-36.8%	-45.6%	-44.9%
240p	-29.2%	-39.7%	-37.6%	-31.1%	-37.1%	-33.9%	-34.7%	-48.0%	-51.6%	-31.6%	-40.2%	-38.8%
Overall	**-31.8%**	**-40.5%**	**-40.3%**	**-36.4%**	**-42.1%**	**-40.3%**	**-43.8%**	**-58.9%**	**-62.4%**	**-37.2%**	**-45.1%**	**-44.6%**

**Table 26: BD-rate of H.265 (HEVC) compared to H.264 (AVC)
for closed GOP structure with 1 sec RAP period**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-30.6%	-41.5%	-42.6%	-40.7%	-41.3%	-40.0%	-52.7%	-66.6%	-69.6%	-42.0%	-47.5%	-47.6%
720p	-30.8%	-36.9%	-35.5%	-35.4%	-36.4%	-34.0%	-46.8%	-61.5%	-65.7%	-37.2%	-41.9%	-41.1%
480p	-28.6%	-35.2%	-34.3%	-32.3%	-36.0%	-34.2%	-41.2%	-57.7%	-61.3%	-33.6%	-40.3%	-39.9%
240p	-26.4%	-34.5%	-32.7%	-27.6%	-29.7%	-26.9%	-31.7%	-45.1%	-48.7%	-28.4%	-34.8%	-33.2%
Overall	**-29.1%**	**-35.6%**	**-35.0%**	**-33.0%**	**-35.0%**	**-33.2%**	**-40.5%**	**-56.6%**	**-60.6%**	**-34.0%**	**-40.1%**	**-39.5%**

**Table 27: BD-rate of H.265 (HEVC) compared to H.264 (AVC)
for closed GOP structure with 2 sec RAP period**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-32.5%	-45.1%	-47.4%	-43.5%	-47.4%	-46.9%	-56.0%	-68.3%	-71.1%	-44.6%	-51.7%	-52.0%
720p	-33.4%	-41.1%	-40.4%	-38.6%	-42.7%	-40.8%	-49.9%	-63.8%	-67.5%	-40.1%	-46.4%	-45.6%
480p	-30.8%	-39.4%	-38.7%	-35.3%	-42.4%	-40.5%	-44.2%	-59.9%	-62.4%	-36.3%	-44.7%	-44.0%
240p	-28.7%	-38.9%	-36.7%	-30.5%	-35.8%	-32.4%	-34.3%	-47.1%	-50.9%	-31.1%	-39.2%	-37.8%
Overall	**-31.3%**	**-39.6%**	**-39.4%**	**-35.8%**	**-40.8%**	**-39.0%**	**-43.4%**	**-58.4%**	**-62.0%**	**-36.6%**	**-44.2%**	**-43.7%**

Figure 40: BasketballDrive 240p sequence under closed GOP structure and 2 sec RAP period for MMS test conditions

Figure 41: BasketballDrive 480p sequence under closed GOP structure and 2 sec RAP period for MMS test conditions
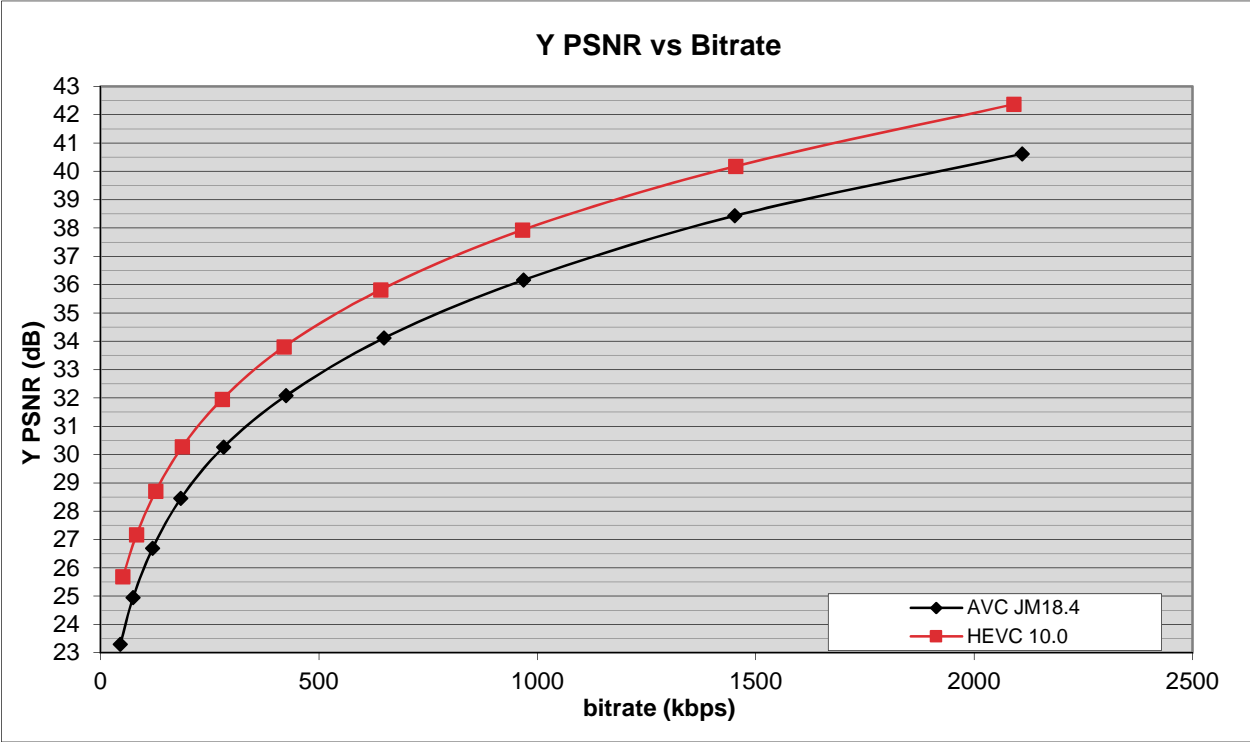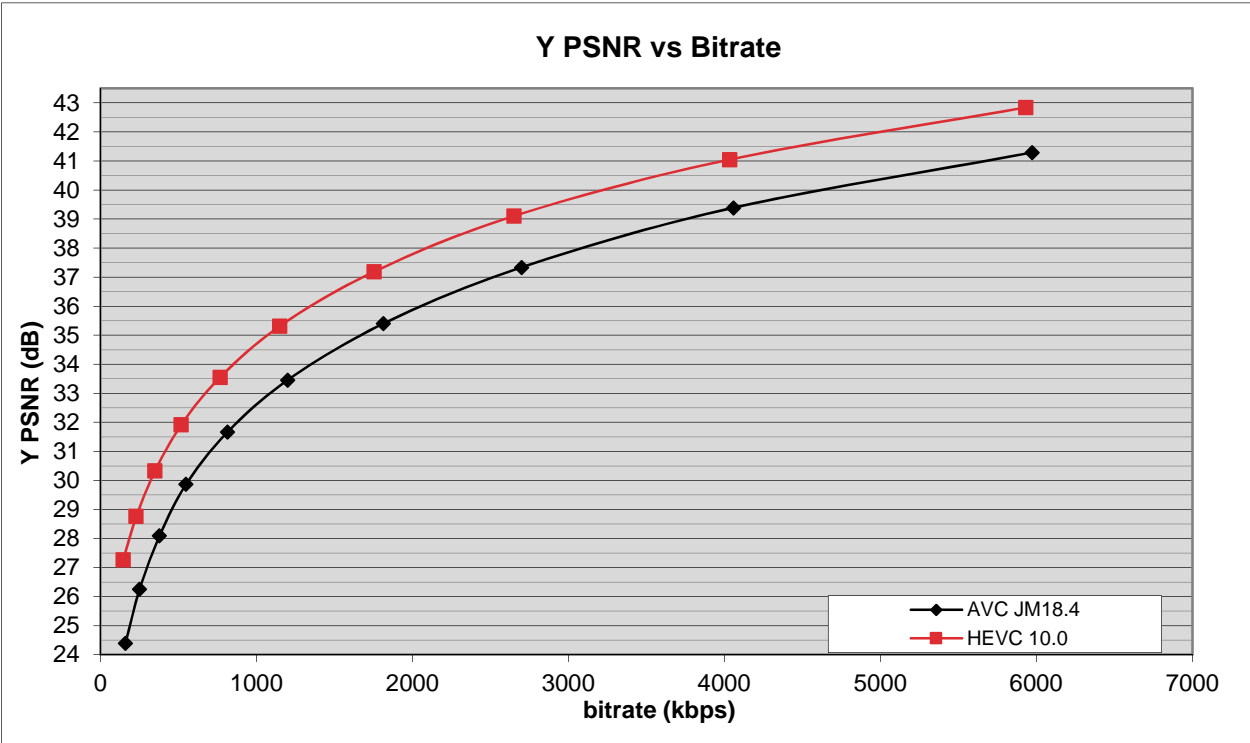
**Figure 42: BasketballDrive 720p sequence under closed GOP structure
and 2 sec RAP period for MMS test conditions**

**Figure 43: BasketballDrive 1080p sequence under closed GOP structure
and 2 sec RAP period for MMS test conditions**
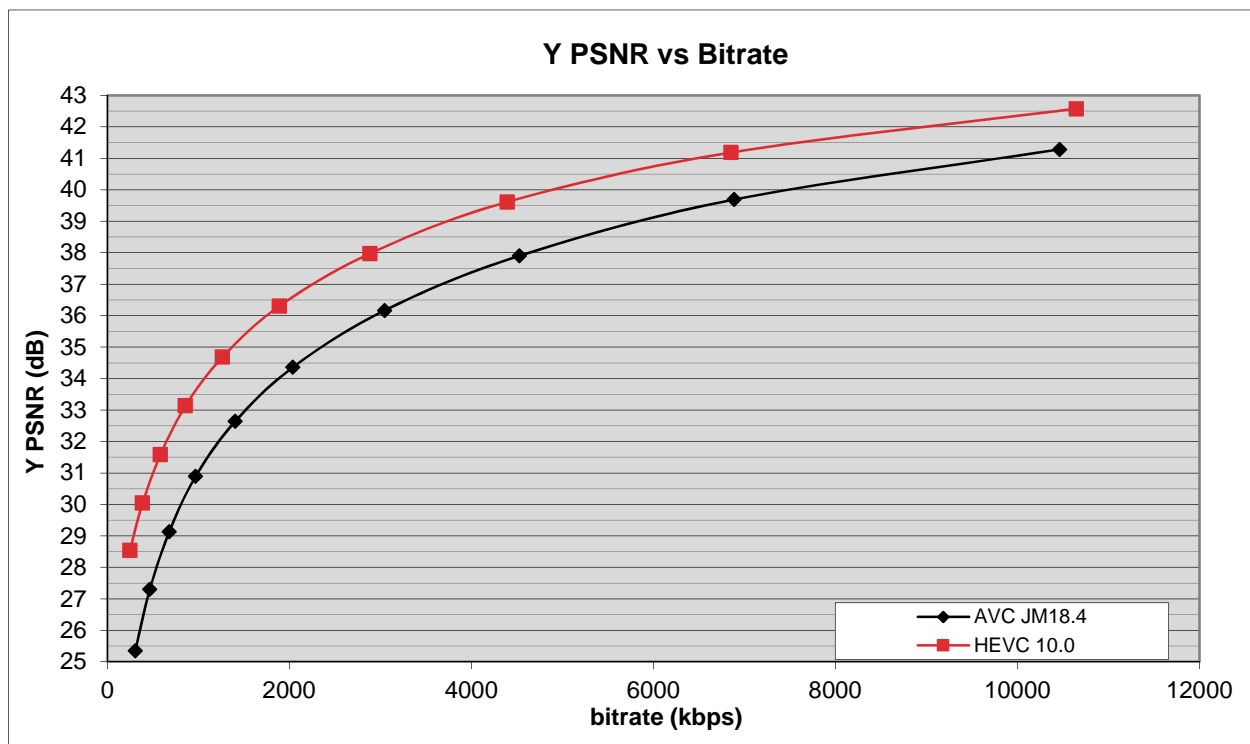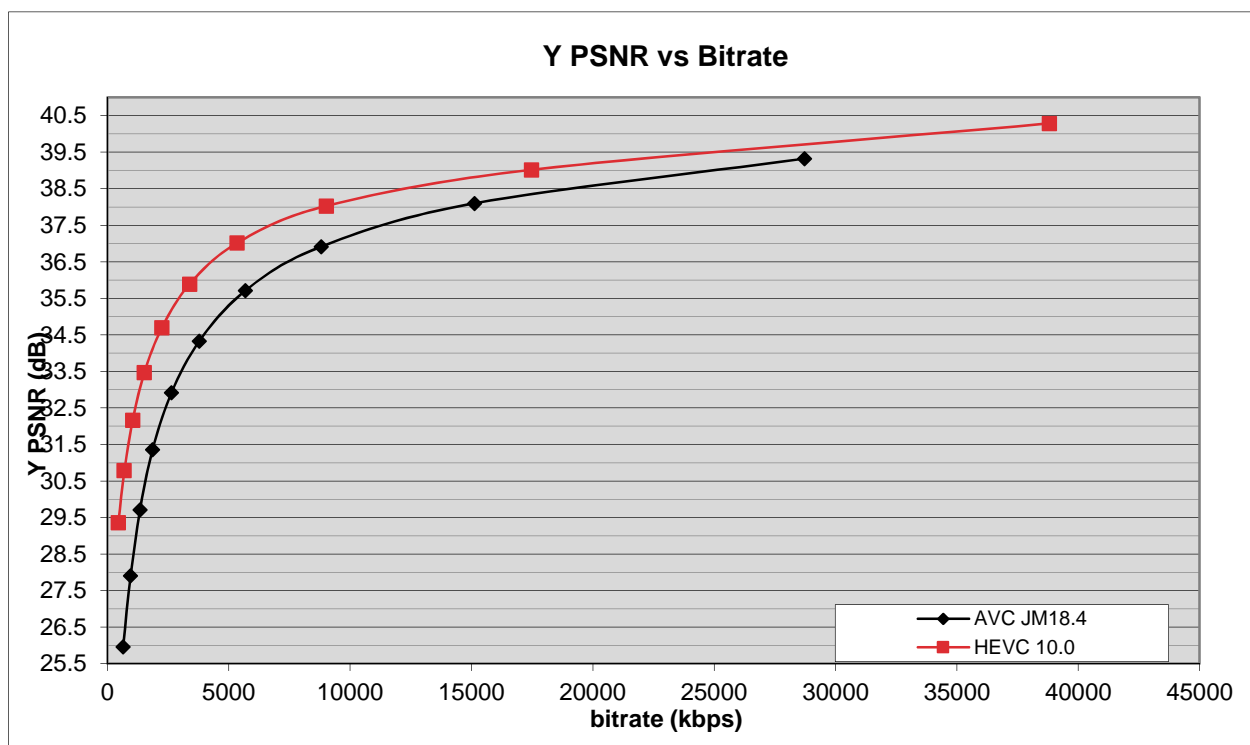
7.7.2   Results for MTSI

The results comparing H.265 (HEVC) and H.264 (AVC) are provided for both the Constrained Baseline and the High profile of H.264 (AVC). The two prediction structures that were tested are described in Table 28.

Table 28: Description of prediction structures used for MTSI test conditions

Test case	Prediction structure
MTSI_1	Temporal scalability is not supported The previous two pictures in decoding order are always used for prediction.
MTSI_2	To test conditions with packet losses. Temporal scalability with 3 temporal sub-layers is supported. Each picA occurring at (or immediately after) the end of one second intervals (the first interval begins from the first picture, which is an IDR picture), uses the picture that precedes the outcame of picA by roughly 300 ms and that belongs to the same or a lower temporal sub-layer as the reference picture for prediction. For all other pictures, the two pictures preceding in decoding order that belong to the same or a lower temporal sub-layer are used for prediction.

7.7.2.1   H.265 (HEVC) Main profile vs. H.264 (AVC) Constrained Baseline profile

Table 29 and Table 30 provide a summary of the BD-rate results for comparison of H.265 (HEVC) Main Profile and H.264 (AVC) Constrained Baseline Profile for the two MTSI test cases. Figure 44 to Figure 47 show the plots of PSNR versus bit rate for two typical 720p video conferencing sequences (Johnny and FourPeople) under MTSI_1 and MTSI_2 prediction structures.

Table 29: BD-rate of H.265 (HEVC) compared to H.264 (AVC) (Constrained Baseline Profile) for prediction structure MTSI_1

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-36.7%	-35.5%	-39.5%	-50.1%	-47.9%	-51.5%	-60.9%	-70.5%	-74.9%	-50.1%	-47.3%	-51.9%
720p	-37.4%	-34.9%	-36.7%	-46.9%	-46.3%	-49.0%	-54.3%	-66.4%	-71.4%	-46.9%	-47.1%	-48.9%
480p	-32.9%	-32.4%	-33.7%	-39.5%	-41.8%	-43.6%	-48.6%	-60.4%	-62.8%	-40.1%	-41.1%	-42.4%
240p	-27.4%	-27.7%	-28.4%	-33.4%	-35.6%	-38.7%	-41.7%	-51.5%	-53.1%	-33.6%	-34.5%	-35.3%
Overall	**-33.1%**	**-32.2%**	**-33.8%**	**-41.3%**	**-42.1%**	**-44.9%**	**-50.0%**	**-61.0%**	**-64.2%**	**-41.6%**	**-41.7%**	**-43.6%**

Table 30: BD-rate of H.265 (HEVC) compared to H.264 (AVC) (Constrained Baseline Profile) for prediction structure MTSI_2

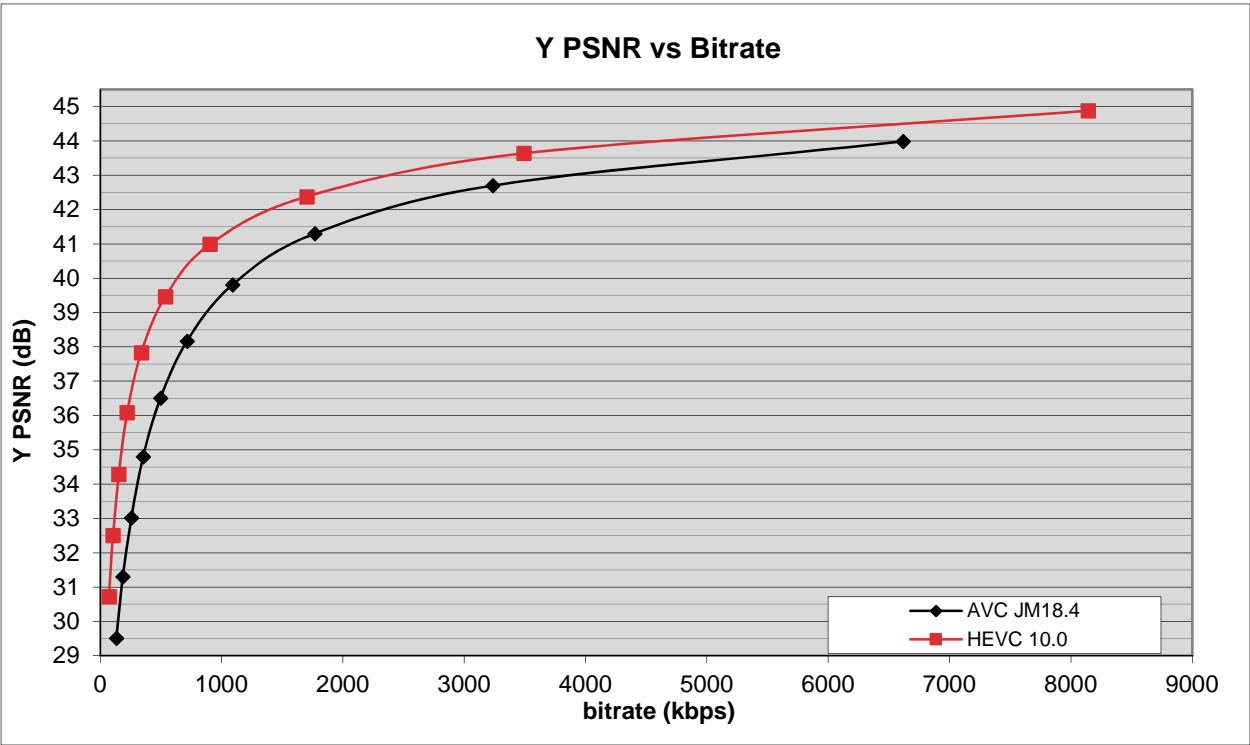|	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-40.9%	-48.1%	-54.4%	-57.6%	-62.7%	-64.6%	-68.2%	-78.0%	-80.6%	-57.2%	-62.1%	-64.8%
720p	-44.8%	-50.1%	-52.2%	-55.8%	-61.3%	-62.7%	-63.6%	-75.7%	-78.7%	-55.5%	-61.1%	-62.5%
480p	-39.1%	-45.0%	-46.3%	-48.3%	-57.4%	-57.8%	-57.8%	-69.4%	-71.6%	-48.5%	-55.7%	-56.3%
240p	-33.8%	-42.6%	-42.6%	-43.7%	-55.0%	-54.0%	-51.7%	-62.2%	-64.2%	-42.7%	-51.4%	-51.8%
Overall	**-39.3%**	**-46.1%**	**-48.0%**	**-50.4%**	**-58.6%**	**-59.1%**	**-59.2%**	**-70.3%**	**-72.7%**	**-50.1%**	**-56.9%**	**-57.9%**

Figure 44: Johnny 720p video conferencing sequence under MTSI test conditions and MTSI_1 prediction structure and AVC Constrained Baseline Profile
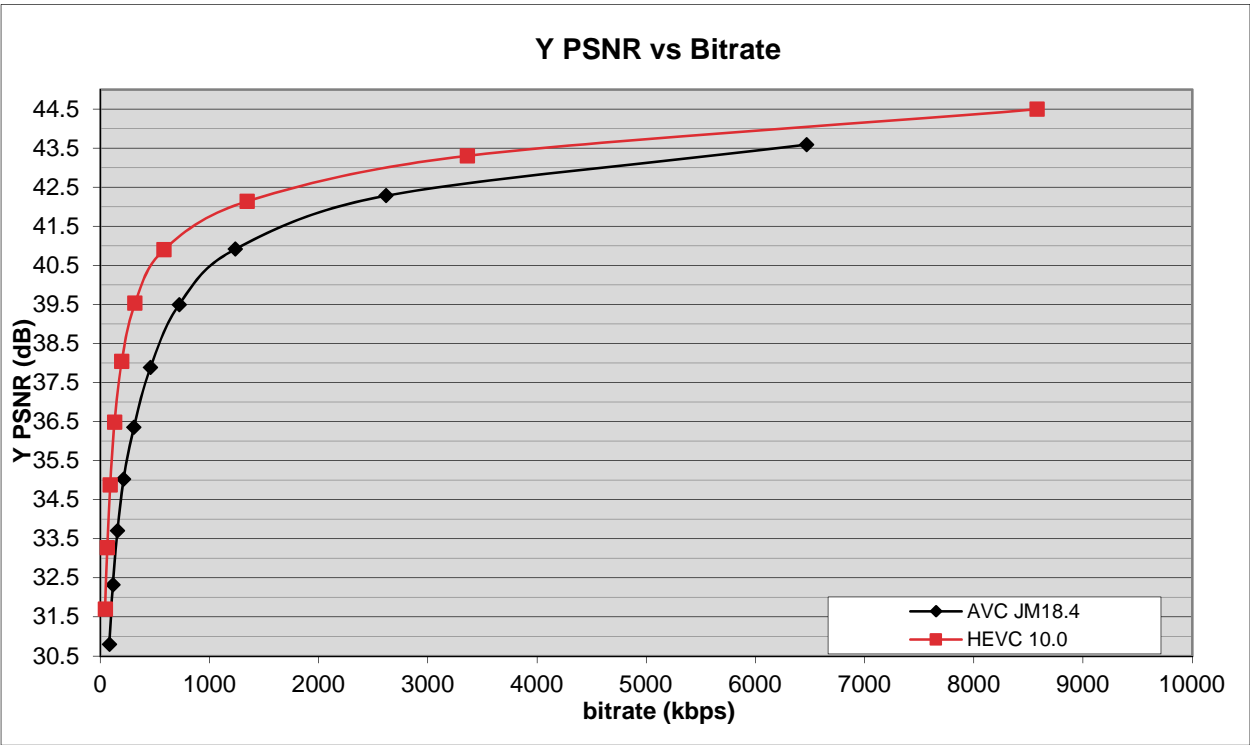

Figure 45: FourPeople 720p video conferencing sequence under MTSI test conditions and MTSI_1 prediction structure and AVC Constrained Baseline Profile
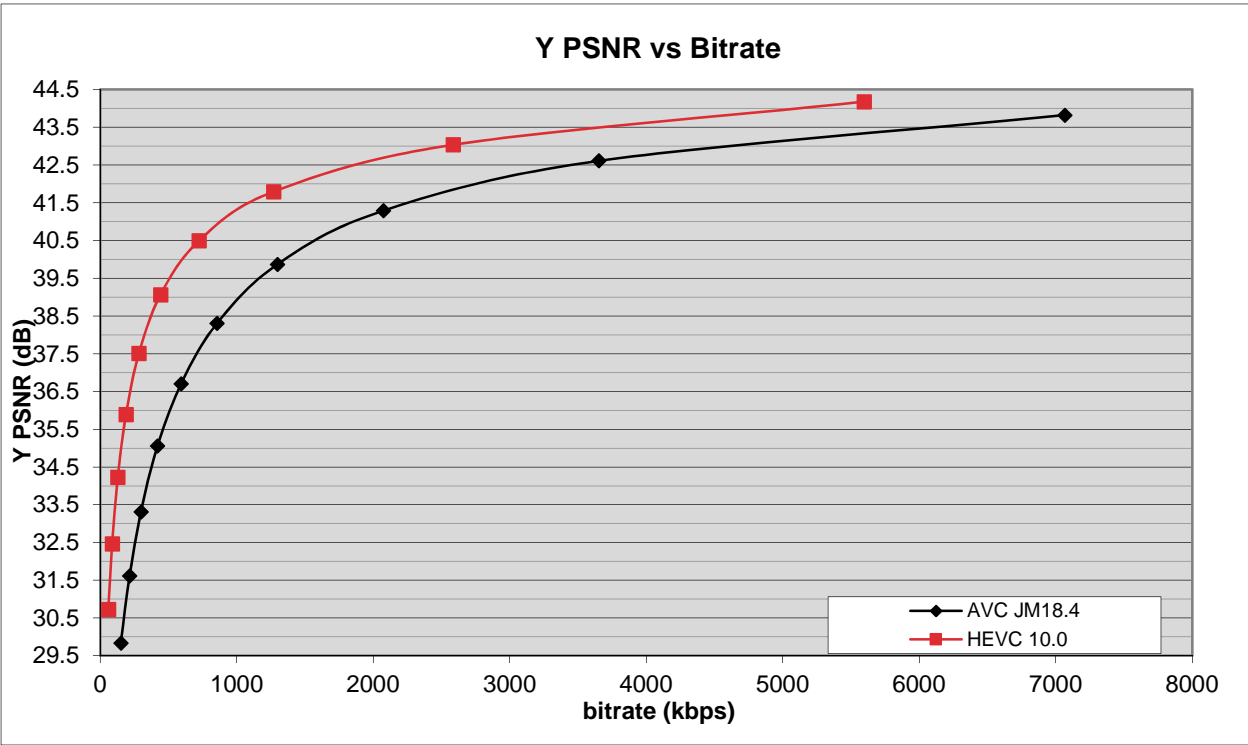
Figure 46: Johnny 720p video conferencing sequence under MTSI test conditions and MTSI_2 prediction structure and AVC Constrained Baseline Profile
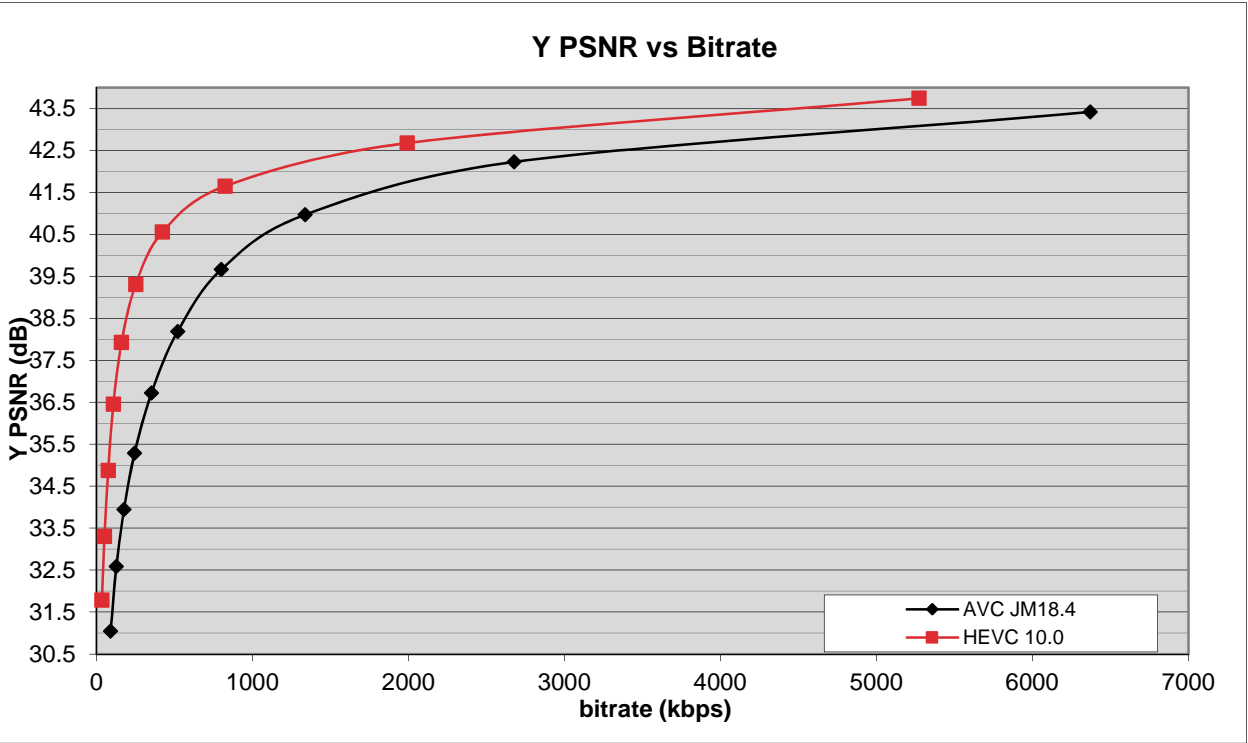

Figure 47: FourPeople 720p sequence under MTSI test conditions and MTSI_2 prediction structure and AVC Constrained Baseline Profile

7.7.2.2   H.265 (HEVC) Main profile vs. H.264 (AVC) High profile

Table 31 and Table 32 provide a summary of the BD-rate results for comparison of H.265 (HEVC) Main Profile and H.264 (AVC) High Profile for the two MTSI test cases. Figure 48 to Figure 51 show the plots of PSNR versus bit rate for two typical 720p video conferencing sequences (Johnny, FourPeople) under MTSI_1 prediction structure.

**Table 31: BD-rate of H.265 (HEVC) compared to H.264 (AVC) (High Profile)
for prediction structure MTSI_1**

	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-25.9%	-30.9%	-35.3%	-39.3%	-41.6%	-45.4%	-49.3%	-66.9%	-71.3%	-38.7%	-41.6%	-46.1%
720p	-27.4%	-29.0%	-31.1%	-36.8%	-39.8%	-43.0%	-42.9%	-62.7%	-67.9%	-36.5%	-41.2%	-43.0%
480p	-24.0%	-26.6%	-28.0%	-29.8%	-36.2%	-38.4%	-38.0%	-57.4%	-59.2%	-30.6%	-35.8%	-37.0%
240p	-19.3%	-22.6%	-23.3%	-24.0%	-30.3%	-33.5%	-31.2%	-47.4%	-48.7%	-24.6%	-29.4%	-30.5%
Overall	**-23.8%**	**-26.8%**	**-28.6%**	**-31.4%**	**-36.3%**	**-39.3%**	**-39.1%**	**-57.4%**	**-60.3%**	**-31.7%**	**-36.2%**	**-38.1%**

**Table 32: BD-rate of H.265 (HEVC) compared to H.264 (AVC) (High Profile)
for prediction structure MTSI_2**

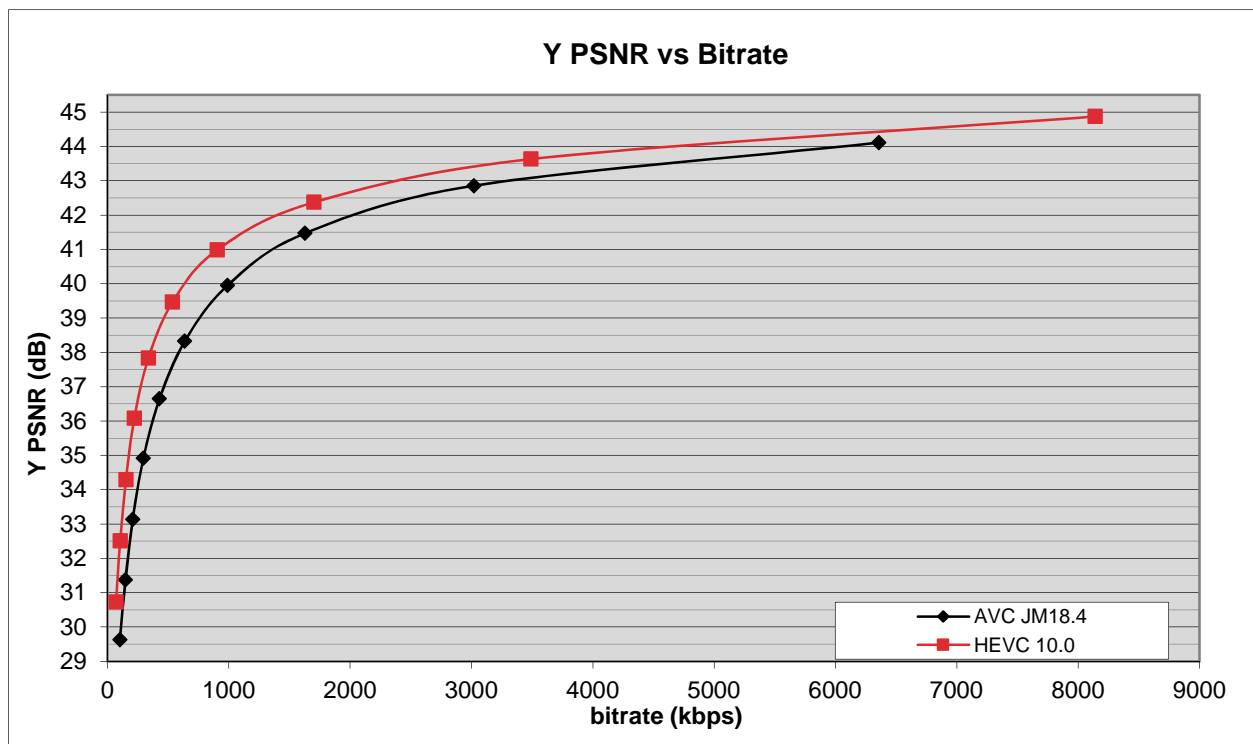|	High bit-rate			Medium bit-rate			Low bit-rate			Overall		
	Y	U	V	Y	U	V	Y	U	V	Y	U	V
1080p	-30.2%	-42.7%	-49.7%	-46.0%	-56.8%	-58.6%	-56.1%	-73.3%	-76.5%	-45.4%	-56.0%	-59.0%
720p	-35.5%	-44.4%	-47.0%	-46.0%	-55.4%	-57.1%	-52.5%	-72.2%	-75.2%	-45.5%	-55.6%	-57.1%
480p	-31.1%	-39.8%	-41.1%	-39.4%	-52.4%	-52.6%	-47.1%	-65.7%	-68.1%	-39.4%	-50.7%	-51.2%
240p	-27.3%	-38.8%	-38.7%	-36.3%	-50.8%	-49.6%	-41.6%	-58.7%	-60.7%	-35.0%	-47.5%	-47.5%
Overall	**-31.0%**	**-41.2%**	**-43.3%**	**-41.3%**	**-53.4%**	**-53.9%**	**-48.4%**	**-66.6%**	**-69.1%**	**-40.7%**	**-51.9%**	**-52.9%**

**Figure 48: Johnny 720p video conferencing sequence under MTSI test conditions and MTSI_1
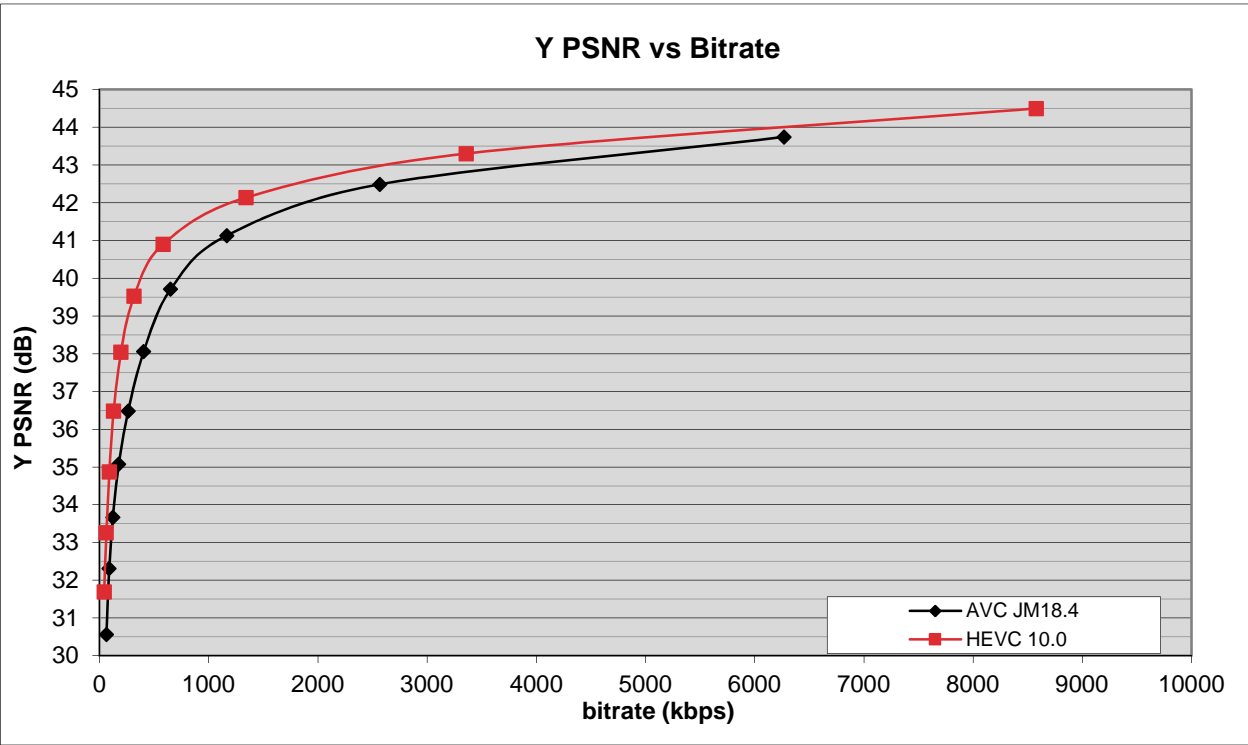prediction structure and AVC High Profile**

Figure 49: FourPeople 720p video conferencing sequence under MTSI test conditions and MTSI_1 prediction structure and AVC High Profile
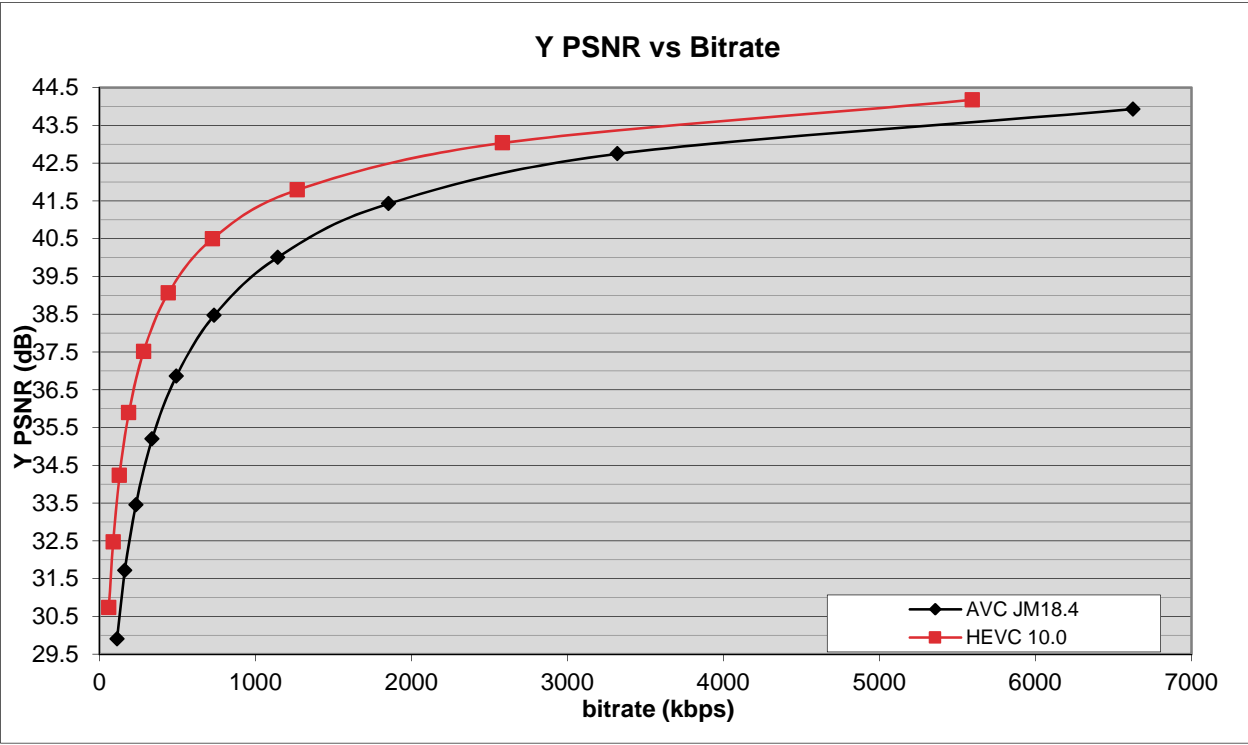

Figure 50: Johnny 720p video conferencing sequence under MTSI test conditions and MTSI_2 prediction structure and AVC High Profile
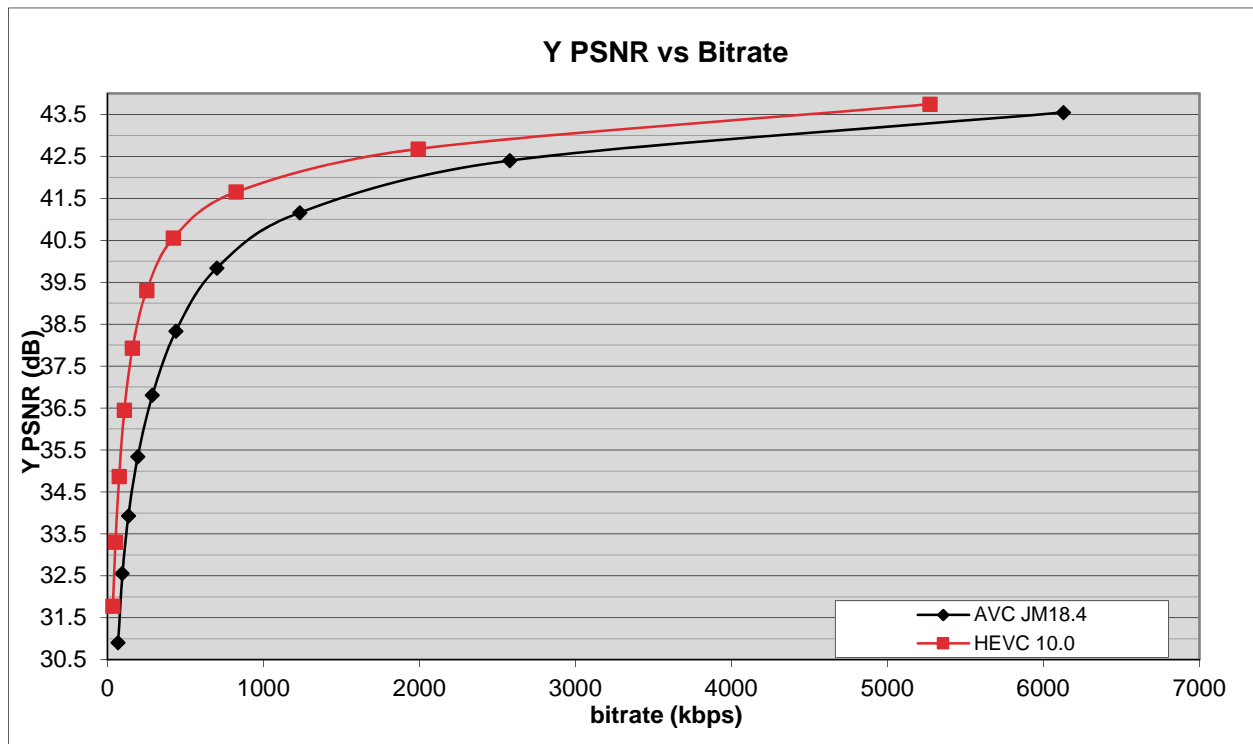
Figure 51: FourPeople 720p video conferencing sequence under MTSI test conditions and MTSI_2 prediction structure and AVC High Profile

8   Test results for image coding

This test evaluates the performance of H.265 (HEVC) for still picture coding. The tests were run by taking into account the guidelines described in [16].

The details of the test are described as follows:

- The first picture of each test sequence was coded with H.265 (HEVC) and JPEG at three different quality levels (40 dB, 36 dB and 32 dB).

- The sizes of the resulting files are then compared.

For JPEG ImageMagick software is used that utilizes the IJG enc–der with -optimize option so that the optimized Huffman tables are used (as suggested by JPEG). The results are summarized below:

Table 33: Results comparing H.265 (HEVC) with JPEG for image coding

	HEVC		JPEG		
	Size	Y-PSNR (dB)	Size	Y-PSNR (dB)	JPEG uses x times more bits
Average (high quality)	63386.39	39.99	125212.86	40.03	1.98
Average (medium quality)	30342.32	36.06	65373.07	36.06	2.15
Average (low quality)	13612.93	31.98	32167.57	32.00	2.36
Average	35780.55	36.01	74251.17	36.03	2.08

As shown above, H.265 (HEVC) brings significant gains over JPEG at various quality levels. The results are consistent with earlier studies, such as gains are higher at lower quality levels.

9 Conclusions

9.1 H.265 (HEVC) for video coding

9.1.1 3GP-DASH

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the 3GP-DASH service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in 3GP-DASH.

9.1.2 PSS

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the PSS service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in PSS.

9.1.3 MBMS

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the MBMS service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in MBMS.

9.1.4 MMS

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the MMS service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in MMS.

9.1.5 MTSI

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the MTSI service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in MTSI.

9.1.6 IMS Messaging and Presence

From the test results summarized in clause 7, it can be seen that H.265 (HEVC) significantly outperforms H.264 (AVC) for the IMS Messaging and Presence service. Therefore, normative specification was agreed to be started to specify the support of H.265 (HEVC) for video coding in IMS Messaging and Presence.

9.2 H.265 (HEVC) for image coding

Although the H.265 (HEVC) still picture profile outperforms the current 3GPP still picture format, the file format support for handling still pictures coded with H.265 (HEVC) is not finalized yet (expected by mid-2014).

Moreover, the H.265 (HEVC) extensions currently under development by the JCT-VC will allow optimizing and extending the use-cases associated with the use of H.265 (HEVC) Still Picture. For example the scalable extension might add quality and spatial scalability features, and the range extension might add high bit-depth support, more efficient coding of screen content, improved lossless coding capability and 4:4:4 coding capability.

Also, thanks to advances in camera hardware, additional use cases related to the ability from UEs to capture image sequences (also known as image bursts) would require further study within the 3GPP context where the H.265 (HEVC) Still Picture profile could be a candidate solution.

Because of the above mentioned reasons it would be desirable to wait until these projects are completed or close to be completed and then consider the most proper support for H.265 (HEVC) for still image coding. It is already identified that these conditions will not be met by the Release 12 timeline.

Annex A:
Change history

Change history							
Date	**TSG #**	**TSG Doc.**	**CR**	**Rev**	**Subject/Comment**	**Old**	**New**
2013-12	62	SP-130570			Presented for information		1.0.0
2014-06	64	SP-140210			Presented for approval	1.0.0	2.0.0
2014-06	64				Approved for Release 12	2.0.0	12.0.0
2015-12	70	SP-150640	0001		Correction to subjective test smartphone figures	12.0.0	12.1.0
2015-12	70				Version for Release 13	12.1.0	13.0.0

Change history							
Date	**Meeting**	**TDoc**	**CR**	**Rev**	**Cat**	**Subject/Comment**	**New version**
2017-03	75					Version for Release 14	14.0.0
2018-06	80					Version for Release 15	15.0.0

History

Document history		
V15.0.0	July 2018	Publication